Replacement Hiring *

Sushant Acharya
FRB New York

Shu Lin Wee
Carnegie Mellon University

May 6, 2017

Abstract

The share of replacement hires has risen over time from 33% in the early 1990s to about 41% in 2015 without a corresponding rise in the share of quits and vacancy posting. We build a model that reconciles these facts and use it to uncover the factors behind the rise in replacement hiring. In our model, matched firms can continually meet new applicants and replace existing workers without having to post a new vacancy or experience a quit. Replacement hires, by allowing firms to replace less productive workers, generate private productivity gains but socially inefficient outcomes as firms fail to internalize the lost value when current workers are released into unemployment. Overall, our calibrated model suggests that an increase in the ratio of old to new vacancies and an increase in firms' outside options are key factors that promoted the increase in the share of replacement hires. Furthermore, we find that welfare in our calibrated economy is 2% lower in terms of output and unemployment is 50% higher relative to the efficient benchmark.

Keywords: Replacement hiring, Unemployment, Efficiency, Labor market churn
JEL Codes: E32, J63, J64

*The authors would like to thank Isabel Cairo, Keshav Dogra, John Haltiwanger and Laura Pilossoph for useful comments and suggestions. 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

Many models of labor search feature firms creating vacancies to attract new job applicants. Some of these job applicants join the firm as new hires while a significant fraction of these job-seekers join the firm as replacement hires. About 36% of all hires are replacement hires. The fraction of total hires that are replacement hires have increased from about 33% in the early 1990s to a high of about 41% in 2015. Importantly, while most models typically treat replacement hiring as hires conducted by firms to replace workers who have quit, we find that the recent upward trend in the share of replacement hires have not been accompanied by a rising share of quits as a fraction of total separations. Rather, the share of separations that are quits are highly procyclical, and do not feature any long term upward trend.[1]

The lack of upward trend in the share of separations that are quits suggests that there is another reason why firms conduct replacement hiring. Intuitively, firms may choose to replace a worker whenever they meet a better job applicant. This can occur while a firm currently has a filled position. Under standard models of labor search, firms seeking to replace their current workers with better applicants have to post more vacancies to attract workers. There, however, has not been a coincident increase in the vacancy posting rate alongside the rise in the share of replacement hires. As we show in Section 2, the lack of upward trend in the share of separations that are quits and in the vacancy posting rates suggests that current labor search models are ill-equiped to address the prevalence of replacement hiring in the data or study the forces behind its upward trend.

In this paper, we provide a new framework of on-the-job search by firms that reconciles how replacement hiring can occur without the incidence of quits or firms having to create new vacancies for each replacement hire. Given our framework, we explore the implications replacement hires have on the aggregate labor market outcomes in the economy. It should be noted that in a world where firms re-hire new applicants for the same position that was not vacated by a quit, replacement hiring only generates gross flows in and out of unemployment but does not do anything to alter the size of the unemployment pool. As replacement hiring forms a greater share of total hiring in the economy, this has implications for the unemployment rate, the value of unemployment and wages in the economy. As the ability of firms to substitute out current workers for better matches with new applicants increases, this in turn increases the volatility of income for workers and reduces their job security.

In our model, search is random and firms pay a fixed cost to create a new vacancy. Vacancies, however, do not expire instantaneously. Rather, they are long-lived. While a vacancy remains active and unexpired, it can continue to attract workers each period. Firms with unfilled and unexpired vacancies only hire workers who are deemed to be a satisfactory match. Matched firms

[1] See Section 2 for details on the empirics.

with unexpired vacancies can continue to meet new applicants but will only hire a new job-seeker if she proves to be a better match than the firm's current worker. Should the matched firm with an unexpired vacancy meet and choose to hire the new applicant, it replaces and releases its current worker into unemployment. This is the sense in which replacement hiring occurs in our model.

While seemingly straightforward, our model introduces features that are typically absent from the standard labor search model. First, the fact that firms pay a fixed cost to create a new vacancy and the fact that vacancies do not expire after one period imply that there is a positive option value to holding a vacancy. This positive option value to holding a vacancy in turn raises a firm's threat point at the stage of bargaining and allows the matched firm to offer lower wages each period relative to a matched firm who has an expired vacancy. Our model then suggests that as the share of hires shifts towards replacement hiring and as the proportion of matched firm-worker pairs with unexpired vacancies increases, the average wage in the economy declines. This in turn leads to a lower labor share even as labor productivity increases as matched firms with unexpired vacancies are able to match with better and better job applicants over time. In Section 4, we verify that average earnings in the data are indeed declining in the composition of matched firm-worker pairs with unexpired vacancies.

Second, our model also has implications about the efficiency of the decentralized economy under Nash Bargaining. Firms post new vacancies in our model until the value of creating new vacancies is equal to its fixed cost. Once a vacancy has been created and so long as it remains unexpired, firms can continue to meet and hire better applicants over time. When deciding whether to hire a new applicant to replace their current worker, firms, in calculating their private gain of re-matching with a new worker, fail to take into account the loss their current worker receives when he moves back into unemployment. The currently matched firm and new worker only consider the gross gain achieved in a replacement hire, and fail to internalize the loss experienced when the firm's existing match is destroyed. In contrast, the social planner takes into account the net gain associated with replacement. As a result, the decentralized economy exhibits too few new jobs created and the share of matched firms with unexpired vacancies relative to unfilled vacancies is too high. We further show that there is no possible split of the Nash-bargained surplus that achieves the social planner's outcome. Instead our model suggests that the introduction of firing costs or employment protections enable the matched firm with an unexpired vacancy to internalize the cost of replacing a current worker, and brings the decentralized economy closer to the social planner's solution.

Third, our model also presents vacancies as a stock as opposed to a flow measure. This representation of vacancies better accords with how the data on job openings is collected. The Bureau of Labor Statistics (BLS) which conducts the monthly Job Openings and Labor Turnover Survey (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"[2]. Standard DMP models of labor search treat vacancies as a flow measure. Our model shows that reformulating the search model to treat vacancies as a stock allows us to address how replacement hires can occur without a corresponding increase in new vacancies. Further, our model decomposes vacancies into two types: filled and unfilled vacancies, and allows us to examine how the changing composition of vacancies can account for the long run decline in match efficiency. In our model, match efficiency is determined by the acceptance rate of firms. Since matched firms with unexpired vacancies only accept new applicants who are better than their current worker, match efficiency is falling whenever the composition of vacancies tilts towards filled, unexpired vacancies.

In our quantitative exercise, we examine the factors that are potentially responsible for the rise in the share of replacement hires. To do this, we calibrate our model to two separate periods in the US labor market. Our benchmark economy is calibrated to match moments for the period 1951:2007. We also calibrate our model to match moments for the period post 2007. We then run the following exercise where we sequentially perturb each of the parameters in the benchmark economy to its post 2007 value. The model suggests that a decline in the expiration rates of vacancies and a rise in firms' outside options are key factors promoting the rise in the share of replacement hires. We also find that the share of replacement hires would have been much higher if the labor market match efficiency had remained at its 1951-2007 average value. The change in labor market match efficiency represents a counteracting force that helped to tamper the rise in the share of replacement hires.

Our paper relates to recent research that has looked at replacement hiring. In a paper closely related to ours, Elsby et al. (2016) also examine replacement hiring and focus on how a quit by a worker necessitates a firm to create a new vacancy in order to re-fill the position. While their paper is concerned with how replacement hiring generates volatility in vacancy posting, we instead focus on the long term trends in replacement hiring and look at a separate reason for why firms conduct replacement hiring even when their positions are currently filled. Our paper is complementary to Elsby et al. (2016) in that we concentrate on the implications of on-the-job search by firms while they focus on the implications of on-the-job search by workers. Fujita and Ramey (2007) also consider a model where firms pay a fixed cost to create job positions which are not destroyed immediately upon worker separation. While they do not focus on replacement hiring and instead examine the sluggish response of the vacancy stock to aggregate shocks, it should be noted that firms with unexpired job positions in their model only re-hire new workers when they are separated from their current worker. By addressing replacement hiring directly, we address how the share of replacement hires can rise without any corresponding increase in the share of separations that are quits. Finally, Menzio and Moen (2010) consider an economy where firms would like to insure workers against income fluctuations but cannot commit to not replacing

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

current workers in a downturn with cheaper new hires. They show that the efficient wage contract is one which exhibits wage rigidity. In contrast, we focus on the efficiency properties of the search model with replacement hiring in steady state when workers can differ in match quality. We find that even in steady state, the decentralized economy under a Nash Bargaining protocol cannot achieve the efficient outcome even when the bargaining weight of firms is set to the elasticity of the matching function with respect to vacancies.

Second, our paper is related to the recent literature on the decline in the labor share. Karabarbounis and Neiman (2013) document that the labor share has declined across countries and argue that improvements in capital producing sectors are the key drivers behind the decline in the labor share as firms move towards using capital intensively in production in place of labor. 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 on the labor share and show how the composition of matched firms affects the wages offered in the economy. Specifically, we demonstrate that as the proportion of firms with unexpired vacancies rises, the labor share falls as more firms have higher outside options.

Lastly, our paper also relates to the recent literature on phantom vacancies. Cheron and Decreuse (2016) argue that phantoms are vacancies that have already found a match and as such cannot generate any more new hires. Instead matching with phantoms represents a loss of time for the unemployed worker as the firm has already filled its position. Our paper views this through an alternative lens: matched firms with unexpired vacancies can still generate new hires. For an unemployed job applicant, however, the probability of being hired by a matched firm with an unexpired vacancy is low relative to the probability of being hired by a firm with an unfilled, unexpired vacancy. In the former case, the unemployed worker must prove to be a better match than the firm's current worker for her to be hired. In the latter case, the unemployed worker must prove to be a satisfactory worker and be above an unfilled vacancy's threshold for hiring. Because matched firms with unexpired vacancies can continue to form new matches in our model, our efficiency results also differ from those identified by Albrecht and Vroman (2015). Albrecht and Vroman (2015) show that in a world with phantom vacancies, firms and workers fail to internalize the externality that whenever they match, they introduce a phantom into the market. In our case, matched firms and new job applicants fail to internalize the loss of value a current worker receives when he is fired and replaced by a new job applicant.

The rest of this paper is organized as follows. Section 2 discusses our empirical findings and shows the recent trends in the share of replacement hires, the share of separations that are quits and the vacancy posting rate. Section 3 introduces the model. Section 4 examines the implications of our model. In section 5, we examine which factors were most important for explaining the rest in the share of replacement hiring. Section 6 explores the efficiency implications of our model while section 7 concludes.

2 Data

Building on the underlying Longitudinal Employer Household Dynamics (LEHD) linked employee-employer database, the Quarterly Workforce Indicators (QWI) provides information on labor market outcomes by state, industry, worker characteristics, employer age and size. In particular, the QWI provides information on the number of total hires at a firm, the total number of separations, the number of job gains and losses as well as earnings. The QWI defines job gains at a firm as the change in employment within a quarter, or in other words a job gain at a firm is equal to the difference between end-of-quarter employment in period t and beginning-of-quarter employment in period t, i.e.

$$\text{Job Gain/Loss} = Emp_t^{end} - Emp_t^{beginning}$$

In contrast, the number of 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 total number of hires measures the gross inflows into a firm, while the measure of job gains measures the net employment change at the firm. Replacement hires are defined as the share of hires that continue into the next quarter in excess of job gains at a firm. The QWI provides a measure of the amount of replacement hires at a state and at each industry level. For our purposes, we focus on nation-wide outcomes. The QWI provides information on

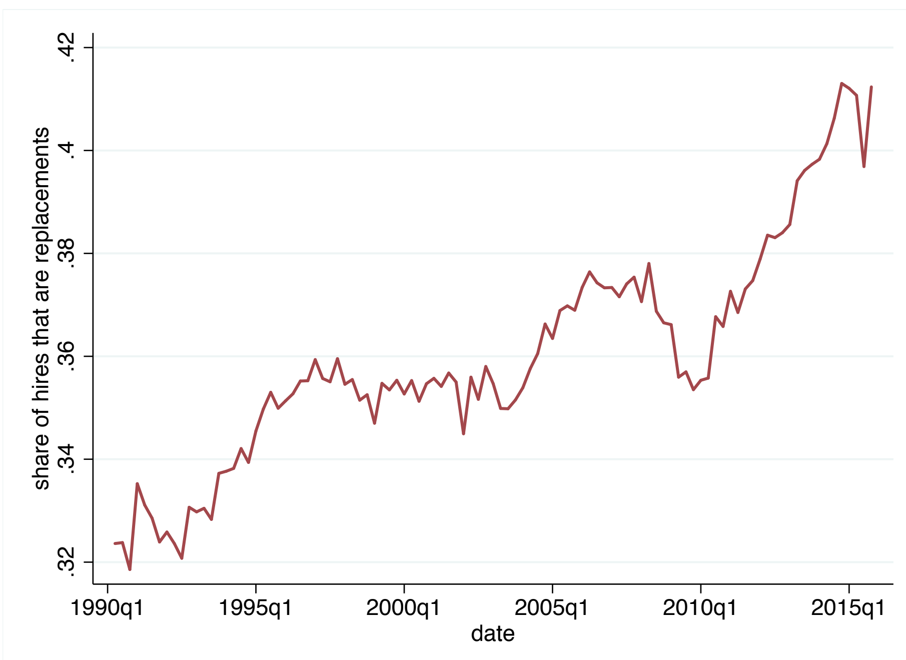

Figure 1: Share of Replacement hires over time

total private hires and total private replacement hires for the entire US economy. We calculate the share of replacement hires as the fraction of total hires that are replacement hires. Figure 1 shows that the share of replacement hires has been trending upwards since 1993. While the rapid uptake in the share of replacement hiring has been observed in the most recent labor market recovery, the upward trend in replacement hiring preceded the Great Recession and can be traced as far back as 1993, the earliest with which we have data on from the QWI. 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* is declining over time. Rather, Figure 1 merely points to the fact that replacement hiring has become a more predominant share of total hiring over time.

There are two reasons why replacement hiring may occur. Firstly, replacement hires can occur whenever a firm seeks to re-fill a position that became vacant due to a voluntary quit. If the primary reason for the occurrence of replacement hiring is to replace workers that have quit, one would expect an upward trend in the share of separations that are quits to coincide with the increase in the share of hires that are replacement hires. As the QWI does not distinguish whether separations that occurred were due to quits or layoffs, we use information from the Jobs Openings and Labor Turnover Survey (JOLTS) and plot quits as a share of total separations. Figure 2 shows the share of separations that are quits over time. Again, it is important to note that the quits rate reported in JOLTS typically measures the quits as a fraction of (average) total employed and is separate from our measure of the fraction of *separations* that are quits. Figure 2 reaffirms previous findings that quits are pro-cyclical and shows that the fraction of total separations that are quits observes no long-run upward trend.

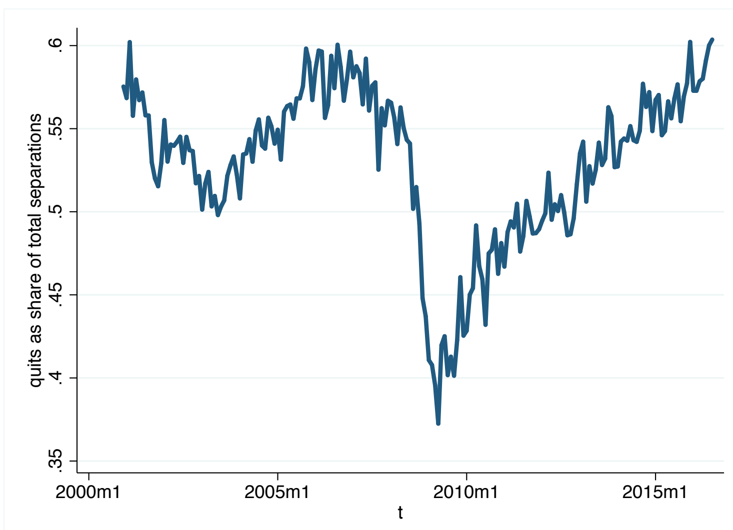

Figure 2: Share of Quits over time

To compare the two time-series, we aggregate the monthly JOLTS data to the quarterly level and normalize the share of quits and replacement hires to be one for 2001q1. Figure 3 shows that while quits (dashed red line) are largely cyclical and can fall below its 2001q1 level, the share of replacement hires has largely kept above one since 2001q1 and has grown by about 15 percent since 2001q1 (solid blue line). Using data from the Job-to-Job Flows (J2J) database which aggregates data on worker flows across main jobs from the LEHD, Figure 3 also plots the separations that occur from job-to-job transitions as a fraction of total separations (black dotted line).[3] Notably, the share of separations from job-to-job flows closely mimick the share of separations that are quits in the JOLTS data, highlighting that quits alone are not sufficient to explain the trends in the share of replacement hiring.

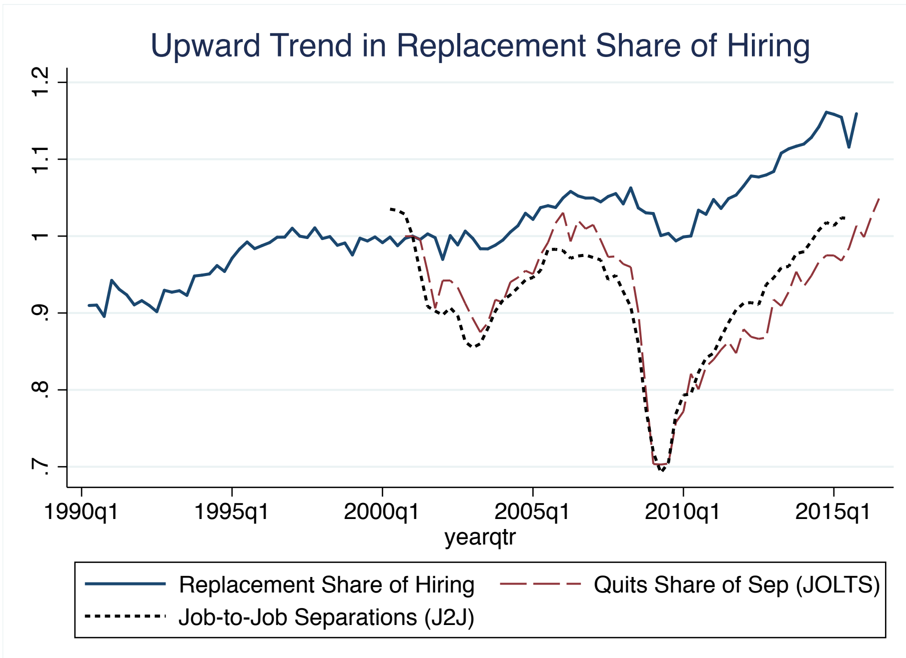

Figure 3: Normalized Share of Quits vs. Replacement Hires over time

The second way a replacement hire could occur is if a firm with a currently filled position meets a better match and chooses to replace his current worker with the new worker. Using the JOLTS micro-data, Elsby et al. (2016) focus on firms with zero net employment change and measure the cumulative hires rate (solid blue line) and cumulative quits rate (dashed-red line) at such firms. While quits do affect the amount of replacement hiring, Figure 4 from Elsby et al. (2016) reveals that a sizable wedge exists between the cumulative hires rate and cumulative quits rates (plus other separations), suggesting that a significant portion of replacement hiring also occurs alongside the event of a layoff.

There are two ways a firm could meet a worker while currently being matched with an employee. Firstly, the firm could create a new vacancy in hopes of attracting a better match to

[3] The J2J database focuses on separations from a main job and not all separations that occur. A job is classified as a main job if it serves as the primary source of income for the worker. A job-to-job separation is recorded when an individual transitions from one job to another with little-to-no observed nonemployment between jobs.

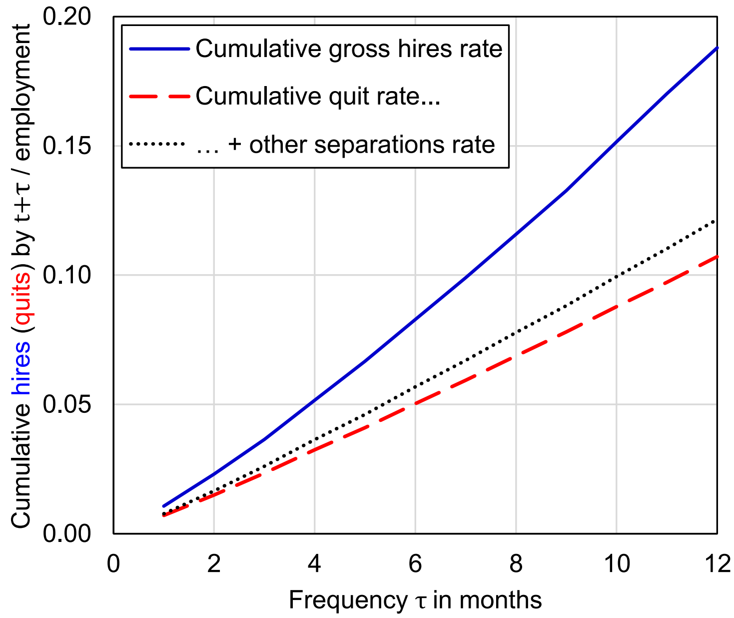

Figure 4: Both Layoffs And Quits Affect Replacement Hiring
Source: Elsby et al. (2016)

replace his current worker. In this case, one would expect a rising trend in vacancy posting rates alongside the rise in the share of replacement hiring. Figure 5 presents plots of the Composite Help-Wanted Index (HWI) constructed by Barnichon (2010) (blue dashed line), the JOLTS vacancy posting rate[4] (green dotted line) and the share of replacement hires (red solid line) normalized to 1 at 2001q1. To compute this figure, we aggregated the monthly Composite HWI and JOLTS vacancy posting rate to a quarterly level. Figure 5 reflects that vacancy posting is highly cyclical and does not exhibit an upward long-term trend.

Given that new vacancy posting has not observed an upward trend while the share of replacement hiring has increased over time, an alternative way firms could attract new applicants without posting new vacancies is if vacancies are long-lived. In this case, a firm currently matched to a worker continues to have the opportunity to meet new applicants so long as he has an unexpired vacancy. The firm with the unexpired vacancy could replace his worker whenever he finds a better match. Under this lens, the stock of vacancies are composed of two types: unfilled vacancies and filled vacancies which have not expired. Our measure of a filled, unexpired vacancy is not inconsistent with 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

[4]The vacancy posting rate is measured as the share of total jobs in the economy that are a vacancy, i.e. $\frac{V}{V+E}$ where V is the measure of vacancies and E measures occupied jobs by the total employed.

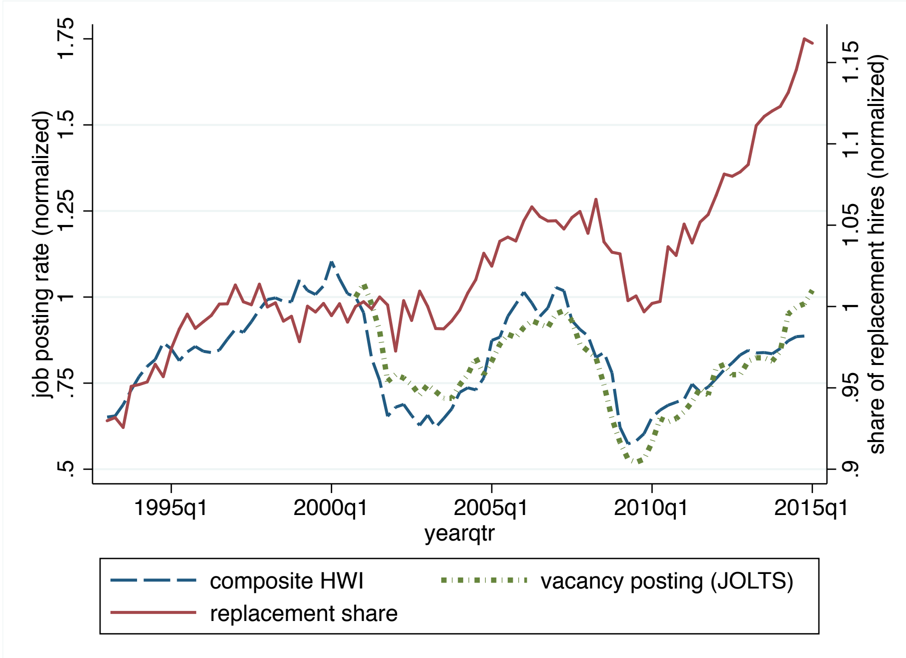

Figure 5: Normalized Vacancy Posting and Replacement Hires over Time

fill a position". [5]

We next examine if the rise in the share of replacement hires is driven by sectoral changes. To do this, we first calculate the share of replacement hires at the 2 digits NAICS industry level, and weight the share of replacement hires at each industry level by its contribution to GDP. Because the industry GDP data is available from the Bureau of Economic Analysis (BEA) at an annual frequency and since the information from the QWI is released at a quarter frequency, we first sum up the total hires and replacement hires within each industry for the entire year, and then calculate the share of replacement hires in a particular industry i as total replacement hires in that industry i divided by total hires in i. Denote rr_{it} as the share of replacement hires in industry i at year t. Then, focusing on the private sector only, our weighted sum of the share of replacement hires is given by:

$$rr_t^{weighted} = \sum_i \frac{VA_{it}}{GDP_{it}^{private}} rr_{it}$$

where $\frac{VA_{it}}{GDP_{it}^{private}}$ is the value-added share of GDP of the private sector from industry i at year t.

[5]Technically, our measure of a filled, unexpired vacancy falls under the JOLTS definition of a job-opening. However, it is far from clear if firms interpret the first condition - 'a position exists' - to mean only unfilled positions. In that case, our model has implications for why 41% of hires occur at establishments with no reported vacancies. Davis et al. (2013) find that accounting for time-aggregation can only partially account for the hires at establishments with no reported vacancies. In our model, we can relate hires at establishments with no reported new vacancies with the measure of replacement hires that occur at matched firms with unexpired vacancies.

(a) Between vs. Within Industry Changes

(b) Between vs. Within Firm Size Changes

Figure 6: Increase in Replacement Share of Hiring Largely Within Firm Size and Industry

We check if compositional changes in sectors could have be the primary factor behind the rise in the share of replacement hires by holding fixed the share of replacement hires in each industry to its 1993 level and allow only each industry's contribution to private GDP to change over time. We refer to this rise in replacement share as being due to between-sector changes and denote this as $rr_t^{between}$ which is formally defined below as:

$$rr_t^{between} = \sum_i \frac{VA_{it}}{GDP_{it}^{private}} rr_{i0}$$

We also plot how the replacement share would have changed over time by holding fixed each industry's contribution to private GDP at its 1993 level and only allow the replacement share within each industry to change over time. We call this the within-sector replacement share and define this formally as:

$$rr_t^{within} = \sum_i \frac{VA_{i0}}{GDP_{i0}^{private}} rr_{it}$$

Figure 6a shows how much of the rise in the share of replacement hires is driven by between sector changes and by within sector changes. Almost all of the increase in the share of replacement hires is driven by within industry changes as opposed to compositional shifts across industries. As such, we conclude that most of change in the share of replacement hiring is coming from within industries as opposed to a compositional shift across industries.

We also verify if the rise in the share of replacement hires is coming from compositional changes amongst firms. Suppose production is given by a decreasing returns to scale production function. This would then imply that larger firms would derive a smaller marginal benefit from replacing a current worker with a new worker who is a better match. If the composition of firms are shifting towards smaller firms, then one might expected part of the rise in the share of replacement hires to be driven by changing composition of firms. The QWI offers five different classification of firm sizes, 0-19 employees, 20-49, 50-249, 259-499, and 500 or more employees. We follow these size definitions and plot how the share of replacement hires would evolve over time if we kept firm sizes fixed to their 1993 level (between) or if we kept replacement shares within each firm size category fixed to their 1993 level (within). Figure 6b plots the replacement share for between firm size changes and within firm size changes. Again, we find that most of the rise in the share of replacement hires can be explained by the within dimension.

Overall, our data exercise suggests that the standard DMP model with on-the-job search by workers may not be sufficient in explaining the rise in the share of replacement hires. We also find that compositional changes between firms of different sizes or of industries are less important in explaining the rise in the share of replacement hires. These findings indicate that the standard labor search model may be missing an important feature which is 'on-the-job' search by firms with unexpired vacancies. We outline in our model section how replacement hiring can occur 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. The production technology of each firm can be written as $F(z, x) = zx$ where z is the level of aggregate productivity which is constant across time and common across all firms. x is the match quality that is drawn at the time of meeting and remains constant for the duration of the match. When a worker and a firm meet, they draw a match-productivity x from a time invariant distribution $\Pi(x)$ on the support $[\underline{x}, \overline{x}]$.[6]

Labor market Search is random. We define a job as a firm-worker pair. A firm that decides to enter the market must incur a fixed cost χ to post a vacancy. Importantly, unlike the standard DMP setup, in our model, unfilled vacancies do not expire instantly. Instead, vacancies expire at a rate $\delta > 0$. Importantly, a firm with an *unexpired* vacancy has the ability to meet and accept new applicants *even if* they are currently matched with a worker. This allows matched firms with unexpired vacancies to replace current workers without having to post a new vacancy or experience a separation. If the matched firm chooses to replace its current worker with the new

[6] We do not restrict support to be bounded. In fact, in our calibrated model, we assume that x is a log-normal distribution and hence the support id unbounded above.

job applicant, it releases the current worker into unemployment. Notice that this is akin to a model in which firms can "search on the job". While the set of vacancies available for matching consist of unfilled unexpired vacancies and currently matched unexpired vacancies, replacement hiring can only occur amongst the latter.

Accordingly, a firm with an expired vacancy cannot meet any new applicants. The expiration of a vacancy, however, is not equivalent to the separation of a matched firm-worker pair. An expiration merely destroys a firm's ability to meet new job applicants. For simplicity, we assume that once a vacancy has expired, a firm cannot re-list the vacancy as long as it is still attached to a worker. Once separated from the worker, such a firm must incur the fixed vacancy posting cost again to post another vacancy. This is not true for a firm who has an unexpired vacancy and has experienced a separation. Such a firm has the continued ability to meet new applicants without incurring any additional costs as long as its vacancy remains unexpired.

Separations into unemployment can be both endogenous and exogenous. In addition to the endogenous separations that occur when a firm conducts replacement hiring, matched firms and workers separate with exogenous rate s regardless of whether they have an unexpired vacancy. All separated workers transition to unemployment but a firm that is separated need not have its vacancy destroyed.

Only unemployed workers can make contact with an unexpired vacancy. Job seekers meet vacancies at a rate p while vacancies meet applicants at a rate q. We assume that the rates p and q depend on the labor market tightness $\theta = (v^m + v^u)/u$ through a matching function which takes total vacancies and the unemployed as its inputs. We denote v^m as the mass of all matched firms with unexpired vacancies, while v^u is the mass of all unexpired unfilled vacancies and u is the mass of unemployed workers. By definition, total vacancies are given by the sum of v^m and v^u. In particular, we assume that the matching-function takes the form:

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

where ξ denotes meeting efficiency. The job-filling rate for firms q is then defined as $M/v = \xi\theta^{-\alpha}$ and the job-filling rate for workers is given by $p = M/u = \xi\theta^{1-\alpha}$.

In order for a meeting to result in a job, 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 create a job as long as the match-specific productivity drawn is above a threshold $\widetilde{x}$, which is determined in equilibrium.[7] However, if the worker instead meets a currently matched firm with an unexpired vacancy, the new match quality x must be at least as large as the match quality between the firm and it's current worker. Thus, even though the probability of an unemployed worker meeting a filled and unfilled vacancy is the same, the probability of a hire is

[7] See Section 3.6 for details.

(weakly) lower for vacancies which are already filled.

3.1 Worker's Problem

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

$$\rho U \;=\; b + p\left\{ \left(\frac{v^u}{v}\right) \int_{\widetilde{x}}^{\overline{x}} \left[W(y,1) - U\right] d\Pi(y) + \left(\frac{v^m}{v}\right) \int_{\widetilde{x}}^{\overline{x}} \int_{\varepsilon}^{\overline{x}} \left[W(y,1) - U\right] d\Pi(y) f^m(\varepsilon) d\varepsilon \right\} \tag{2}$$

where $f^m(\epsilon)$ denotes the density of firms with unexpired vacancies who are currently attached to a worker with match quality ε. $W(y,1)$ is the value of being employed at a firm with an unexpired vacancy, denoted by second argument of 1, and with match quality y.

The discounted value of being unemployed can be decomposed into two terms: b, the flow utility associated with home production and the second term in equation (2) which denotes the expected change in value that the worker enjoys in the event that he transitions to employment in the future. With probability p, an unemployed worker meets an unexpired 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 v^m/v, the unemployed worker encounters a currently matched-firm with an active vacancy. In this case the unemployed worker must draw a x greater than the match quality that the firm shares with its current worker. The second term inside the parenthesis captures the acceptance decision for a current matched firm-worker pair of match quality ε weighted by the probability of meeting a matched firm with match quality ε.

Worker employed at a firm with an expired vacancy The asset value of an employed worker at a firm with an expired vacancy and with match quality x can be written as:

$$\rho W(x,0) = w(x,0) - s\left(W(x,0) - U\right) \tag{3}$$

The asset value can be decomposed into two parts: (i) the wage $w(x,0)$ that the worker earns today and (ii) the change in value the worker receives whenever he is exogenously separated at a rate s and transitions into unemployment. The second argument of both the value-function $W(x,0)$ and the wage $w(x,0)$ indicates that the worker is situated in a firm with an expired vacancy.

Worker employed at a firm with an unexpired vacancy The asset value of an employed worker at a firm with an unexpired vacancy and with match quality x can be written as:

$$\rho W(x,1) = w(x,1) - \Big(s + q\left[1 - \Pi(x)\right]\Big)\left[W(x,1) - U\right] + \delta\left[W(x,0) - W(x,1)\right] \quad (4)$$

where $w(x,1)$ denotes the wages paid to such a worker. As before, the second argument taking a value of 1 indicates that the worker is situated in a firm with an unexpired vacancy. There are three events that such a worker can experience in the future. Firstly, at rate δ, the vacancy of the firm expires and the worker receives a change in value $W(x,0) - W(x,1)$. Secondly, the worker is exogenously displaced into unemployment at rate s and receives a change in value of $U - W(x,1)$. Finally, the firm can meet a new applicant at a rate q and they draw a match quality above x. In this case the current worker is released into unemployment and receives the change in value $U - W(x,1)$.

Notice that for a given match quality x, the value of being employed at a firm with an unexpired vacancy can differ from that of being employed at a firm with an expired vacancy. The worker at the former faces an additional risk of becoming unemployed as his firm has the ability to replace him. In contrast, a worker at the latter firm faces no risk of being replaced since its firm does not have the opportunity to meet new applicants.

3.2 Firm's Problem

Matched firms with expired vacancies Consider a firm that is matched with productivity x but does not have a vacancy. The value of such a firm can be written as:

$$\rho J(x,0) \;=\; zx - w(x,0) + s\left[J(0) - J(x,0)\right] \quad (5)$$

where the '0' refers to the fact that the firm does not have a vacancy. The firm receives its current profit $zx - w(x,0)$ and receives a change of value whenever it separates from its worker. $J(0)$ refers to the value of creating a new vacancy.

New firms/vacancies If a firm decides to post a vacancy, it incurs a cost χ and becomes an unfilled vacancy instantly. Then the value of a vacancy $J(0)$ can be written as:

$$\rho J(0) \;=\; -\chi + J(1) - J(0) \quad (6)$$

where $J(1)$ is the value of an unfilled unexpired vacancy.[8] Equivalently, the above also implies

[8] Note that this is different than the standard DMP specification of flow costs of vacancy posting. Here χ is a one time fixed cost that the firm must pay in order to get the privilege of hiring. This privilege vanishes at a rate δ. Furthermore, note that we do not explicitly assume that on paying the cost, the firm becomes an active

that all new vacancies created are subsumed into the stock of unfilled and unexpired vacancies.

Matched firms with unexpired vacancies The value of such a firm with current match quality x can be written as:

$$\begin{aligned}\rho J(x,1) &= zx - w(x,1) + q\int_{x}^{\overline{x}} \left[J\left(y,1\right) - J\left(x,1\right)\right] d\Pi(y) + \delta\left[J(x,0) - J\left(x,1\right)\right] \\ &\quad + s\left[J(1) - J(x,1)\right] \end{aligned} \tag{7}$$

The firm receives current profits $zx - w(x,1)$ and can undergo three possible events in the future. First, its vacancy expires at a rate δ and it receives a change of value $J(x,0) - J(x,1)$. Second, it may undergo an exogenous separation at a rate s with the associated change in value $J(1) - J(x,1)$. Note that unlike the firm with an expired vacancy, the firm here does not need to create a new vacancy in order to meet new applicants. Finally, a firm may meet a new applicant at rate q and enjoys a change in value $J(y,1) - J(x,1)$ as long as the new match quality y exceeds his current match quality x.

Firms with unfilled and unexpired vacancies The value of an unmatched firm with an an unexpired vacancy can be written as:

$$\rho J(1) = q\int_{\widetilde{x}}^{\overline{x}} \left[J\left(y,1\right) - J\left(1\right)\right] d\Pi(y) + \delta\left[J(0) - J(1)\right] \tag{8}$$

In the current period, since the firm is unmatched, it has zero production and zero current profits. The vacancy might expire at rate δ in which case it must create a new vacancy in order to meet new applicants. At a rate q, the firm meets an unemployed applicant and forms a match as long as the match quality that the pair draws is above the reservation productivity value $\widetilde{x}$.

The gains from matching with a worker with match quality x for a firm with an unexpired vacancy are given by $J(x,1) - J(1)$. Notice that a firm with an unexpired vacancy would never accept a match that yields a gain to matching that is below 0. Thus, the reservation match-productivity $\widetilde{x}$ can be defined as the value of x which satisfies:

$$J(\widetilde{x},1) - J(1) = 0 \tag{9}$$

---

unattached vacancy. This is simply because of continuous time. Suppose for a moment that each period was of length Δ. Then we can write the equation above as:

$$J\left(0\right) = -\chi + q\Delta\mathbb{E}_x[J(x,1) \mid x \geq \tilde{x}] + (1-q\Delta)(1-\rho\Delta)\Big[(1-\delta\Delta)J\left(1\right) + \delta\Delta J\left(0\right)\Big]$$

Taking the limit $\Delta \to 0$, the equation above collapses to (6). The broad intuition is that in discrete time the value functions are written at the production stage and search-and-matching happens after that, so a firm has to wait one period before it can match, so it stays unmatched for the first period.

3.3 Wage Formation

Wages are determined via Nash Bargaining:

$$w(x,i) = \text{argmax}_{w(x,i)} \Big[J(x,i) - J(i)\Big]^{1-\eta} \Big[W(x,i) - U\Big]^{\eta} \text{ for } i = 0,1 \tag{10}$$

where $\eta \in [0,1]$ denotes the bargaining power of a worker. Under this setup, a firm with an unexpired vacancy pays the same wage to a new applicant of match quality x regardless of the firm's history. In other words, a newly hired applicant of match quality x receives the same wage regardless of whether he was hired at a firm with an unfilled, unexpired vacancy or by a matched firm with an unexpired vacancy. Implicitly, we are assuming that whenever a matched firm chooses to hire a new applicant, he releases his current worker into unemployment and there are no recalls. As such, when both the firm and new applicant bargain over wages, the firm's outside option is simply the positive option value of an unfilled and unexpired vacancy.

3.4 Free entry

Under free entry, firms enter the labor market until the value of creating a new vacancy, $J(0)$, is driven down to zero. Equation (6) then implies that that the value of a firm with an unfilled and unexpired vacancy is given by:

$$J(1) = \chi > 0 \tag{11}$$

The above equation pins down q, the rate at which firms with unexpired vacancies meet job applicants, for a given $\widetilde{x}$. Notice that the above equation reinforces the fact that since vacancies do not expire immediately and unexpired vacancies can continually contact applicants, the value of an unfilled vacancy is no longer zero. Thus, a firm with an unexpired vacancy has a positive option value.

3.5 Surplus

The joint payoff to a match is the total surplus shared by the firm and worker relative to continuing to search.[9] Appendix A.2 shows that the surplus for a firm-worker pair with match quality x and who have an expired vacancy is given by:

$$(\rho + s)S(x,0) \quad = \quad zx - \rho U \tag{12}$$

[9] We follow the standard definition of surplus:

$$S(x,i) = J(x,i) - J(i) + W(x,i) - U \text{ for } i \in \{0,1\}$$

where ρU is as defined in equation (2). Since the firm with an expired vacancy has zero option value from creating a new vacancy and since the worker's gain from a match is his benefit from a match less the value of remaining unemployed, we get the standard result that the joint surplus, appropriately discounted, is given by the total benefit of a match, here output, less the worker's outside option, the value of unemployment.

In contrast, the gain a firm with an unexpired vacancy derives from a match is his benefit to matching relative to the value of continuing to search for a better worker, where the latter is encapsulated by the value of an unfilled and unexpired vacancy. In this case, for a firm-worker pair with match quality x, the surplus for such a match is given by:[10]

$$\Big(\rho + \delta + s + q\left[1 - \Pi\left(x\right)\right]\Big) S\left(x,1\right) \;=\; zx + \delta S(x,0) - \rho U - q\left(1-\eta\right)\int_{\tilde{x}}^{x} S\left(y,1\right) d\Pi\left(y\right) \tag{13}$$

Focusing on the RHS of the first line of equation (13), the total benefits from a match include not only (the discounted value of lifetime) output produced but also take into account that at rate δ, the vacancy expires and the match achieves the surplus of a firm-worker pair with an expired vacancy. From these benefits, the surplus nets out the relative value of continuing to search which in this case is given by both the worker's outside option of unemployment *and* the firm's value from continuing to search for a worker. Notice when a firm-worker pair of match quality x is formed, the matched firm with an unexpired vacancy can still meet workers who draw better or worse match quality but is only willing to accept a new applicant who draws match quality greater than x. If the firm had instead chosen not to match, he would have had the additional ability to match with any applicant who drew a match quality between $\widetilde{x}$ and x. This is captured by the term $q\left(1-\eta\right)\int_{\tilde{x}}^{x} S\left(y,1\right) d\Pi\left(y\right)$.

3.6 Closing the Model

The entire model can be summarized by the surplus equations above but in turn these depended critically on $\widetilde{x}$ and θ. We now focus on the key conditions that determine these in equilibrium.

Reservation Match Productivity ($\widetilde{x}$) Recall that equation (9) defined $\widetilde{x}$ as the minimum level of match quality for which a firm with an unexpired, unfilled vacancy would be indifferent between creating a match or continuing to search. In addition, Appendix A.1 shows that in equilibrium, a firm receives $1-\eta$ share of the surplus of a match under Nash-Bargaining.[11] As

[10] See Appendix A.2 for derivation of the expression.

[11] Appendix A.1 shows that for all $x \geq \widetilde{x}$, $J(x,i) - J(i) = (1-\eta)S(x,i)$ for $i \in \{0,1\}$.

a result, we can rewrite equation (9) as:

$$J(\tilde{x}, 1) - J(1) = (1 - \eta) S(\widetilde{x}, 1) = 0 \tag{14}$$

Thus, at match quality $\widetilde{x}$, the surplus of a match between a firm-worker pair is also zero. Consequently, evaluating equation (13) at $x = \widetilde{x}$, we can implicitly derive the reservation match quality $\widetilde{x}$:

$$z\widetilde{x} = \rho U \tag{15}$$

Labor market-tightness Given θ, equation (15) uniquely pins down $\widetilde{x}$. Next, using equations (11) and using the fact that the firm receives $1 - \eta$ share of the match surplus, we can rewrite the free-entry condition (8) to yield the following expression for the job-filling rate q:

$$q = \frac{(\rho + \delta)\chi}{(1 - \eta) \int_{\widetilde{x}}^{\overline{x}} S\left(y, 1\right) d\Pi(y)} \tag{16}$$

Given a matching function, equation (16) pins down equilibrium job-filling rate q and hence, the labor market tightness θ.

3.7 Labor Market Flows

Unemployed The steady state rate of unemployment u can be written as:

$$s(1 - u) = qv^u \left[1 - \Pi\left(\tilde{x}\right)\right] \tag{17}$$

The LHS denotes the inflows into unemployment driven by exogenous separations. The RHS denotes the outflows from unemployment. These occur when an unfilled, unexpired vacancy meets an unemployed worker at rate q and they draw a match quality above the reservation value. While unemployed job seekers also leave the pool of unemployed whenever they meet a matched firm with an unexpired vacancy and draw a match quality that is above the current firm-worker pair's match quality, the firm in choosing to match with this new worker releases its current worker into unemployment. Although this contributes to gross flows in and out of unemployment, it nets out to zero. As such, the steady state unemployment is unaffected by replacement hiring.

A fundamental difference in this model relative to the DMP setup is that vacancies have a backward-looking component. Both unfilled, unexpired vacancies and matched firms with unexpired vacancies represent stocks as opposed to flows.

Matched firms with unexpired vacancies Recall that since all matched firms with unexpired vacancies can continue to attract applicants, they contribute towards total vacancies and therefore, market tightness. The steady state stock of matched vacancies can be expressed as:

$$(s+\delta)v^m = q[1-\Pi(\widetilde{x})]v^u \tag{18}$$

The LHS represents the outflows from the stock of matched vacancies due to exogenous separation or expiration of a vacancy while the RHS represents the inflows. These inflows occur whenever an unfilled, unexpired vacancy meets an applicant and forms a match. In addition to the total stock of matched vacancies, we can also characterize the steady state distribution of matched vacancies across the match-quality dimension. In equilibrium, the density of this distribution can be written as:

$$f^m(x) = \begin{cases} \frac{(s+\delta)[q[1-\Pi(\tilde{x})]+s+\delta]}{1-\Pi(\tilde{x})} \frac{\Pi'(x)}{(q[1-\Pi(x)]+s+\delta)^2} & \text{for } x \in [\widetilde{x}, \overline{x}] \\ 0 & \text{else} \end{cases} \tag{19}$$

It is important to recognize that the distribution of matched vacancies by x is informative about the share of replacement hiring in total hiring. If the distribution of matched firms are skewed towards low values of x, then there is substantial room for such firms to find a better match in the future and conduct replacement hiring. In contrast, if most firm-worker pairs are of very high match quality, then it is very hard for a new applicant to replace an existing worker.

Unfilled and unexpired vacancies Having described the steady state stock of matched vacancies and the unemployed, the final component in the labor market tightness that remains to be described is the stock of unfilled unexpired vacancies. This mass can be described by the equation below:

$$\Big(q\left[1-\Pi\left(\tilde{x}\right)\right]+\delta\Big)v^u = v^{new} + sv^m \tag{20}$$

where v^{new} represents the new vacancies that are created in the current period. The LHS of equation (20) represents the outflows from the stock of unfilled vacancies. This consists of those unfilled vacancies that expire at a rate δ and also includes those unfilled vacancies which successfully convert a meeting with an applicant into a hire. The RHS represents the inflows which come from two sources. First, all newly created vacancies add to the future stock of unfilled, unexpired vacancies. Second, all matched firms with unexpired vacancies which undergo an exogenous separation join the stock of unfilled, unexpired vacancies.

Overall, at any point in time, unemployed workers can only match with existing/old vacancies.

New vacancies only add to the stock of unfilled vacancies in the future.[12] In the standard DMP setup, all vacancies are jump variables but in our model, total vacancies have a backward-looking component and v^{new} is the jump variable. Using equations (17), (18) and (20), the ratio of unfilled old vacancies to new vacancies, and the ratio of old matched vacancies to new vacancies can be written as:

$$\frac{v^u}{v^{new}} = \frac{1}{\delta}\left[\frac{\delta + s}{q[1-\Pi(\widetilde{x})] + \delta + s}\right] \tag{21}$$

$$\frac{v^m}{v^{new}} = \frac{1}{\delta}\left[\frac{q[1-\Pi(\widetilde{x})]}{q[1-\Pi(\widetilde{x})] + \delta + s}\right] \tag{22}$$

The equations above imply that in the model the ratio of new vacancies to old vacancies is given by δ. Ceteris paribus, this suggests that the higher the rate of expiration of existing vacancies, the more new vacancies you would have relative to older vacancies. In a similar fashion, an expression for the unemployment rate can be derived in terms of $v^n ew$:

$$u = 1 - \frac{q\left[1-\Pi\left(\tilde{x}\right)\right]}{s\delta}\left[\frac{\delta + s}{q\left[1-\Pi\left(\tilde{x}\right)\right] + \delta + s}\right] v^{new} \tag{23}$$

Finally, we can derive the mass of new vacancies posted using the free entry condition (16) and the definition of the matching function (1):

$$v^{new} = \frac{\delta\left(\frac{\xi}{q}\right)^{\frac{1}{\alpha}}}{1 + \frac{\delta+s}{s}\left(\frac{\xi}{q}\right)^{\frac{1}{\alpha}}\left[\frac{q[1-\Pi(\widetilde{x})]}{q[1-\Pi(\widetilde{x})+\delta+s}\right]} \tag{24}$$

3.8 Wages and Replacement Hiring

Recall from equation (11) that the firm with the unexpired vacancy has a positive option value. In contrast, a firm with an expired vacancy has 0 option value. These differential option values affect the total gains to matching. Specifically, using equations (12) and (13), one can show that

[12] Notice that v^{new} is not counted as part of the vacancies available for matching today. The total vacancies available for matching at time t are given by :

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

where Δ is the length of one period. $(1-\delta\Delta)(v^u_{t-\Delta} + v^m_{t-\Delta})$ 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 $v^{new}_t\Delta$. As we take the limit $\Delta \to 0$, this term disappears from θ. Thus, in the continuous time limit $\theta_t = (v^u_t + v^m_t)/u_t$. Thus, unemployed workers can only match with existing/old vacancies. New vacancies only add to the stock of unfilled vacancies in the future.

$\forall x \geq \widetilde{x}$, we have:

$$S(x,1) - S(x,0) = -\frac{\frac{q[1-\Pi(x)]z(x-\widetilde{x})}{\rho+s} + (1-\eta)q\int_{\widetilde{x}}^{x} S(y,1)d\Pi(y)}{\rho+s+\delta+q[1-\Pi(x)]} \leq 0$$

where we have equation (15) to substitute out ρU. The inequality is strict for $x > \widetilde{x}$. Intuitively, the positive outside option of firms with unexpired vacancies lowers the total gain to matching. Relative to the surplus of a matched pair with an expired vacancy, the surplus of a pair with an unexpired vacancy is lower because of the potential value lost whenever the firm with an unexpired vacancy does a replacement hire. These differential payoffs to matching are also reflected in the wages that firms with expired vacancies pay relative to an otherwise identical firm with an unexpired vacancy. The Proposition below summarizes the differential wage outcomes:

Proposition 1 (Wages)**.** *Consider a worker-firm pair with match quality x.*

1. *The wage the worker receives from a firm with an expired vacancy is given by:*

$$w(x,0) \quad = \quad z\widetilde{x} + \eta z(x-\widetilde{x}) \tag{25}$$

2. *The wage discount that worker receives from a firm with an unexpired vacancy relative to one with an expired vacancy is given by:*

$$w(x,1) - w(x,0) = -\eta\,(1-\eta)\,q\int_{\tilde{x}}^{x} S\,(y,1)\,d\Pi\,(y) \leq 0 \tag{26}$$

Proof. See Appendix A.3. □

Proposition 1 highlights the fact that there are in fact two dimensions of wage dispersion present in this economy. First, wages differ across jobs with differing match quality. Second, and unique to this model, wages differ across firms with different outside options. Intuitively, the positive outside option for a firm with an unexpired vacancy implies that it has a higher threat point, allowing it to pay lower wages than an identical firm with an expired vacancy. The wage discount in (26) comes precisely from the fact that the firm with an unexpired vacancy has a higher outside option. An implication of Proposition 1 is that an economy with a larger fraction of firms with unexpired vacancies would have lower average wages. This suggests that the upward trend in the share of replacement hires that we observe in U.S. data, which is coincident with slow wage growth might be related.

3.9 Match Efficiency and Replacement Hiring

A rising share of replacement hiring also has implications for match efficiency. In our model, the job-finding rate of an individual is given by:

$$p \underbrace{\left(\frac{v^u}{v} \left[1 - \Pi(\widetilde{x})\right] + \frac{v^m}{v} \int_{\widetilde{x}}^{\overline{x}} \left[1 - \Pi(\epsilon)\right] f^m(\epsilon) d\epsilon \right)}_{\text{acceptance rate}}$$

where p is the rate at which a job-seeker meets a vacancy and the second term in parentheses is the average acceptance rate. In a search model which assumes a constant returns-to-scale matching function, the above acceptance rate would typically be treated as the match efficiency component of the job-finding rate.[13] Proposition 2 details how the share of total vacancies that is composed of matched firms with unexpired vacancies, i.e. $\frac{v^m}{v}$, affects match efficiency in the economy.

Proposition 2 (Match Efficiency)**.** *For a given $\widetilde{x}$ and probability density $f^m(x)$, match efficiency is declining in $\frac{v^m}{v}$*

Proof. See Appendix A.4. □

Intuitively, holding all else constant, the rate at which job-seekers are accepted is declining in the share of matched unexpired vacancies, $\frac{v^m}{v}$. This is because matched firms with unexpired vacancies only accept a job applicant whenever the new match quality drawn surpasses that of the match quality of its current worker. In contrast, unfilled vacancies accept a job-seeker so long as the match quality drawn surpasses the reservation productivity, $\widetilde{x}$. As the replacement share of hiring is positively correlated with the share of matched unexpired vacancies, an implication of an economy with a rising share of replacement hiring is that it should also observe a declining match efficiency.

Given our model's implications for match efficiency and wage dynamics with respect to replacement hiring, we check whether these implied trends hold in the data in Section 4.

4 Testing the Predictions of the Model

Our model provides some testable predictions which we now verify against the data. For this purpose, we use the information in the QWI and The Conference Board Help-Wanted Online (HWOL) dataset to test some of the predictions of the model. The HWOL provides monthly information on the total volume of job advertisements posted online for the periods spanning May 2005 to the present at both the state and the national level. In addition, the HWOL data contains information on the volume of job postings that are new. A job posting is counted as a 'new' job advertisement only in the month that it first appears. Because different states

[13] For example, if the matching function is Cobb-Douglas and has the form $m = Au^{\alpha}v^{1-\alpha}$, then the job-finding rate would be given by $\frac{m}{u} = A\left(\frac{v}{u}\right)^{1-\alpha}$. In this context, $p = \left(\frac{v}{u}\right)^{1-\alpha}$ is the meeting rate, and A is the residual component of hires not explained by total vacancies and total unemployed.

participated in the QWI and started supplying data at different points in time, we focus on cross-sectional variation rather than variation across time when comparing data from the HWOL to data from the QWI.[14]

4.1 Expiration Rates and Replacement Share

Our model provides an exact mapping from the ratio of new to old vacancies to the vacancy expiration rate, δ. In our model, the amount of new vacancies, v^{new}, adds to the stock of unfilled vacancies available for matching in the future while $v^m + v^u$ form the total stock of vacancies today. Thus, the ratio of new to old vacancies in our model is given by

$$\frac{v^{new}}{v^m + v^u} = \delta \tag{27}$$

Thus, while we do not have direct measures of the expiration rate, we can compute the model implied expiration rate by using information on the volume of old and new vacancies. Using HWOL data, we can construct an estimate of average quarterly δ at the state level.[15] Intuitively, lower vacancy expiration rates tilt the composition of vacancies towards matched firms with unexpired vacancies. The ratio of matched firms with unexpired vacancies to total vacancies, $\frac{v^m}{v}$, is given by:

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

While q and $\widetilde{x}$ are endogenous objects and will change whenever δ changes, note that the term $q\left[1 - \Pi(\widetilde{x})\right]$ appears in both the numerator and denominator of $\frac{v^m}{v}$. Thus, no matter the sign and size of the change in $q\left[1 - \Pi(\widetilde{x})\right]$, a decline in δ always makes $\frac{v^m}{v}$ larger. Since replacement hiring is positively correlated with $\frac{v^m}{v}$, our model would predict that the share of replacement hiring is declining in δ since matched firms with unexpired vacancies have more opportunities to re-match with new applicants when δ is low. The left panel of Figure 7 shows how the average share of replacement hiring is falling with the associated ratio of new-to-old vacancies in each state, giving credence to our model's implication that the replacement share of hiring decreases when vacancies expire more quickly and firms have a shorter window to re-match with better applicants.

[14] States like Massachusetts joined the QWI only in 2010 while the earliest data in the QWI dates back to 1990.

[15] As the HWOL is a monthly series, we first sum up the total volume of job ads and the total volumes of new jobs ads within a quarter. We then compute the δ from these quarterly measures. Following this, we calculate the average quarterly δ in each state across the time periods 2005-2015

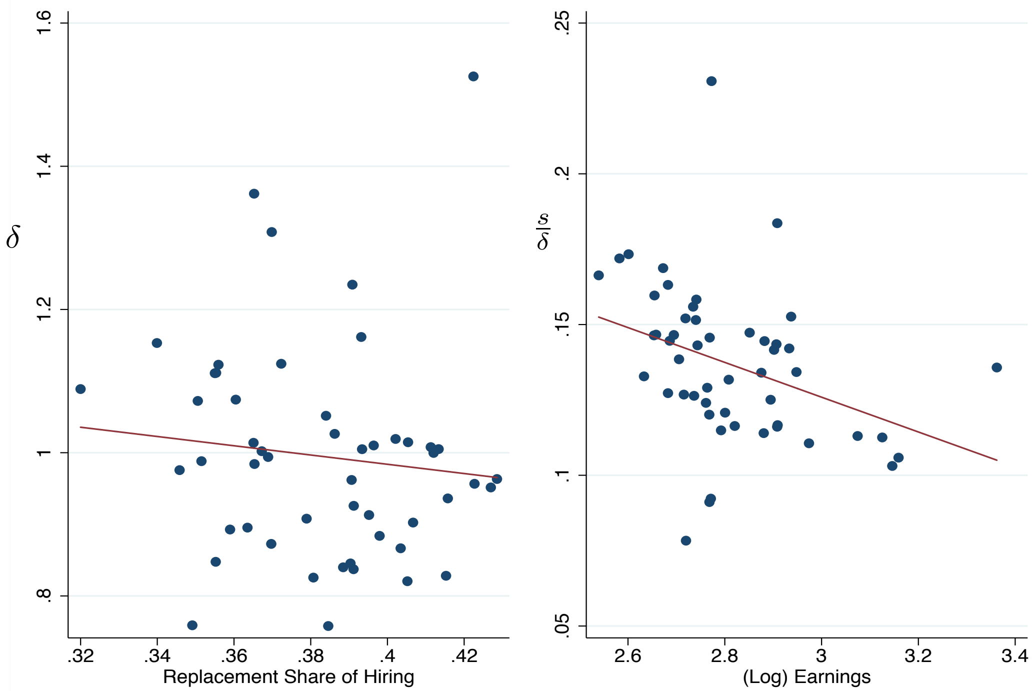

Figure 7: **Left:** Replacement Share declines in δ. **Right:** Earnings decline in s/δ.

4.2 Earnings and Composition of Matched Firms

Our model also predicts that average earnings are declining in the share of firm-worker pairs with unexpired vacancies as firms with unexpired vacancies observe a positive option value of a vacancy, allowing them to offer their workers wage discounts. The average wage rate and earnings in the economy are lower whenever the composition of firms tilts towards matched firm-worker pairs with unexpired vacancies. In our model, the ratio of matched firm-worker pairs with unexpired vacancies to matched firm-worker pairs with expired vacancies, v^m/v^i is directly related to separation and expiration rates[16]:

$$\frac{v^m}{v^i} = \frac{s}{\delta}$$

The QWI provides information on the beginning-of-quarter separation rates as well as information on the average earnings of all employees. We use the beginning-of-quarter separation rates as our measure of s. Combining this measure with our estimate of δ from the HWOL data, our model suggests that states with higher shares of matched firms with unexpired vacancies, i.e. a high s/δ, should have lower wages on average. We deflate the measure of average earnings of all employees by the Consumer Price Index (CPI) and take the natural log of deflated earnings.

[16]In steady state, total inflows into firms with no vacancies must equal total outflows, $\delta v^m = s v^i$.

The right panel of Figure 7 shows that (log) earnings are in fact declining in our estimated measure of v^m/v^i. Regressing (log) deflated earnings on s/δ reveals that a 1 percentage point increase in v^m/v^i lowers earnings of all employees by 2.2%.[17] Overall, our model suggests a different avenue in accounting for the decline in the labor share and provides accounting identities that link the predominance in replacement hiring to the decline in labor share.

4.3 Match Efficiency and the Composition of Vacancies

Finally, our model suggests that match efficiency is declining in the share of matched firms with unexpired vacancies, $\frac{v^m}{v}$. Because $\tilde{x}$ is an endogenous object and $q\left[1-\Pi(\tilde{x})\right]$ describes only the job-filling rate of unfilled vacancies, we do not have direct measures of $\frac{v^m}{v}$ in the data. Nonetheless, for a given $\tilde{x}$, the replacement share of hiring is positively correlated with $\frac{v^m}{v}$. Recall that in our model, only matched firms with unexpired vacancies can conduct replacement hiring, while unfilled vacancies, by construction, cannot give rise to replacement hires.[18] As such, we can look at how match efficiency varies with the replacement share of hiring over time. Using information from JOLTS to construct a measure of match efficiency, i.e. the residual variation in hires which is not explained by data on unemployed job-seekers and vacancies[19], Figure 8 shows how match efficiency is declining over time while the replacement share of hiring is increasing over the same period. These trends suggest that part of the long-run fall in match efficiency could be related to the changing composition of vacancies over time.

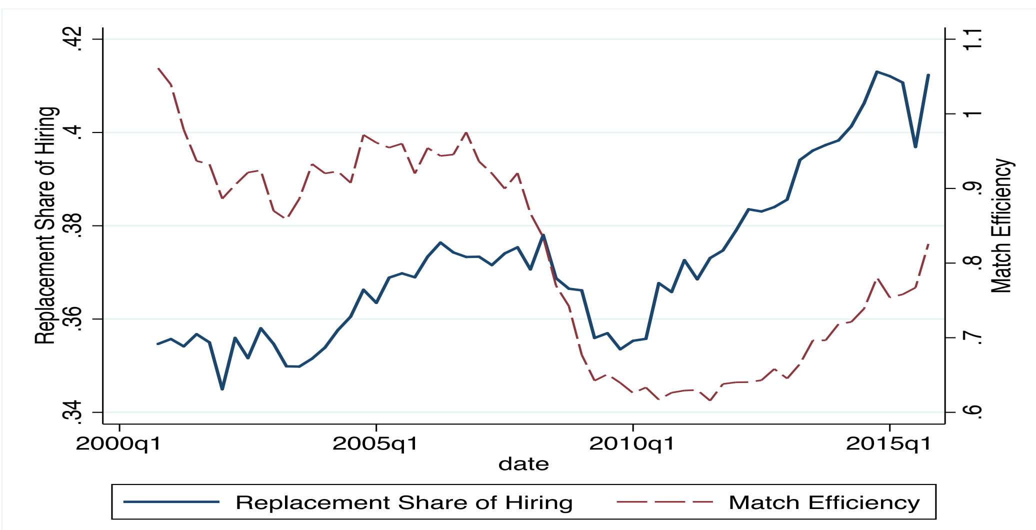

Figure 8: Match Efficiency Declines While Replacement Share Rises Over time

Having examined the implications of our model, we now examine which factors are important

[17]This estimate is statistically significant with a standard error of 0.822.

[18]This is discussed in further detail in Section 5.

[19]Assuming a Cobb-Douglas Matching Function and an elasticity of 0.5 on unemployment, we calculate (log) match efficiency as $\log(A) = \log(hires) - 0.5\log(unemployed) - 0.5\log(vacancies)$.

for generating a rise in replacement hiring share over time.

5 Numerical Exercise

We use our model to back out which factors were most important towards contributing to the rising share in replacement hires. To do this, we consider two separate time periods. We first calibrate our model to moments from 1951 to 2007, and then re-calibrate our model to fit moments from post 2007. Once we have calibrated the model to these two separate time periods, we start from the benchmark economy calibrated to 1951-2007 moments and perturb one parameter at a time to its post 2007 value. This allows us to create counterfactuals and assess which factor was most important towards contributing to the rise in replacement hires.

Before moving to the calibration section, it is important to make clear how we measure replacement hires in our model. We follow the definition of replacement hiring as in the QWI and calculate it as the amount of hires *in excess* of net employment gains. In our model, replacement hiring then corresponds only to hiring by matched firms with unexpired vacancies and not by unfilled vacancies. We use the following two examples to make clear why this is the correct accounting for our model

Example 1 Suppose a matched firm with an unexpired vacancy undergoes a separation today, it becomes an unfilled and unexpired vacancy and observes a net employment change of minus one. In the next period, this same unfilled vacancy meets an unemployed job applicant who draws a match quality above $\widetilde{x}$ and the firm hires the worker. In this case, the next period observed a net change in employment of +1, a gross hire of +1 and a replacement hire of zero in the next period.

Example 2 Suppose instead a matched firm with an unexpired vacancy today meets an unemployed job applicant who surpasses the match quality of his current worker. The matched firm decides to replace his current worker and hire the new job applicant. In this case, in the current period, there was a gross separation of +1, a gross hire of +1, and a net employment gain of zero. In the current period, this also amounted to a replacement hire of +1.

Notice that in the first example, a matched firm with an unexpired vacancy that became an unfilled vacancy which subsequently hired a worker does not fit into the definition of replacement hires. This is because the number of gross hires in the period where the unfilled vacancy matched with a job applicant is exactly equal to the net gain in employment. In contrast, in the second example, the matched firm with an unexpired vacancy who chooses to hire a new applicant and release his current worker into unemployment observes a gross hire of one in excess of his net employment gain of zero. As such, we would record a replacement hire for the latter case.

5.1 Calibration

We set a period in our model to be one quarter. The discount rate ρ is accordingly set to 0.012 to reflect an annual interest rate of about five percent. We normalize the aggregate productivity parameters z to be 1 in steady state and set the bargaining weight of workers, η to be 0.5. We set the elasticity of the matching function with respect to unemployment α to be 0.5 as standard in the literature.

As the empirical literature has typically found that wages are log-normally distributed, we assume that the distribution of match quality, $\Pi(x)$ is given by the log of $N(-\frac{\sigma_x^2}{2}, \sigma_x^2)$. Overall, this gives us five key parameters left to calibrate: 1) the expiration rate of vacancies, δ, 2) the exogenous separation rate s, 3) the unemployment benefit b, 4) the meeting efficiency parameter, ξ and 5) the dispersion in match quality σ_x. The last parameters of the model, the vacancy posting cost, χ, is calculated as the residual to equation (16).

The set of key parameters, $(s, b, \delta, \sigma_x, \xi)$ affect the transition rates of individuals in and out of both employment and unemployment, as well as the amount of replacement hiring. Notice that s, δ and σ_x have implications for the rate at which individuals exit employment, s affects how people are exogenously separated, while the latter two parameters affect the extent of replacement hiring conducted and thus how many current matches are endogenously dissolved. As such, one of the moments we target concern the exit rate of individuals from employment. The parameters b and ξ affect the rate at which individuals are able to transition into employment and consequently the unemployment rate. Thus, we also target the job-finding rate in the economy as well as the unemployment rate. It should be noted that unlike the DMP model, there is a distinction between gross and net flows in this model - replacement hiring adds to gross flows in and out of unemployment but does not affect the size of the unemployment pool on net. As such, 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 share of replacement hires depends crucially on the expiration rate, δ and the matching efficiency, ξ, of the economy. The former affects the longevity of a vacancy and the firm's ability to conduct a replacement hire while the latter affects the rate at which firms are able to contact new applicants for replacement hiring. Finally, the flow value individuals receive while unemployed can be linked to the amount of unemployment insurance. Thus, we jointly calibrate these key parameters, $(s, b, \delta, \sigma_x, \xi)$, by targeting the moments in the data concerning the exit probability of employed individuals, the job finding probability of the unemployed, the share of replacement hires, the unemployment rate and the unemployment insurance.

We conduct two separate exercises. We first calibrate the model to data moments for the period 1951m1 to 2007m12. We then re-calibrate the model to data moments for the period post 2007 and ask how unemployment and wage outcomes would differ if the expiration rate of vacancies in post 2007 was the same as the expiration rate for the period 1950m1 to 2007m12.

For both time periods, we follow Shimer (2005) and target a 40 percent unemployment insurance replacement rate. Using data from the CPS on employment, unemployment and short term unemployment, we find, for the period 1950m1-2007m12, 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.[20]. 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%. We use data from the Quarterly Workforce Indicators (QWI) to calculate the share of replacement hires as a fraction of total hires. Since the QWI has data available from 1993Q1 onwards, we calculate that the average share of replacement hires for the period 1993Q1 to 2007Q4 is about 0.35. For the period post 2007, we find an average unemployment rate of 7.3%, an average exit probability of 0.02, the average job-finding probability of 0.25 and share of replacement hires of about 0.4. The table below summarizes our calibrated parameters:

Our calibration exercise gives us reservation match qualities, $\tilde{x}$, of 0.89 and 0.81 for the periods 1951-2007 and post 2007 respectively. In addition, the value of an unfilled active vacancy, $J(1)$, is given by the vacancy posting cost, χ, which for the period 1951-2007 is given by $\chi = 0.69$ and for the period post-2007 by $\chi = 2.01$. Recall that $J(1) = \chi$. The increase in χ to its post 2007 value is thus equivalent to the value of creating a new job increasing over the two periods. This increase is consistent with stock market valuations of firms over the same time period. As per Hall (2017), the value of a new job confers information on the financial valuation of the firm. In our model, the rise in firm value over time shows up as an increase in χ.

Some predictions of the model While we did not target these moments, we find that our model has implications for the ratio of new to old vacancies. In the post 2007 period, our model suggests this ratio takes a value of 1.15. Using the information on new and total job advertisements from the HWOL data, we find that the ratio of new to old vacancies for the period of post 2007 is about 1.01.[21] While our model suggests a higher share of new to old vacancies, our model is still close to the HWOL statistic. Furthermore, it should be noted that the HWOL does not capture the whole universe of vacancy postings and only reflects the new to old ratio of vacancies posted online. In calibrating our model, we have targeted moments that represent the aggregate outcomes for the entire economy.

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

[21] The HWOL series begins in 2005. Prior to this, the now-discontinued Conference Board Help Wanted Advertising Index did not make a distinction between old and new vacancies.

Table 1: Model Parameters

Fixed Parameters		
Parameter	Description	Value
ρ	discount rate	0.012
z	agg. productivity	1
η	bargaining weight	0.5
α	matching function elasticity	0.5

Calibrated Parameters (1951-2007)			
Parameter	Target	Value	Model
b	40% UI replacement rate	0.3800	0.40
ξ	job finding rate of 1.74	3.6969	1.74
s	exit rate of 0.098	0.0649	0.099
σ_x	unemployment rate of 0.05	0.0630	0.054
δ	share of replacement hires of 0.35	2.9132	0.35

Calibrated Parameters (post 2007)			
Parameter	Target	Value	Model
b	40% UI replacement rate	0.3661	0.40
ξ	job finding rate of 0.863	2.0226	0.837
s	exit rate of 0.061	0.0377	0.064
σ_x	unemployment rate of 0.073	0.0662	0.071
δ	share of replacement hires of 0.40	1.1463	0.41

Our model also features a decline in the labor share between the two time periods. For the period 1951-2007, the computed labor share was 0.92, while for the period post 2007, the computed labor share was 0.87. This five percentage point decline in the labor share is accompanied by a one percentage point increase in the fraction of matched firm-worker pairs that possess an unexpired vacancy, i.e. $\frac{v^m}{1-u}$. These outcomes reflect that as the share of employed individuals working in firms with an unexpired vacancy increases, the average wage declines as more individuals are working for firms who have a higher outside option. Correspondingly, the divide between the mean wage and average labor productivity is greater for the post 2007 period. In the period prior to 2007, the mean wage in our model is about 0.95 while average labor productivity is about 1.03. In the period post 2007, the mean wage in our model is about 0.91 and the average labor productivity is about 1.04.

5.2 Counterfactuals

We now analyze the factors that could have contributed to the rise in the share of replacement hires between 1993-2016. For this exercise, we set as our benchmark economy the model with replacement hires parameterized to match aggregate moments for the period 1951-2007. We then sequentially 'turn on' each of the exogenous parameters of the model to its post 2007 value and see how this affects the key aggregate moments of our model economy. Table 2 summarizes the results of our counterfactual exercise. The first column of Table 2 lists the aggregate outcomes in the benchmark economy parameterized to match moments for the period 1951-2007 while the last column lists the aggregate outcomes for the economy parameterized to match moments for the post 2007 period.

Column 2 of Table 2 shows what happens to the benchmark economy when the dispersion in match quality, σ_x, is raised to its 2007 level. As can be seen from Column 2, our results suggest that the change in the dispersion of match quality is not a primary factor behind the rise in the share of replacement hires. None of the outcomes change if we shift the dispersion in match quality to its post 2007 calibrated value.

One hypothesis for the rising share of replacement hires in the data is that with the rise of online job boards and ease with which companies can post vacancies on their own websites, the expiration rate of vacancies has declined over time. Indeed, moving δ down to its post 2007 calibrated value brings the replacement share close to its post 2007 value. This is perhaps not surprising since intuitively, a lower δ would naturally raise the replacement share of hiring as vacancies now last for longer and firms have more opportunities to conduct replacement hiring. Since vacancies now last for longer, firms can choose to raise their hiring standards since rejecting a worker is less costly today if the vacancy is likely to survive into the future. This raises $\tilde{x}$, which in turn lifts the mean wage and the labor productivity in the economy. The change in δ to its post 2007 value however, predicts too low an unemployment rate for the economy. The lower unemployment rate arises because the value of an unfilled vacancy is pinned down by the vacancy posting cost χ which is still fixed at its pre-2007 value. From Equation (8), a fall in δ with no change in χ, necessitates a decrease in q. Since each vacancy (filled or unfilled) now lasts for longer, there is more competition for each job-seeker and the labor market is tighter. Thus, while the decline in the expiration rate of vacancies is likely to be a factor driving the increase in the share of replacement hires, it predicts a much lower unemployment rate than actually observed for that period.

We next perturb the vacancy posting cost, χ, to its post 2007 value. Increasing χ to its post 2007 calibrated value predicts too high a share of replacement hires in the economy. The high amount of replacement hiring in the economy also shifts the composition of vacancies towards matched vacancies. This, together with the lower reservation match quality leads to depressed mean wages and labor shares far lower than their post 2007 values. Because jobs are now costly

Table 2: Counterfactuals

Variable	pre2007	$\sigma_{x,07}$	$\sigma_{x,07}, \delta_{07}$	$\sigma_{x,07}, \delta_{07}, \chi_{07}$	$\sigma_{x,07}, \delta_{07}, \chi_{07}, s_{07}$	$\sigma_{x,07}, \delta_{07}, \chi_{07}, s_{07}, \xi_{07}$	post2007
$\tilde{x}$	0.89	0.89	0.92	0.77	0.85	0.81	0.81
q	4.64	4.64	2.62	4.27	3.95	2.62	2.59
share of repl. hires	0.35	0.34	0.39	0.48	0.47	0.41	0.41
u	0.054	0.054	0.039	0.084	0.045	0.072	0.071
mean wage	0.95	0.95	0.97	0.87	0.94	0.91	0.91
$y/(1-u)$	1.03	1.04	1.05	1.06	1.06	1.04	1.04
labor share	0.92	0.92	0.92	0.83	0.89	0.87	0.87
mkt output	0.98	0.98	1.01	0.97	1.01	0.97	0.97

to post, the unemployment rate also rises to too high a level. Overall, one can conclude that the fall in δ and the rise in χ to their post 2007 values put strong upward pressure on the share of replacement hires. To arrive at the observed level of replacement hiring for the post 2007 period, there had to be an additional key factor that reduced the amount of replacement hiring.

Column 5 of Table 2 shows that reducing s to its post 2007 value does little to change the share of replacement hires, although the decline in s does have strong impact on reducing the unemployment rate. Since the inflows into unemployment are determined by s, the unemployment is naturally lower when we reduced s and the value of the unemployed increases, raising $\widetilde{x}$, the mean wage and the labor share.

Column 6 of Table 2 reveals that reducing ξ to its post 2007 is the key counteracting factor that brings the share of replacement hires back to its post 2007 value. Since the meeting efficiency is now lower, matched vacancies now face a lower rate of meeting unemployed job applicants, reducing their opportunity to conduct a replacement hire and thus the share of replacement hires drops to its post 2007 level.

Thus, in our counterfactual exercise, we find that the forces that supported the rise in the share of replacement hires stem from a lower expiration rate of vacancies and higher new vacancy creation costs as pinned down by χ. The fall in meeting efficiency over the two periods is key to dampening the rise in replacement hires. With respect to the labor share, the increase in χ to its post 2007 calibrated value is the key factor that acts towards lowering the labor share. Recall that new vacancies form part of the inflows into the stock of unfilled vacancies. By making the cost of new vacancies more expensive, the composition of vacancies shifts strongly toward matched vacancies, and this lowers the average wage and labor share in the economy.

6 Is Replacement Hiring Efficient?

Given the rising trend in the share of replacement hires, a natural question that arises is whether the outcomes in such an economy are socially efficient? Is the amount of churn in the labor market too high or too low and are firms being too selective or lax in their hiring standards? We explore these questions through the lens of a constrained social planner's problem. Since all agents are risk-neutral, the planner seeks to maximize lifetime discounted value of output net of vacancy creation costs by choosing the number of new vacancies and the reservation productivity $\widetilde{x}$. The objective function can be written as:

$$\max \int_0^\infty e^{-\rho t} \left\{ -\chi v_t^{new} + b u_t + z \int_{\underline{x}}^{\overline{x}} \epsilon \mu_t^m (\epsilon) \, d\epsilon + z \int_{\underline{x}}^{\overline{x}} \epsilon \mu_t^i (\epsilon) \, d\epsilon \right\} dt$$

The first term in the parenthesis is the cost associated with posting new vacancies while the second term denotes the value of home-production from the unemployed. Because the universe of matched firms consists of firms with expired and unexpired vacancies, market output produced comes from two sources. The third term represents the output produced by matched firms with unexpired vacancies while the last term represents that produced by firms with expired vacancies. It is also convenient to define the function $\mu^m(x) := f_t^m(x) v_t^m$ where $f_t^m(x)$, as previously defined, is the density of matched firm-worker pairs of match quality x who possess an unexpired vacancy. Thus, $\mu_t^m(x)$ refers to the corresponding mass of firm-worker pairs with match quality x. Similarly, $\mu_t^i(x)$ refers to the mass of firm-worker pairs with match quality x who possess an expired vacancy.

In maximizing the objective function above, the planner must respect the matching technology which are captured by the following constraints:

$$\dot{\mu}^m (x) = \left[q\pi (x) \left(v^u + \int_{\widetilde{x}}^{x} \mu^m (\epsilon) \, d\epsilon \right) - \left(s + \delta + q \left[1 - \Pi (x) \right] \right) \mu^m (x) \right] \forall x \geq \widetilde{x} \tag{28}$$

$$\dot{\mu}^i (x) = \delta \mu^m (x) - s \mu^i (x) \tag{29}$$

$$\dot{v}^u = v^{new} + s \int_{\underline{x}}^{\overline{x}} \mu^m (\epsilon) \, d\epsilon - \left(q \left[1 - \Pi (\widetilde{x}) \right] + \delta \right) v^u \tag{30}$$

$$u = 1 - \int_{\underline{x}}^{\overline{x}} \mu^m (\epsilon) \, d\epsilon - \int_{\underline{x}}^{\overline{x}} \mu^i (\epsilon) \, d\epsilon \tag{31}$$

where $q_t = \xi \left(v_t / u_t \right)^{-\alpha}$. Equation (28) denotes how the mass of firm-worker pairs with match quality x and an unexpired vacancy evolve over time given the choices of the planner. Similarly, equation (29) describes how the mass of firm-worker pairs with match quality x and an expired vacancy evolve over time. Inflows into this stock originate from matched firm worker pairs who previously had an unexpired vacancy which just expired, while outflows comprise of pairs with expired vacancies which experience a exogenous separation. Equation (30) describes the evolution of unfilled vacancies which the planner can affect by choosing v^{new}. Equation (31) simply defines the level of unemployment as the mass of individuals in the economy minus those who are employed. These constraints are the same as those which describe the decentralized economy but outcomes may differ due to differences in the planner's choice of $\widetilde{x}$ and v^{new}.

The solution to the social planner's problem is detailed in Appendix B.1. In a similar fashion to the decentralized model, the efficient outcomes can also be summarized by the surplus equations, the reservation threshold $\widetilde{x}$ and labor market-tightness θ. Lemma 1 below summarizes these conditions.

Lemma 1. *In steady state, the social planner's optimal decisions can be summarized by the following equations:*

$$\mathbb{S}(x,0) = \frac{z(x-\tilde{x})}{\rho+s} \tag{32}$$

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

$$(\rho+\delta)\chi = q\left[\int_{\tilde{x}}^{\overline{x}}\mathbb{S}(x,1)\,d\Pi(x)-\xi(z\tilde{x}-b)\left(\frac{q}{\xi}\right)^{\frac{1-\alpha}{\alpha}}\right] \tag{34}$$

$$z\widetilde{x} = b+\alpha p\left[\left(\frac{v^u}{v}\right)\int_{\widetilde{x}}^{\overline{x}}\mathbb{S}(y,1)\,d\Pi(y)+\left(\frac{v^m}{v}\right)\int_{\widetilde{x}}^{\overline{x}}\int_{x}^{\overline{x}}\left[\mathbb{S}(y,1)-\mathbb{S}(x,1)\right]d\Pi(y)\,f^m(x)\,dx\right] \tag{35}$$

where $\mathbb{S}(x,0)$ refers to the social value of a firm-worker pair with match quality x but without an unexpired vacancy. Similarly, $\mathbb{S}(x,1)$ refers to the social value of a firm-worker pair with match quality x with an unexpired vacancy.

Proof. See Appendix B.1. □

In order to evaluate whether the outcomes in the decentralized economy are efficient, it is sufficient to check whether the $\widetilde{x}$ and θ in the decentralized economy are the same as that in the planner's problem. Lemma 2 below reproduces the analogous equations from the decentralized economy to facilitate a comparison between the outcomes in the decentralized economy and the social planner's choices

Lemma 2. *In steady state, the outcomes in the decentralized economy can be summarized by the* *following equations:*

$$S(x,0) = \frac{z(x-\tilde{x})}{\rho+s} \tag{36}$$

$$S(x,1) = z\left(\frac{\delta+\rho+s}{\rho+s}\right)\int_{\tilde{x}}^{x}\frac{\Big(\rho+\delta+s+q\left[1-\Pi(y)\right]\Big)^{\eta-1}}{\Big(\rho+\delta+s+q\left[1-\Pi(x)\right]\Big)^{-\eta}}dy \tag{37}$$

$$(\rho+\delta)\chi = q(1-\eta)\int_{\tilde{x}}^{\overline{x}} S(x,1)\,d\Pi(x) \tag{38}$$

$$z\widetilde{x} = b+\eta p\left[\left(\frac{v^u}{v}\right)\int_{\widetilde{x}}^{\overline{x}} S(y,1)\,d\Pi(y) + \left(\frac{v^m}{v}\right)\int_{\widetilde{x}}^{\overline{x}}\int_{x}^{\overline{x}} S(y,1)\,d\Pi(y)\,f^m(x)\,dx\right] \tag{39}$$

where equation (36) describes the surplus of a match with an expired vacancy in the decentralized economy and is the same as (12). Equation (37) describes the surplus of a match with an unexpired vacancy in the decentralized and is the solution to the equation (13).[22] *Equation (38) is the free-entry condition and is the identical to equation (16) and (39) defines the hiring threshold in the decentralized economy and is identical to equation (15).*

In many variants of the standard DMP setup, decentralized outcomes can be rendered efficient by the appropriate choice of the bargaining weight η. In such decentralized economies, firms impose a "congestion" externality on other firms by reducing the rate at which the other vacancies meet firms. However, in posting a vacancy, the firm imposes a positive externality on workers by making it easier for workers to meet vacancies. These externalities can often be exactly offset by an appropriate choice of the bargaining weight η which effectively makes agents internalize the effects of their decisions (Hosios, 1990).

Simply, if setting η to an appropriate value makes equations (32)-(35) identical to (36)-(39), then the outcomes of the decentralized economy can be efficient. Notice that setting $\eta = 0$,or eliminating the workers bargaining power, makes equations (32)-(34) identical to (36)-(38) but equations (35) and (39) do not agree with each other. In fact, Proposition 3 below establishes that there is no $\eta \in [0,1]$ for which the outcomes in the decentralized economy are the same as the choice of the social planner.

Proposition 3 (Generic Inefficiency with Nash Bargained wages)**.** *There exists no $\eta \in [0,1]$ such that the $\widetilde{x}$ and θ in the decentralized economy are the same as the social planners choices.*

Proof. See Appendix B.2. □

In order to appreciate why this is the case, inspection of equations (35) and (39) reveals that the very nature of replacement hire introduces another externality which Nash-Bargaining is unable to price. Equation (35) is analogous to the value of an unemployed worker in the

[22] See Appendix A.2 for details.

decentralized case (39)[23]. In the decentralized economy, setting $\eta = \alpha$ almost makes these equations the same but not quite. While the planner only counts the net gain, $\mathbb{S}(y,1) - \mathbb{S}(x,1)$, when an unemployed worker who draws match productivity y replaces an existing worker who had match productivity $x < y$, the unemployed agent and firm in the decentralized economy only care about the gross change, $S(y,1) - 0$. In other words, the decentralized economy does not internalize that a replacement hire also destroys value by removing a worker with match quality x from the workforce.

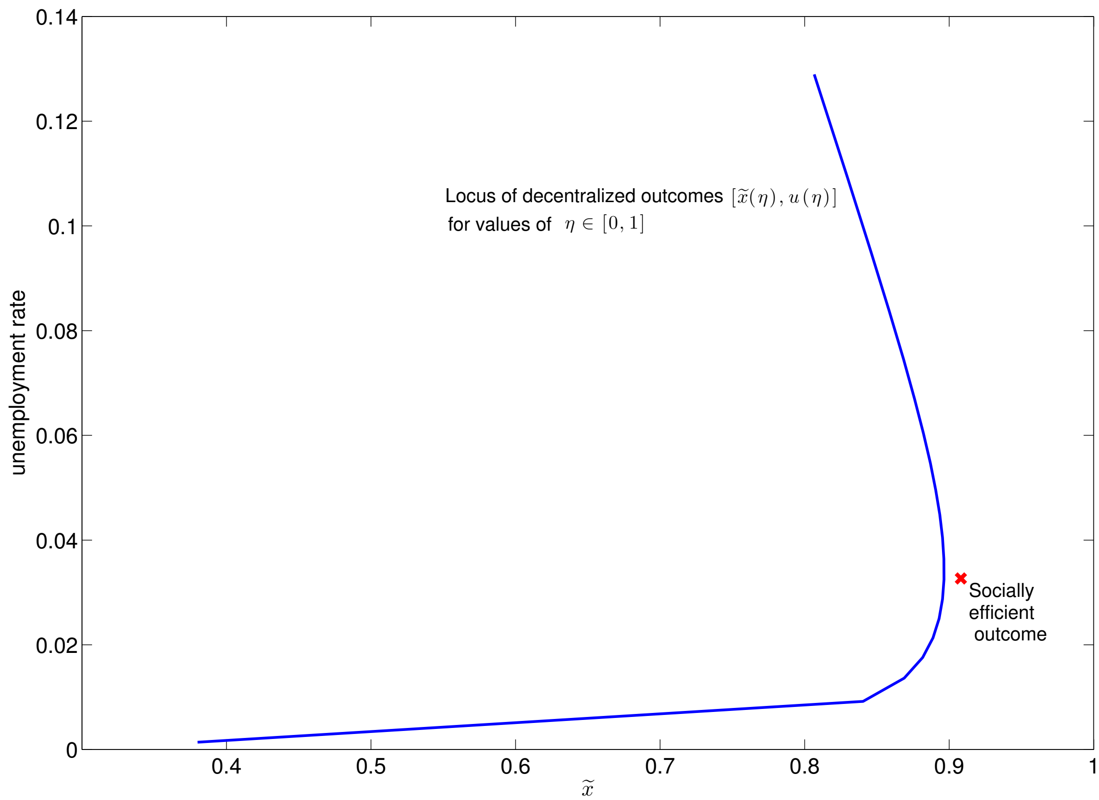

Figure 9: Efficient vs Decentralized Outcomes

In fact, a policy maker in the decentralized economy who can choose η freely is able to choose the same θ as the social planner but will be unable to implement a high enough $\widetilde{x}$ as the social planners choice. The following Proposition formalizes this.

Proposition 4. *Consider a policymaker in the decentralized economy who is free to choose any $\eta \in [0,1]$. This planner can choose η such that $\theta_{decentralized} = \theta_{efficient}$ but such a choice implies that the threshold match quality in the decentralized economy $\widetilde{x}_{decentralized} < \widetilde{x}_{efficient}$.*

[23]Notice that equation (39) is identical to (2) in equilibrium since $z\widetilde{x} = \rho U$

Proof. See Appendix ??. □

Proposition 4 shows that choosing η to eliminate the congestion effects of posting vacancies pushes down the bargaining power of workers so much that firms are able to employ them at relatively low wages. These lower wages, in turn allow even less profitable firms to become profitable and lowers the match productivity threshold below which a firm would choose not to create a job. As a result, there are too many low-quality jobs in this economy. These low quality jobs in turn imply that it is easy for firms with active vacancies to replace their current low quality workers with better workers. The social planner would instead prefer a higher $\widetilde{x}$ as that would reduce the number of replacement hires allowing agents since the planner only values these replacement hires for the net productivity gains they generate rather than the private agents who only value these for the gross gains.

Figure 9 graphically depicts the content of Proposition 4 by plotting the planners choice of $(\widetilde{x}, \theta)$ relative to the locus of decentralized outcomes of the same quantities for $\eta \in [0, 1]$.[24] Compared to our benchmark economy calibrated to match the data moments for the period 1951-2007, we find that the efficient level of output is 2 percent higher. Moreover, the decentralized economy created too few jobs relative to the efficient outcome and features an unemployment rate which is 50% higher than the socially efficient level of unemployment. Under the social planner's choices the composition of vacancies tilts towards more unfilled vacancies relative to match vacancies. The decentralized economy features 60% of all vacancies being matched vacancies while the social planner chooses the same share to be 51%. Figure 10 plots the cumulative distribution function of the match-quality for matched firms with unexpired vacancies. As can be seen, the decentralized economy features a lower $\widetilde{x}$ and a larger fraction of jobs with low match quality relative to the efficient outcome.

7 Conclusion

We document that the share of replacement hires in the US has risen over time without any corresponding increase in the share of separations that are quits or vacancy posting rates. These facts are hard to reconcile through the lens of the standard DMP model and we provide a framework that can account for how replacement hires can occur without the event of a quit or a new vacancy posting. Our model suggests that the decline in vacancy expiration and the rise in firms' outside options are the key drivers behind the increase in the share of replacement hires. Further, our model predicts that lower wages would be observed in an economy where replacement hiring is prevalent and where the composition of matched firms shifts towards firms that have unexpired vacancies. We also investigate the efficiency properties of our model and

[24] We use the parameters which describe the model for the period 1951-2007 and vary η to construct the decentralized outcomes and also the efficient outcome.

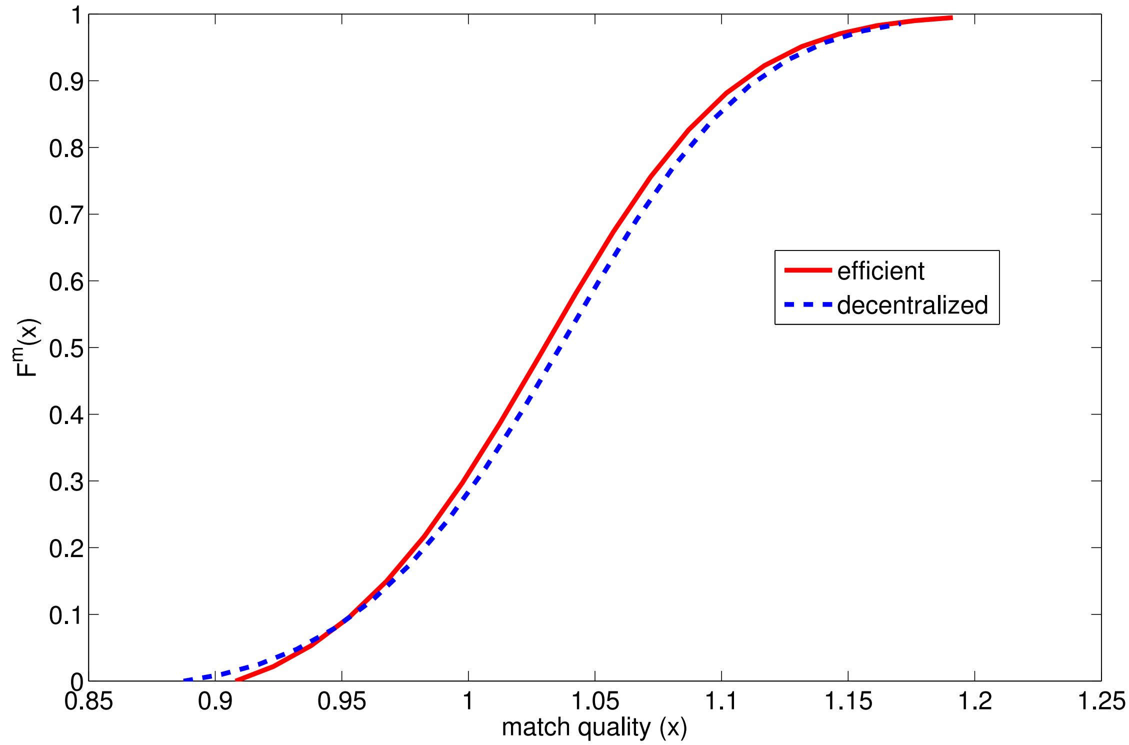

Figure 10: Distribution of Match Quality for Firms

find that while firms generate private productivity gains from replacement hiring, firms do not internalize the value lost by their current worker whenever she is replaced with a better match.

References

Albrecht, James and Susan Vroman, "Directed Search with Phantom Vacancies," Technical Report 2015.

Barnichon, Regis, "Building a composite Help-Wanted Index," *Economics Letters*, December 2010, *109* (3), 175–178.

Cheron, Arnaud and Bruno Decreuse, "Matching with Phantoms," *The Review of Economic Studies*, 2016, p. rdw032.

Davis, Steven J., R. Jason Faberman, and John C. Haltiwanger, "The Establishment-Level Behavior of Vacancies and Hiring," *The Quarterly Journal of Economics*, 2013, *128* (2), 581–622.

Elsby, Michael, Bart Hobijn, and Ayseful 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.

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

Hall, Robert E., "High Discounts and High Unemployment," *American Economic Review*, February 2017, *107* (2), 305–30.

Hosios, Arthur J., "On The Efficiency of Matching and Related Models of Search and Unemployment," *Review of Economic Studies*, 1990, *57* (2), 279–298.

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

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

Shimer, Robert, "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, March 2005, *95* (1), 25–49.

Appendix

A Surplus and Wage Determination

A.1 Nash Bargaining

Wages are determined by Nash Bargaining and wages can be expressed as a solution to the problem:

$$\max_{w(x,i)} \Big[J(x,i) - J(i)\Big]^{1-\eta} \Big[W(x,i) - U\Big]^{\eta} \text{ for } i = 0, 1 \tag{40}$$

The Nash-Bargaining solution has the feature that the surplus of a match between a firm and worker satisfies:

$$(1-\eta)[W(x,i) - U\Big] = \eta\Big[J(x,i) - J(i)\Big] \tag{41}$$

Then using the definition of surplus $S(x,i) = J(x,i) - J(i) + W(x,i) - U$ for $i \in \{0,1\}$, under the Nash Bargaining solution, the worker gets a fraction η of the surplus while the firm receives the remaining $1-\eta$:

$$W(x,i) - U = \eta S(x,i) \tag{42}$$

$$J(x,i) - J(i) = (1-\eta)S(x,i) \tag{43}$$

for $\in \{0,1\}$.

A.2 Surplus

Surplus of a match with an inactive vacancy We start of by deriving the expression for the surplus of a firm-worker pair who do not possess an unexpired vacancy. Notice that by imposing free entry, $J(0) = 0$, equation (5) can be written as:

$$(\rho + s)J(x,0) = zx - w(x,0) \tag{44}$$

. Also, subtracting ρU from both sides of equation (3), we can re-write that equation as:

$$(\rho + s)\,[W(x,0) - U] = w(x,0) - \rho U \tag{45}$$

Then using the definition of surplus $S(x,0) = J(x,0) - J(0) + W(x,0) - U$, we can write an expression for the surplus of a match with an inactive vacancy as:

$$(\rho + s)S(x,0) = zx - \rho U \tag{46}$$

which is the expression in equation (12). Notice that since equation (15) states that $z\widetilde{x} = \rho U$, we can re-write the surplus equation as:

$$(\rho + s)S(x,0) = z(x - \widetilde{x}) \tag{47}$$

which is the same as equation (36) in the text. Notice that just like a match with an active vacancy, the surplus of a match with an inactive vacancy also goes to zero at $\widetilde{x}$.

Surplus of a match with an active vacancy In order to derive an expression for the surplus of a firm-worker pair with an active vacancy, we start by subtracting equation (8) from (7) and rearranging:

$$\begin{aligned}\Big(\rho + \delta + s + q\,[1 - \Pi(x)]\Big)\,[J(x,1) - J(1)] \;&=\; zx - w(x,1) - q\int_{\widetilde{x}}^{x} [J\,(y,1) - J(1)]\,d\Pi(y) \\ &\quad + \delta\,[J(x,0) - J(0)]\end{aligned} \tag{48}$$

Similarly, subtracting equation ρU from both sides of (4) and rearranging yields:

$$\Big(\rho + \delta + s + q\,[1 - \Pi(x)]\Big)\,[W(x,1) - U] = w(x,1) + \delta\,[W(x,0) - U] - \rho U \tag{49}$$

Adding the two expressions above yields an expression for the surplus:

$$\Big(\rho + \delta + s + q\,[1 - \Pi(x)]\Big)S(x,1) \;=\; zx - \rho U - q\int_{\widetilde{x}}^{x} [J\,(y,1) - J(1)]\,d\Pi(y) + \delta S(x,0) \tag{50}$$

where $S(x,1) = J(x,1) - J(1) + W(x,1) - U$. Next using the Nash Bargaining solution (43), the equation above can be re-written as:

$$\Big(\rho + \delta + s + q\,[1 - \Pi(x)]\Big)S(x,1) \;=\; zx - \rho U - q(1-\eta)\int_{\widetilde{x}}^{x} S\,(y,1)\,d\Pi(y) + \delta S(x,0) \tag{51}$$

which is identical to equation (13) in the main text.

Next, take the derivative of (51) with respect to x:

$$-\frac{\eta q \Pi'(x)}{\rho + \delta + s + q\,[1 - \Pi(x)]} S\,(x,1) + S'(x,1) \;=\; \left[\frac{\delta + \rho + s}{\rho + S}\right] \frac{z}{\rho + \delta + s + q\,[1 - \Pi(x)]} \tag{52}$$

which is an ordinary differential equation in x. Using the boundary condition $S(\widetilde{x},1)=0$, we can solve for $S(x,0)$ for all $x\geq\widetilde{x}$:

$$S\left(x,1\right) = z\left[\frac{\delta+\rho+s}{\rho+s}\right]\left(\rho+\delta+s+q\left[1-\Pi\left(x\right)\right]\right)^{-\eta}\int_{\widetilde{x}}^{x}\left(\rho+\delta+s+q\left[1-\Pi\left(y\right)\right]\right)^{\eta-1}dy \tag{53}$$

which is the same as equation (37) in the text.

A.3 Wages

Wages paid by a firm with an inactive vacancy and match productivity x In order to define the expressions for the wages, we start by using the fact that $J(x,0)=(1-\eta)S(x,0)$ in equation (47) and rearranging:

$$(\rho+s)J(x,0)=(\rho+s)(1-\eta)S(x,0)=z(1-\eta)(x-\widetilde{x}) \tag{54}$$

Using equation (44), we can rewrite the above equation as:

$$(\rho+s)J(x,0)=z(1-\eta)(x-\widetilde{x})=zx-w(x,0) \tag{55}$$

Rearranging:

$$w(x,0)=z\widetilde{x})+\eta z(x-\widetilde{x})\text{ for }x\geq\widetilde{x} \tag{56}$$

which is identical to equation (25), Proposition 1 in the main text.

Wages paid by a firm with an active vacancy and match productivity x In a similar fashion as above, using $J(x,1)-\chi=(1-\eta)S(x,1)$ in equation (51) and rearranging:

$$\begin{aligned}\left(\rho+\delta+s+q\left[1-\Pi(x)\right]\right)[J(x,1)-\chi] &= (1-\eta)z(x-\widetilde{x})-q(1-\eta)\int_{\widetilde{x}}^{x}[J\left(y,1\right)-\chi]d\Pi(y)\\ &\quad +\delta J(x,0)\end{aligned}$$

Using equation (48) with the equation above and rearranging, we can get the expression for wages:

$$\begin{aligned}w(x,1) &= z\widetilde{x}+\eta z(x-\widetilde{x})-\eta(1-\eta)q\int_{\widetilde{x}}^{x}[J\left(y,1\right)-\chi]d\Pi(y)\\ &= w(x,0)-\eta(1-\eta)q\int_{\widetilde{x}}^{x}S\left(y,1\right)d\Pi(y)\end{aligned} \tag{57}$$

which is the same expression as (26), Proposition 1 in the main text.

A.4 Match Efficiency

Denote the average acceptance rate as A which is defined as:

$$A = \frac{v^u}{v}\left[1 - \Pi(\tilde{x})\right] + \frac{v^m}{v}\int_{\tilde{x}}^{\overline{x}}\left[1 - \Pi(\epsilon)\right] f^m(\epsilon)d\epsilon \tag{58}$$

Use the fact that $\frac{v^u}{v} = 1 - \frac{v^m}{v}$ and take the derivative of A wrt $\frac{v^m}{v}$:

$$\frac{\partial A}{\partial\left(v^m/v\right)} = \int_{\tilde{x}}^{\overline{x}}\left[\Pi(\tilde{x}) - \Pi(\epsilon)\right] f^m(\epsilon)d\epsilon \leq 0 \tag{59}$$

where the inequality is strict if $\tilde{x} \neq \overline{x}$.

B Efficient Benchmark

B.1 Social Planner's Problem

Rather than solving the problem of the planner as choosing the reservation productivity $\widetilde{x}$ and new vacancies v^{new}, it is easier to set up the problem in which instead of choosing the threshold $\widetilde{x}$, the planner chooses an *acceptance function* $a(x)$ for each level of match productivity x. This acceptance function $a(x) \in [0,1]$ for all $x \in [\underline{x}, \overline{x}]$. By definition, $\widetilde{x}$ is the value of match productivity above which $a(x) = 1$. In terms of the social planner's problem, this reformulation means that the constraints given by equations (28) and (30) have to be altered to:

$$\dot{\mu}_t^m(x) = \left[a_t(x)\, q_t \pi(x)\left(v_t^u + \int_{\underline{x}}^{x} \mu_t^m(\epsilon)\, d\epsilon\right) - \left(s + \delta + q_t\int_{x}^{\overline{x}} a_t(\epsilon)\,\pi(\epsilon)\, d\epsilon\right)\mu_t^m(x)\right] \tag{60}$$

and

$$\dot{v}_t^u = v_t^{new} + s\int_{\underline{x}}^{\overline{x}} \mu_t^m(\epsilon)\, d\epsilon - \left(q_t\int_{\underline{x}}^{\overline{x}} a(\epsilon)\,\pi(\epsilon)\, d\epsilon + \delta\right) v_t^u \tag{61}$$

respectively. The rest of the problem remains the same. The problem of the social planenr can be expresses as a Hamiltonian:

$$
\begin{aligned}
\mathcal{H}_t &= -\chi v^{new} + \int_{\underline{x}}^{\overline{x}} (zx-b)\,\mu^m(x)\,dx + \int_{\underline{x}}^{\overline{x}} (zx-b)\,\mu^i(x)\,dx + b \\
&+ \int_{\underline{x}}^{\overline{x}} \Lambda(x,1)\left[a(x)\,q\pi(x)\left(v^u + \int_{\underline{x}}^{x} \mu^m(\epsilon)\,d\epsilon\right) - \left(s+\delta+q\int_{x}^{\overline{x}} a(\epsilon)\,\pi(\epsilon)\,d\epsilon\right)\mu^m(x)\right]dx \\
&+ \int_{\underline{x}}^{\overline{x}} \Lambda(x,0)\left[\delta\mu^m(x) - s\mu^i(x)\right]dx \\
&+ \varphi\left[v^{new} + s\int_{\underline{x}}^{\overline{x}} \mu^m(\epsilon)\,d\epsilon - \left(q\int_{\underline{x}}^{\overline{x}} a(x)\,\pi(x)\,dx + \delta\right)v^u\right] \\
&- \psi\left[\ln q - \ln\xi + \alpha\ln\left(v^u + \int_{\underline{x}}^{\overline{x}} \mu^m(\epsilon)\,d\epsilon\right) - \alpha\ln\left(1 - \int_{\underline{x}}^{\overline{x}} \mu^m(\epsilon)\,d\epsilon - \int_{\underline{x}}^{\overline{x}} \mu^i(\epsilon)\,d\epsilon\right)\right]
\end{aligned}
$$

where the choice variables are v^{new}, q and $a(x)$. The state variables are given by the density of matched pairs of type x with active an inactive vacancies and the stock of unfilled vacancies. $\Lambda(x,1)$ is the co-state associated with the state $\mu^m(x)$ and denotes the gross social value of a type x matched job with an active vacancy. Similarly, $\Lambda(x,0)$ is the co-state associated with the state $\mu^i(x)$ and denotes the gross social value of a type x matched job with an inactive vacancy. φ_t is the co-state associated with v^u and denotes the benefit of an additional unfilled vacancy. ψ denotes the multiplier on the matching function.

The optimal decisions of the planner in steady-state can be described by:

$$
\chi = \varphi \tag{62}
$$

$$
\begin{aligned}
\frac{\psi}{q} &= \int_{\underline{x}}^{\overline{x}} \Lambda(x,1)\left[a(x)\,\pi(x)\left(v^u + \int_{\underline{x}}^{x} \mu^m(\epsilon)\,d\epsilon\right) - \int_{x}^{\overline{x}} a(\epsilon)\,\pi(\epsilon)\,d\epsilon\,\mu^m(x)\right]dx \\
&\quad -\varphi\left(\int_{\underline{x}}^{\overline{x}} a(x)\,\pi(x)\,dx\right)v^u
\end{aligned} \tag{63}
$$

$$
a(x) = \begin{cases} 0 & \text{if } q\pi(x)\left[\Lambda(x,1)\,v^u + \int_{\underline{x}}^{x}\left[\Lambda(x,1)-\Lambda(\epsilon,1)\right]\mu^m(\epsilon)\,d\epsilon - \varphi v^u\right] < 0 \\ 1 & \text{if } q\pi(x)\left[\Lambda(x,1)\,v^u + \int_{\underline{x}}^{x}\left[\Lambda(x,1)-\Lambda(\epsilon,1)\right]\mu^m(\epsilon)\,d\epsilon - \varphi v^u\right] \geq 0 \end{cases} \tag{64}
$$

$$
\begin{aligned}
\rho\Lambda(x,1) &= zx - b + q\int_{x}^{\overline{x}} \Lambda(\epsilon,1)\,a(\epsilon)\,\pi(\epsilon)\,d\epsilon - \Lambda(x,1)\left(s+\delta+q\int_{x}^{\overline{x}} a(\epsilon)\,\pi(\epsilon)\,d\epsilon\right) \\
&\quad +\delta\Lambda(x,0) + s\varphi - \alpha\left[\frac{\psi}{v} + \frac{\psi}{u}\right]
\end{aligned} \tag{65}
$$

$$
\rho\Lambda(x,0) = zx - b - s\Lambda(x,0) - \alpha\frac{\psi}{u} \tag{66}
$$

$$
\rho\varphi = \int_{\underline{x}}^{\overline{x}} \Lambda(x,1)\,a(x)\,q\pi(x)\,dx - \varphi\left[q\int_{\underline{x}}^{\overline{x}} a(x)\,\pi(x)\,dx + \delta\right] - \alpha\frac{\psi}{v} \tag{67}
$$

where equations (62)-(64) are the first order conditions w.r.t. v^{new}, q and $a(x)$ respectively. Equations (65)-(67) are the equations which describe the evolution of the co-state variables $\Lambda(x,1), \Lambda(x,0)$ and φ respectively (with steady state imposed).

Using equation (64), one can define $\widetilde{x}$ implicitly as the x which satisfies:

$$\Lambda\left(\widetilde{x},1\right)=\chi=\varphi \tag{68}$$

In other words, the planner would not want to maintain a match which is worth less than a unfilled vacancy. Imposing the optimal $a(x)$ in the system of equations above, equations (63), (70) and (67) can be rewritten as:

$$\frac{\psi}{u} = q\theta\left[\left(\frac{v^u}{v}\right)\int_{\tilde{x}}^{\overline{x}}\left[\Lambda\left(x,1\right)-\Lambda\left(\tilde{x},1\right)\right]\pi\left(x\right)dx+\frac{v^m}{v}\int_{\tilde{x}}^{\overline{x}}\int_{x}^{\overline{x}}\left[\Lambda\left(\epsilon,1\right)-\Lambda\left(x,1\right)\right]d\Pi\left(\epsilon\right)\frac{\mu^m\left(x\right)}{v^m}dx\right] \tag{69}$$

$$\left(\rho+s+\delta+q\left[1-\Pi\left(x\right)\right]\right)\Lambda\left(x,1\right) = zx-b-\frac{\alpha\psi}{u}+\delta\Lambda\left(x,0\right)+s\chi+q\int_{x}^{\overline{x}}\Lambda\left(\epsilon,1\right)d\Pi\left(\epsilon\right)-\frac{\alpha\psi}{v} \tag{70}$$

$$\left(\rho+\delta+q\left[1-\Pi\left(\tilde{x}\right)\right]\right)\chi=q\int_{\tilde{x}}^{\overline{x}}\Lambda\left(x,1\right)d\Pi\left(x\right)-\alpha\frac{\psi}{v} \tag{71}$$

Using equations (71) and (66) in (70), we can get:

$$\left(\rho+\delta+s+q\left[1-\Pi\left(x\right)\right]\right)\mathbb{S}\left(x,1\right) = \left(\frac{\rho+s+\delta}{\rho+s}\right)\left[zx-b-\frac{\alpha\psi}{u}\right]-q\int_{\tilde{x}}^{x}\mathbb{S}\left(y,1\right)d\Pi\left(y\right) \tag{72}$$

where $\mathbb{S}(x,1)=\Lambda(x,1)-\chi$ is the social value of a matched vacancy net of the value of an unfilled vacancy. Differentiating this equation with respect to x yields:

$$\left(\rho+\delta+s+q\left[1-\Pi\left(x\right)\right]\right)\mathbb{S}'\left(x,1\right) = \left(\frac{\rho+s+\delta}{\rho+s}\right)z \tag{73}$$

We also know that $\mathbb{S}(\widetilde{x},1)=\Lambda(\widetilde{x},1)-\chi=0$. This boundary value problem has a unique solution given by:

$$\mathbb{S}\left(x,1\right)=z\left(\frac{\rho+s+\delta}{\rho+s}\right)\int_{\tilde{x}}^{x}\frac{1}{\rho+s+\delta+q\left[1-\Pi\left(y\right)\right]}dy \tag{74}$$

which is identical to equation (33), Lemma 1 in the main text.

In addition, evaluating equation (72) at $x=\widetilde{x}$, it must be the case that $z\widetilde{x}=b+\frac{\alpha\psi}{u}=0$.

Using this fact, we can rewrite equation (66) as:

$$\mathbb{S}(x,0) = [\Lambda(x,0) - 0] = \frac{z(x-\widetilde{x})}{\rho+s} \tag{75}$$

which is identical to equation (32), Lemma 1 in the main text. Next, using the definition of $\mathbb{S}(x)$ and rearranging equation (71), we can derive the analog to the free entry condition:

$$(\rho+\delta)\chi \quad = \quad q\int_{\tilde{x}}^{\overline{x}} \mathbb{S}\left(x,1\right) d\Pi\left(x\right) - \frac{\alpha\psi}{v} \tag{76}$$

Since we know that $z\widetilde{x} = b + \frac{\alpha\psi}{u}$, we can rewrite the expression above as:

$$\begin{aligned} (\rho+\delta)\chi \quad &= \quad q\int_{\tilde{x}}^{\overline{x}} \mathbb{S}\left(x,1\right) d\Pi\left(x\right) - \frac{z\widetilde{x}-b}{\theta} \\ &= \quad q\int_{\tilde{x}}^{\overline{x}} \mathbb{S}\left(x,1\right) d\Pi\left(x\right) - \xi(z\widetilde{x}-b)\left(\frac{q}{\xi}\right)^{\frac{1-\alpha}{\alpha}} \end{aligned} \tag{77}$$

which is identical to equation (34), Lemma 1 in the main text.[25] Relative to the decentralized free-entry condition, the planner's valuation of creating a new vacancy is lower by the term $\alpha\psi/u$ because the planner internalizes that removing a worker from unemployment and putting her into employment affects the market-tightness. Private firms do not internalize this effect when making their decisions.

Using the definition of $\mathbb{S}(x,1)$ in equation (69) and rearranging yields:

$$z\tilde{x} = b + \alpha q\theta\left[\left(\frac{v^u}{v}\right)\int_{\tilde{x}}^{\overline{x}} \mathbb{S}\left(y,1\right) d\Pi\left(y\right) + \frac{v^m}{v}\int_{\tilde{x}}^{\overline{x}}\int_{x}^{\overline{x}} \left[\mathbb{S}\left(y,1\right) - \mathbb{S}\left(x,1\right)\right] d\Pi\left(y\right)\frac{\mu^m\left(x\right)}{v^m}dx\right] \tag{78}$$

which is identical to equation (35), Lemma 1 in the main text.

B.2 Proof of Proposition 3

Let us suppose that $\exists\eta \in [0,1]$, such that $\tilde{x}$ and q generated by the planners choice are the same in the decentralized economy. Under this assumption and the fact that $(\tilde{x},1) = 0$ and $\mathbb{S}\left(\tilde{x},1\right) = 0$

[25] where we have used the definition of the matching function in going from the first line to the second.

we know that:

$$\begin{aligned}\frac{dS(x,1)}{dx} &= z\left(\frac{\delta+\rho+s}{\rho+s}\right)\frac{1}{\rho+\delta+s+q[1-\Pi(x)]}+\frac{\eta q\pi(x)S(x,1)}{\rho+\delta+s+q[1-\Pi(x)]}\\ &\geq z\left(\frac{\rho+s+\delta}{\rho+s}\right)\frac{1}{\rho+s+\delta+q[1-\Pi(x)]}\\ &= \frac{d\mathbb{S}(x,1)}{dx}\end{aligned}$$

In other words, the private surplus in the decentralized economy for any $x > \widetilde{x}$ is greater than the social surplus. Next, subtracting equation the decentralized free entry condition (38) from the socially optimal free entry (34), we get:

$$-\xi^{\frac{1}{\alpha}}(1-\eta)\int_{\tilde{x}}^{\overline{x}}[S(y,1)-\mathbb{S}(x,1)]\,d\Pi(y)=(z\tilde{x}-b)\,q^{\frac{1-\alpha}{\alpha}} \tag{79}$$

Since $S(x,1)\geq\mathbb{S}(x,0)$, the LHS of the equation above is non-positive for the equation to hold, it must be the case that

$$z\widetilde{x}-b\leq 0$$

We start by considering the case $\eta\in(0,1]$. Then equation (79) requires that $z\widetilde{x}-b<0$. Consequently, equation (39) implies that:

$$\left(\frac{v^u}{v}\right)\int_{\tilde{x}}^{\overline{x}}S(y,1)d\Pi(y)+\left(\frac{v^m}{v}\right)\int_{\tilde{x}}^{\overline{x}}\left[\int_x^{\overline{x}}S(y,1)d\Pi(y)\right]f^m(x)dx<0 \tag{80}$$

However, this is a contradiction since $S(x,1)\geq 0$ for $x\geq\widetilde{x}$.

Finally, we consider the case when $\eta=0$. Notice that for $\eta=0$, equations (33) and (37) are identical. Since $\widetilde{x}$ and q are the same by assumption across the decentralized and efficient outcomes, it must be that $S(x,1)=\mathbb{S}(x,1)$ for all $x\geq\widetilde{x}$ as long as $\eta=0$. Since equation (79) requires that $z\widetilde{x}<b$, equation (35) implies that

$$\left(\frac{v^u}{v}\right)\int_{\tilde{x}}^{\overline{x}}\mathbb{S}(y,1)\,d\Pi(y)+\left(\frac{v^m}{v}\right)\int_{\tilde{x}}^{\overline{x}}\int_x^{\overline{x}}[\mathbb{S}(y,1)-\mathbb{S}(x,1)]\,d\Pi(y)\,\frac{\mu^m(x)}{v^m}dx<0 \tag{81}$$

which requires that (using $S(x,1)=\mathbb{S}(x,1)$)

$$\int_{\tilde{x}}^{\overline{x}}\left(\mathbb{E}_{\mu^m}[S(x,1)]-\mathbb{E}_{\mu^m}[S(y,1)\mid y\geq x]\right)[1-\Pi(x)]\,\frac{\mu^m(x)}{v^m}dx>0$$

which cannot hold since $S(x,1)$ is a strictly increasing function for $x\geq\widetilde{x}$.