ISSN 1936-5330

THE FEDERAL RESERVE BANK of KANSAS CITY
RESEARCH WORKING PAPERS

Job Duration and Match Quality Over the Cycle

José Mustre-del-Río

November 2012; Updated September 2016
RWP 12-08

RESEARCH WORKING PAPERS

Job Duration and Match Quality Over the Cycle *

José Mustre-del-Río
Federal Reserve Bank of Kansas City

First Version: November 2012
This Version: September 2016

Abstract

This paper examines how job quality varies over the cycle. Empirical evidence from the National Longitudinal Survey of Youth (NLSY) suggests match quality is procyclical. This interpretation is corroborated in a calibrated model with on-the-job search. In the model, more high quality matches are observed in expansions because of improved reallocation through on-the-job search. This effect, however, is dominated by lower job destruction in expansions, which preserves matches at the bottom of the quality distribution.
Keywords: Business cycles, search and matching, directed search on-the-job , match quality, job duration.
JEL classification: E24, E32, J21, J22, J62, J63, J64.

*The author thanks Mark Bils, Yongsung Chang, Mary Daly, Jessica Ford, William Hawkins, Toshihiko Mukoyama, Nick Sly, Willem Van Zandweghe, Benjamín Villena-Roldán, and seminar participants at Rochester and the Kansas City Fed for useful discussions and comments. Excellent research assistance was provided by William Xu. The views expressed in this paper are solely those of the author and do not necessarily represent the views of the Federal Reserve Bank of Kansas City or the Federal Reserve System.
E-mail: jose.mustre-del-rio@kc.frb.org.

A central question in economics is how business cycles affect the allocation of resources. Focusing on the labor market, it is a priori unclear how cycles affect productive arrangements. If matches are heterogeneous in their quality and are allowed to dissolve endogenously, recessions may "cleanse" the labor market from a certain *type* of matches. Following a negative productivity shock, some jobs at the bottom of the quality distribution fail to generate positive surplus and are destroyed. As reservation match quality is now higher, only exceptionally high quality matches are formed, and as a result average quality rises. In contrast, recessions may "sully" the labor market by reducing the *frequency* of reallocation (e.g. **?**). In recessions, the job-to-job transition rate falls as firms post fewer vacancies. This reduces the rate at which workers reallocate into better matches, and hence average quality falls.

This paper examines how changes in the *frequency* of reallocation and differences in the *types* of matches created (or destroyed) affect match quality over the business cycle. Empirical evidence from the National Longitudinal Survey of Youth (NLSY) 1979-2010 suggests match quality is procyclical. Viewing these empirical results through the lens of a search model reveals the evidence is consistent with match quality improving in expansions because increased reallocation through job-to-job transitions. However, the model also provides quantitative statements regarding the types of matches created over the cycle. In the model, low quality matches are created (or preserved) in expansions because reservation quality falls. Overall, this effect dominates so that average match quality actually falls in expansions. Thus, the analysis provides new insights regarding the empirical evidence and quantitative significance of how business cycles affect productive arrangements in the labor market.

The empirical results are based on data from the NLSY using the duration of a match as a measure of quality. These estimates suggest matches starting in expansions are systematically longer than matches formed in recessions. For example, median duration rises from 25 months to 34 months when the unemployment rate at the time of match formation falls by one standard deviation (1.5 percentage points). This effect is slightly stronger when looking at matches formed by workers who were previously employed. Additionally,

the finding is robust to including initial wages as an explanatory variable, suggesting higher match quality in expansions is not fully captured by starting wages.

Next, a version of Menzio and Shi's (2011) model of directed on-the-job search is presented to provide a theoretical context to the empirical results. This model is ideal for the present analysis as it is readily solvable out of steady state, and hence provides predictions at cyclical frequencies. Applying the same estimation strategy used on the NLSY sample to model simulated data delivers similar coefficients. A deeper inspection reveals these model based results are indeed driven by a "sullying effect" operative in the model: in expansions on-the-job search increases reallocation away from lower quality matches into better ones. Overall, however, the model implies that expansions are associated with lower average match quality because of increased match creation at the bottom of the quality distribution; i.e. the "cleansing effect" dominates.

This paper is related to several strands of the empirical and quantitative macro labor literature. The empirical analysis is closely related to the work of **?** who finds that matches starting in recessions are of shorter duration.[1] She also uses data from the NLSY, but only through 1988. This paper complements her work by considering a much longer time horizon with greater cyclical variation, accounting for individual unobserved heterogeneity, and distinguishing between matches formed by previously employed versus nonemployed workers. Accounting for individual unobserved heterogeneity is critical for the finding that match quality is not solely internalized by initial wages. Distinguishing matches by prior employment status of the worker further ties the empirical analysis with the calibrated model of on-the-job search. Lastly, this paper shows that from this empirical evidence one cannot necessarily conclude the sullying effect dominates at the aggregate level.

Also related is the more recent empirical work on sullying effects. Using firm-level data, **?** also finds that employment relationships that start in recessions are short-lived. However, once firm heterogeneity is taken into ac-

[1] The notion that a good match is represented by a lengthy duration is inspired by **?**.

count this effect is reversed, suggesting the importance of firm differences in explaining the sullying effect. This paper compliments hers by focusing on the worker side. Using U.S. matched employer-employee data, **?** find that downturns hinder the progression of workers toward higher paying firms. This paper is complimentary to theirs as it provides a quantitative accounting of the cleansing versus sullying effects at the aggregate level using a model of on-the-job search. Lastly, **?** and **?** find large and persistent earnings declines for new graduates entering the labor market during a recession (i.e. a sullying effect).[2] They also find that the effect is strongest for the least skilled workers, reminiscent of a cleansing effect. Relative to these papers, the current paper is silent about long-term individual consequences of entering the labor market in a recession versus boom, and focuses instead on aggregate consequences.

On the quantitative side, the calibrated model used in this paper is taken from **?** (MS henceforth). This paper is complimentary to theirs as it shows their model generates empirically plausible cyclical fluctuations in job duration, aside from a correct description of worker transitions across employment states. This finding is important since in their model there is a direct relationship between job duration and match quality.[3] Also related is the work of **?** and **?**. **?** presents a model of random on-the-job search. Unlike this paper, he finds that recessions are, on net, sullying when wages are determined at a fixed piece rate. **?** develop a rich model of random on-the-job search with ex-ante heterogeneity both on the firm and worker side. They find increased sorting from on-the-job search during expansions.

The next section discusses the estimation strategy and the data used for the empirical analysis. Section 2 presents the estimation results. Section 3 outlines the MS model and presents the corresponding empirical and quantitative results. Section 4 concludes.

[2]See also, **?** who finds similar results when using the NLSY.

[3]See also, **?**.

1 Estimation and data description

This section first outlines the estimation procedure used to measure cyclical variation in match quality. Next, a description of the data is provided, followed by a discussion of the representativeness of the sample.

1.1 Estimation strategy

A proportional hazard model with time-varying regressors is estimated to empirically assess the cyclicality of match quality. This type of model is chosen as it allows for the inclusion of censored observations in the estimation without imposing additional assumptions on the hazard function.[4] Specifically, the hazard takes the form:

$$\lambda(t|\mathbf{X}(t)) \quad = \quad \lambda_0(t)\exp(\beta'\mathbf{X}(t)) \tag{1}$$

Here $\lambda_0(\tau)$ represents the baseline hazard of a job ending at time t; β is a coefficient vector to be estimated, and $\mathbf{X}(t)$ is a vector of individual and aggregate characteristics at time t. Over the duration of the job spell, these characteristics are either constant (e.g. race, education, and cyclical conditions at the start of the job) or time-varying (e.g. labor market experience, and current labor market measures). Analysis time for this hazard begins when the match is first observed and ends when the match dissolves (or is no longer observed).

Following **?**, cyclical conditions in the labor market are proxied by the national unemployment rate.[5] Like **?**, the vector $\mathbf{X}(t)$ includes: the national unemployment rate when the match begins u_0, the current unemployment rate u_t, and the square of the current unemployment rate u_t^2. Unlike her, $\mathbf{X}(t)$ also includes an interaction term between the initial and current unemployment

[4] See **?** for a related application.

[5] A measure more consistent with the model presented in Section 3 would be the vacancy-to-unemployment ratio, or labor market tightness. Because a consistent measure of vacancies (as defined by the model) is not available until late 2000 (e.g. the Job Openings and Labor Turnover Survey— JOLTS), the estimation strategy uses only unemployment rates as cyclical indicators.

rates $u_0 \times u_t$, and year fixed-effects. Demographics, education, and experience are controlled for using indicators for race, educational attainment, and a cubic in labor market experience.

Key to the analysis is the coefficient on the initial unemployment rate u_0. The models of **?** and MS suggest that recessions (expansions) decrease (increase) allocative efficiency by moving fewer (more) workers towards jobs where they are more productive; i.e. a sullying effect. Thus, a positive coefficient on u_0 provides evidence in favor of a sullying effect where matches observed in recessions (expansions) are low (high) quality and short (long) in duration. If the sullying effect operates through the on-the-job search channel, then restricting the estimation of $\lambda(t|\mathbf{X}(t))$ to matches formed from job-to-job transitions should also lead to a positive estimated coefficient on u_0. In contrast, the cleansing effect of recessions (expansions) means low quality matches are destroyed (created). Additionally, new matches must now meet a higher (lower) minimum productivity threshold. A negative coefficient on u_0 would be suggestive of this effect. In recessions (expansions), more (less) stringent hiring requirements lead to high (low) quality matches that are longer (shorter) in duration.

Next, the current unemployment rate, u_t, captures how current conditions affect hazard rates independent of match quality. Like **?** the current unemployment rate is introduced in a non-linear manner, u_t^2, to help distinguish between the pro-cyclicality of voluntary switches and the counter-cyclicality of involuntary separations or layoffs.

Meanwhile, the interaction term, $u_0 \times u_t$ captures how match quality and current conditions interact. Including this variable follows the implications of MS and **?**: low quality matches are formed in expansions only to be destroyed as soon as conditions deteriorate.

Lastly, year-fixed effects capture the unbalanced nature of expansions versus recessions. Expansions are more frequent and longer lasting than recessions, and by construction more matches will be observed in expansions. Not accounting for this unbalanced nature will tend to bias the effects of improved reallocation toward expansions, and thus finding a significant sullying channel.

Beyond including measures of cyclical conditions and observable individual characteristics, Equation 1 is estimated using individual fixed-effects. Controlling for fixed-effects alleviates biases arising from unobserved fixed heterogeneity. Additionally, each individual's spells are weighted by the inverse of the number of spells observed for them normalized by the number of survey waves to which they respond.[6] This allows each individual to contribute equally in the likelihood estimation of Equation 1.

The proposed sampling scheme uses as much data as possible. Alternatively, one could use only the first two observed spells for each individual, as suggested by **?**, or use one randomly chosen spell as suggested by **?**. These sampling schemes, in general, will lead to less efficient estimation.[7]

1.2 Data

The data used in this study come from the National Longitudinal Survey of Youth (NLSY), survey years 1979 through 2010. The NLSY is a nationally representative sample of 12,686 young men and women who were 14-21 years old when first interviewed in 1979. Interviews were conducted annually through 1994 and biennially thereafter.

The NLSY has important advantages over other surveys for studying match duration. Compared to address based surveys, such as the Current Population Survey (CPS), individuals do not drop out of the sample following a change in geographical location, which may be highly correlated with job duration. During each interview participants report information for up to five jobs that can be linked across consecutive interviews. Thus, the NLSY's format allows for more consistent construction of duration variables when compared to surveys such as the Panel Study of Income Dynamics (PSID).[8] Importantly, the

[6] Normalizing by the number of survey waves helps distinguish between individuals who report few long duration jobs lasting over several years versus individuals who report few jobs because of attrition. Additionally, individual fixed-effects are included in the estimation.

[7] **?** finds that under conditions of moderate censoring (e.g. the majority of individuals experience at least two events) the fixed-effects estimator is nearly always better than the conventional partial likelihood estimator when applied to repeated events with unobserved heterogeneity.

[8] See **?** for an in depth discussion of the issues when measuring job tenure in the PSID.

NLSY should capture short matches in between interview dates. Lastly, the NLSY has a much longer panel dimension in comparison to other longitudinal surveys such as the Survey of Income and Program Participation (SIPP), which only follows individuals for four years.

Following **?**, the sample is restricted to males from the cross-sectional samples and only job spells that start when the individual is at least 18 years old and not in school are included. Individuals must report valid wages and work at least 15 hours per week. Spells that end prior to 1979 or lasting less than a month are dropped. Unlike **?**, all spells of an individual are considered, rather than restricting the sample to a single random spell per individual. Hence, the sample not only covers more years but also more information per individual.

To construct the main sample of jobs and their respective durations, data from the Employer Roster Survey is used. Using the variables that contain the start and stop dates for each job report, the start of the match is defined as the week when the job is first recorded. The end of the match is defined as the week when the job is last linked.[9] Gaps within the duration of a match are ignored. This distinguishes this paper's measure of match duration to continuous tenure on the job. Given an ultimate interest in match quality, this broader definition of duration is used as repeated meetings between a worker and firm likely reflect a good match.

To facilitate comparison with the theoretical model outlined later, the sample is restricted to primary jobs. Specifically, spells that are contained within the duration of another job are dropped. Job-to-job transitions are defined as whenever the worker was non-employed for at most two weeks in between jobs and the previous job ended because of a quit.

All jobs satisfying the previous requirements are used in the estimation. The resulting sample consists of 9,025 spells from 1,310 individuals, or 6.9 spells per individual on average. Table 1 presents summary statistics for the main sample. For this sample, the average match lasts 2.2 years, which underscores the importance of using a dataset with a long panel dimension. The

[9]The duration of a job is defined as right-censored whenever the individual is currently working at the job during the time of the interview when the match is last reported.

Table 1. Summary statistics

Average job duration (in months)	26.27746
Average age when job starts (age_0)	27.78387
Average unemployment rate when job starts (u_0)	6.661848
Average current unemployment rate (u_t)	5.960606
% jobs that end in quit	62.89992
% non-white	21.87111
% less than high school	20.84743
% high school	47.08919
% some college	17.55385
% college or more	14.50952
# individuals	1,310
# spells	9,025
Average spells per individual	6.889313

Note: Source NLSY 1979-2010 men in cross-sectional samples.

majority of individuals are white and have at least a high school degree. Note, this sample excludes individuals with only one spell as fixed-effects cannot be estimated with such individuals. It also excludes spells where the reason the previous job ended is unknown. Appendix B shows the main results do not change when the sample is expanded by relaxing this last restriction.

1.3 Representativeness

It is important to note that the analysis sample is not representative of the entire U.S. population as it follows men of a particular cohort. In particular, this cohort is young in the late 80s and early 90s, which are relatively tranquil times, and more established in their career paths in the 2000s, which includes two recessions including the Great Recession. Since job mobility declines with age, results based on this NLSY sample may imply less cyclical variation in job duration (and hence match quality) compared to a more representative sample of individuals over the same period.

2 Empirical results

This section begins by presenting the main empirical results of the paper, which confirm that recessions are associated with lower match quality. The second part of this section shows this result is robust to the inclusion of starting wages, suggesting the procyclicality of match quality (at least initially) appears not to be fully internalized by the market.

2.1 Baseline results

Table 2 presents the results from estimating the hazard model and show the importance of initial conditions for explaining variation in job duration. The first column shows the initial national unemployment rate, u_0, has a positive and statistically significant effect on the hazard rate. Consistent with a sullying effect, matches starting in recessions are of shorter expected duration. Using a smaller sample and narrower time frame, **?** estimates a coefficient on the initial national unemployment rate of 0.0497, which is very similar to the coefficient presented in Column 1. Importantly, though, this column does not account for individual unobserved heterogeneity in spite of using multiple spells per individual. Column 2 adds worker fixed-effects to the estimation and shows that accounting for unobserved fixed heterogeneity across individuals raises the size and significance of the coefficient on u_0. Accounting for this heterogeneity also makes the coefficient on the interaction term significant (only at the 5% level). The negative sign on the interaction term suggests the sullying effect is dampened as aggregate conditions worsen. For future reference, Column 2 will be referred to as the preferred specification.

To illustrate the magnitude of the estimated coefficient on u_0, the most significant coefficient, Figure 1 presents the survivor functions implied by Column 2. The solid line represents the survivor function for a white male, with a high school diploma and 12 years of labor market experience, holding a job that starts when the unemployment rate is at its sample mean and setting the current unemployment rate to its sample mean. The "- -" line represents the survivor function for the same person, only now facing a more favorable ini-

Table 2. Hazard estimates: NLSY sample

	(1)	(2)	(3)	(4)
	all	all	from NE+layoff	from E+quit
u_0	0.04225*	0.12185***	0.14785***	0.09975***
	(0.022)	(0.026)	(0.056)	(0.036)
u_t	-0.04078	-0.05069	-0.13461	-0.00279
	(0.058)	(0.064)	(0.120)	(0.090)
u_t^2	0.01323	0.01086	0.04121	0.04093*
	(0.013)	(0.016)	(0.030)	(0.025)
$u_0 \times u_t$	-0.00375	-0.02571*	-0.05431**	-0.03393*
	(0.011)	(0.014)	(0.026)	(0.020)
Worker fixed-effects	–	yes	yes	yes
No. of obs.	292,032	292,032	101,693	190,339

Note: u_0 denotes the unemployment rate at the time when the match begins. u_t denotes the time-varying current unemployment rate. $u_0 \times u_t$ denotes the interaction between the initial and current unemployment rate. Standard errors are clustered by time and appear in parentheses. Columns 1 and 2 are estimated with all spells. Column 3 restricts the estimation to spells where the individual was previously not employed and was laid off from her most recent job. Column 4 only uses spells where the individual was previously employed and quit her most recent job. Regressors not reported: cubic in experience, year fixed-effects, and indicators for race, less than high school education, some college, and college graduate (or more). $+, *, **, ***$ indicate statistical significance at 10%, 5%, 1%, and 0.1% levels.

tial unemployment rate that is one standard deviation (1.5 percentage points) below its mean (i.e. boom). The "-.." line is the survivor function when the initial unemployment rate is one standard deviation above its mean (i.e. recession). The solid line implies the median match lasts 25 months under standard conditions. Median expected duration falls to 19 months (24% decrease) in a recession and rises to 34 months (36% increase) in a boom.

The previous estimates suggests recessions decrease match quality, but do not provide evidence whether prior employment status matters. Columns 3 and 4 consider, respectively, matches formed with previously nonemployed workers (and who were laid off from their last job) versus those switching between jobs (and who quit their last job). The positive and statistically significant coefficient on u_0 in Columns 3 and 4 suggests recessions are associated

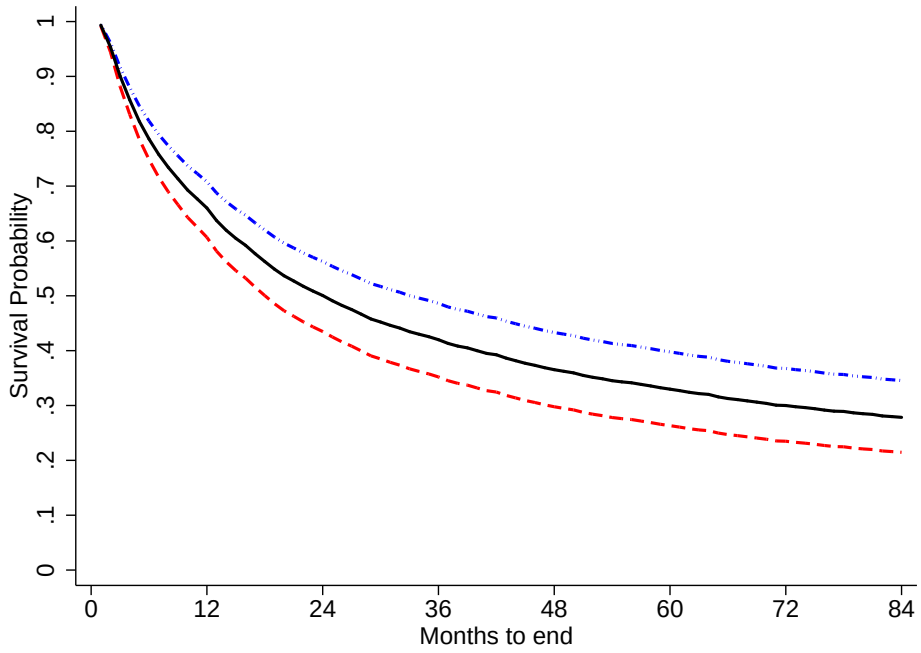

Figure 1. Survivor functions for the initial unemployment rate at its mean (—) and one standard deviation increase/decrease (- -/- ..).

with lower job duration regardless of prior employment status. Quantitatively, the effect is slightly stronger for matches where the worker was previously employed: a one standard deviation decrease in u_0 increases duration of matches formed by previously employed workers from 44 to 58 months (32% increase), whereas the duration of matches formed by previously nonemployed workers rises from 14 to 18 months (29% increase).

The remaining rows in Columns 3 and 4 show the interaction of initial and current conditions is most significant for matches with previously nonemployed workers. In other words, a further deterioration in current economic conditions tends to offset the negative impact that initial aggregate conditions have on duration, particularly for previously nonemployed workers.

2.2 Are starting wages important?

Overall, the previous results provide evidence of procyclical match quality. Given the large body of evidence showing individual wages are procyclical, it is natural to ask whether the cyclicality of match quality found here is internalized through wages?[10]

Table 3 repeats the estimation of Equation 1, but includes the initial log real wage lnw_0 as an explanatory variable. The first column of this table replicates the **?** finding that initial wages make the estimated coefficient on the initial unemployment rate, u_0, decline in magnitude and become insignificant. The second column reveals that once individual fixed-heterogeneity is taken into account initial wages still matter, but so does the initial unemployment rate. The estimated coefficient is now 0.11 compared to 0.12 as reported in Table 2. Additionally, the inclusion of initial wages makes the interaction term, $u_0 \times u_t$, increase in size and significance. Columns 3 and 4 confirm that the coefficient on u_0 is robust to the inclusion of initial wages regardless of prior employment status. Controlling for initial wages, however, changes the quantitative importance of u_0 for the previously nonemployed versus employed. Column 3 suggests a one standard deviation decrease in u_0 increases median duration for the previously nonemployed from 15 to 21 months (40% increase). Meanwhile, Column 4 suggests median duration of the previously employed increases from 36 to 48 months (33% increase), given a similar change in u_0. Hence, conditional on initial wages, initial aggregate conditions tend to matter more for matches with previously nonemployed workers.

Focusing on the direct impact that initial wages have on duration, the estimates from Columns 3 and 4 suggest starting wages are more important for the duration of matches with previously employed workers. While a one standard deviation increase in initial wages raises the duration of a match with a previously employed worker from 36 to 100 months (nearly a 200% increase), it only raises the duration of a match with a previously nonemployed worker from 15 to 20 months (one-third increase). The large impact that initial wages

[10] See **?** for a summary of the evidence on the cyclicality of wages.

Table 3. Hazard estimates: NLSY sample

	(1)	(2)	(3)	(4)
	all	all	from NE+layoff	from E+quit
u_0	0.02568	0.11331***	0.17457***	0.09809***
	(0.023)	(0.027)	(0.059)	(0.037)
u_t	-0.03242	-0.03046	-0.15289	-0.02190
	(0.061)	(0.072)	(0.133)	(0.104)
u_t^2	0.01114	0.00556	0.05553*	0.05709**
	(0.014)	(0.017)	(0.032)	(0.027)
$u_0 \times u_t$	-0.00900	-0.02981**	-0.07338***	-0.04145*
	(0.011)	(0.014)	(0.028)	(0.021)
lnw_0	-0.67278***	-0.77970***	-0.43265***	-0.93258***
	(0.038)	(0.053)	(0.093)	(0.086)
Worker fixed-effects	–	yes	yes	yes
No. of obs.	277,811	277,811	96,941	180,870

Note: u_0 denotes the unemployment rate at the time when the match begins. u_t denotes the time-varying current unemployment rate. $u_0 \times u_t$ denotes the interaction between the initial and current unemployment rate. lnw_0 denotes the initial log real wage. Standard errors are clustered by time and appear in parentheses. Columns 1 and 2 are estimated with all spells. Column 3 restricts the estimation to spells where the individual was previously not employed and was laid off from her most recent job. Column 4 only uses spells where the individual was previously employed and quit her most recent job. Regressors not reported: cubic in experience, year fixed-effects, and indicators for race, less than high school education, some college, and college graduate (or more). $+, *, **, ***$ indicate statistical significance at 10%, 5%, 1%, and 0.1% levels.

have on matches with previously employed workers may be consistent with a job ladder model: job-to-job transitions allow employed individuals to reach better paying and longer lasting matches.

3 Quantitative importance

To provide a more structural interpretation to the empirical results from the previous section, this section asks what estimation results would be recovered if the true data-generating process is the model of directed on-the-job search of MS. Aside from providing a theoretical foundation to the results from Sec-

tion 2, the model also provides quantitative conclusions regarding the relative importance of the cleansing versus sullying effects of recessions. This section begins with a brief outline of the model. Next, the model-generated empirical results are presented, followed by a discussion of the quantitative implications.

3.1 Model outline

In the MS model, search is directed—as in **?** and **?**—rather than random. Firm-worker matches differ in their quality, which is constant over the duration of the match. Firms choose how many and what type of vacancies to create, while workers choose what type of vacancies to search. Importantly, workers can search off- and on-the-job and meet with vacancies according to a constant returns to scale matching function. Upon meeting, a worker and a firm observe a signal about the idiosyncratic productivity of their match. If the signal meets the conditions pre-specified by the vacancy's type, the worker and the firm begin to produce and, eventually, observe the actual quality of their match. If the signal does not meet those conditions, the worker returns to her previous employment position. The current analysis only considers the case where the signal is completely uninformative; i.e. matches are *experience goods*.[11]

A distinguishing feature of the MS model that makes it ideally suitable for the present analysis is that the equilibrium is block recursive. Agents' value and policy functions depend on the aggregate state of the economy only through the realization of the aggregate shock, and not through the entire distribution of workers across employment states. Specifically, the job destruction threshold, below which matches are endogenously destroyed, only depends on the aggregate shock. The same is true for the tightness of the sub-markets where workers search for new jobs.

With the structure of the model in hand, the cleansing and sullying effects are more precisely defined. In the MS model, a negative (positive) productivity shock raises (lowers) the job destruction threshold and hence low

[11] If the signal is completely informative, i.e. matches are *inspection goods*, the effect of aggregate productivity shocks on the labor market is negligible. Thus, this scenario is not considered.

quality matches are destroyed (created)—a cleansing effect. Meanwhile, a sullying effect arises because a negative (positive) productivity lowers (raises) the probability that a worker in a low quality match finds a better job.

3.2 Results

This section begins by confirming the NLSY results using model generated data from the MS model. Next, using the underlying distribution of the MS model, the cleansing effect is shown to dominate. Lastly, a discussion is provided of what the regressions are capturing.

3.2.1 Comparison with NLSY

This section presents the results of estimating Equation 1 on panel data generated from simulating the MS model. Details on how the panel was constructed appear in Appendix C.

Table 4 presents the estimation results using model generated data. All estimations include individual fixed-effects, mirroring the preferred specification in Table 2. The first column of the table uses all observations, while columns 2 and 3 distinguishes spells by prior employment status of the worker.

The first column of Table 4 reveals that model based data delivers results similar to those using NLSY data. Indeed, the coefficient on initial unemployment rate, u_0, is positive and significant, implying that in model generated data recessions are associated with lower job duration, or equivalently, expansions are associated with higher job duration.

Quantitatively, the coefficient on u_0 is slightly larger than the one in Table 2. This is partly due to scale reasons. In particular, the standard deviation of unemployment in model generated data is much smaller than in the NLSY sample (0.22 versus 1.6 percent, respectively), hence the coefficient on u_0 will tend to be larger in order to explain similar changes in job duration. Additionally, the model has only one shock, so variables will tend to be more correlated compared to the data where multiple shocks are at play. Finally, the MS model is calibrated to match moments from a representative sample

of the U.S., while the NLSY sample is more selected.

To provide quantitative context to this coefficient, a one standard deviation decrease in u_0 increases median job duration from 207 to 224 months; i.e. an 11% increase. Recall, the NLSY estimate implies a 36% increase. The key take-away, however, is that both sets of results suggest match quality is procyclical.

The second and third columns of Table 4 show that match quality rises in booms irrespective of the worker's prior employment status, which is also consistent with the findings from the NLSY sample. Quantitatively, though, these results imply relatively more cyclical responsiveness from matches with previously unemployed workers. A one standard deviation decrease in u_0 increases median job duration for matches of individuals who were previously unemployed from 188 to 216 months (i.e. a 15% increase). Meanwhile, a standard deviation decrease in u_0 raises median job duration for matches of individuals who were previously employed from 222 to 244 months—a 10% increase. Looking at the other coefficients, only the linear term of the contemporaneous unemployment rate, u_t, is significant. As discussed below, this appears to be capturing a cleansing effect where high quality matches, created in recessions, survive adverse contemporaneous conditions.

3.2.2 Which effect dominates?

The previous results show that the positive and significant effect of initial aggregate conditions on expected job duration (suggestive of a sullying effect) is consistently reproduced in model simulated and actual data. However, because the MS model also has an active cleansing effect, the quantitative question of which effect dominates remains?

Table 5 presents summary statistics of the match quality distributions depending on whether the aggregate technology shock is at its mean or one standard deviation above its mean (i.e. expansion). The first row of this table reveals that expansions are associated with lower average match quality. Like MS report, a one percentage point increase in aggregate productivity decreases average match quality by roughly 0.3 percent—in other words, the cleansing

Table 4. Hazard estimates: model generated data

	(1)	(2)	(3)
	all	from U	from E
u_0	0.22505***	0.34012***	0.24294**
	(0.076)	(0.108)	(0.102)
u_t	-0.04006	-0.21117**	0.00134
	(0.053)	(0.083)	(0.076)
u_t^2	-0.06832	-0.07095	0.05992
	(0.185)	(0.306)	(0.248)
$u_0 \times u_t$	0.17075	-0.18966	0.17652
	(0.170)	(0.288)	(0.229)
Worker fixed-effects	yes	yes	yes
No. of obs.	1,786,738	837,752	948,986

Note: u_0 denotes the unemployment rate at the time when the match begins. u_t denotes the time-varying current unemployment rate. $u_0 \times u_t$ denotes the interaction between the initial and current unemployment rate. Standard errors are clustered by time and appear in parentheses. Column 1 is estimated with all spells. Column 2 restricts the estimation to spells where the individual was previously unemployed. Column 3 only uses spells where the individual was previously employed. Regressors not reported: year fixed-effects. $+, *, **, ***$ indicate statistical significance at 10%, 5%, 1%, and 0.1% levels.

effect dominates. Importantly, the cleansing effect manifests itself near the bottom of the quality distribution. For example, the 1st percentile declines by 2%, while the 5th percentile declines by 30%. Thus, while the baseline regressions from the previous section detect a sullying effect, the cleansing effect promotes enough creation at the low end of the quality distribution that average match quality falls.

3.2.3 Highlighting the mechanisms at play

Aside from reaffirming the NLSY based estimation and providing quantitative conclusions, the structure of the model provides deeper insights as to what the empirical results are capturing. Using the productivity distribution and policy functions of the model, one can categorize matches into: *high quality matches*, that only end because the worker executes a job-to-job transition or

Table 5. Match quality distributions by aggregate shock

	(1)	(2)	
	Average	High	% difference (2)-(1)
Mean	.2826628	.2819353	-0.25737
Std. dev.	.1385793	.1395819	0.72348
1st percentile	-0.22066	-0.22538	-2.13904
5th percentile	0.010222	0.00713	-30.24848
50th percentile	0.28763	0.28763	0.00000
95th percentile	0.47649	0.47649	0.00000
99th percentile	0.5316	0.5316	0.00000

Note: Statistics are derived from simulated data used in Table 4. Average column is based on data when the aggregate shock is at its unconditional mean. High column is based on data when the aggregate shock is one percentage point above the mean.

because of an exogenous separation; and *low quality matches*, that end mostly through endogenous destruction leaving the worker unemployed.[12]

Table 6 uses this quality classification to reveal the positive coefficient on u_0 is capturing a sullying effect. Restricting the estimation to high quality matches (Column 2), which are subject to the sullying effect, renders a positive and highly statistically significant coefficient on u_0—much like the baseline results imply. In other words, in recessions reduced reallocation through job-to-job transitions reduces quality within this group of higher quality jobs.

In sharp contrast, the same estimation restricted to low quality matches delivers results more in line with a cleansing hypothesis. As can be seen in Column 3, the estimated coefficient on u_0 is negative and significant (at the 5% level). Thus, in recessions the higher threshold for endogenous job destruction actually raises the duration of matches at the lower end of the quality distribution.

The remaining coefficients in Column 3 provide greater insights on the

[12]In the MS model it is true that some low quality matches end in expansions because the worker executes a job-to-job transition. Quantitatively, however, the reduction in the probability of endogenous job destruction is much larger, compared to the increased probability of a job-to-job transition. Hence for these matches, changes in the job destruction threshold are more important for understanding how duration changes when an expansion begins.

Table 6. Hazard estimates by match quality: model generated data

	(1)	(2)	(3)
	all	high quality	low quality
u_0	0.22505***	0.16152***	-2.17276*
	(0.076)	(0.054)	(1.167)
u_t	-0.04006	0.01460	2.28949**
	(0.053)	(0.054)	(1.147)
u_t^2	-0.06832	-0.17576	-11.82356**
	(0.185)	(0.190)	(5.103)
$u_0 \times u_t$	0.17075	0.21474	11.75126**
	(0.170)	(0.162)	(5.105)
Worker fixed-effects	yes	yes	yes
No. of obs.	1,786,738	1,757,305	29,433

Note: u_0 denotes the unemployment rate at the time when the match begins. u_t denotes the time-varying current unemployment rate. $u_0 \times u_t$ denotes the interaction between the initial and current unemployment rate. Standard errors are clustered by time and appear in parentheses. Column 1 is estimated with all spells. Column 2 is estimated only on high quality matches, while Column 3 is estimated only on low quality matches (as defined in the text). Regressors not reported: year fixed-effects. $+, *, **, ***$ indicate statistical significance at 10%, 5%, 1%, and 0.1% levels.

duration of low quality matches. A higher contemporaneous unemployment rate raises the hazard of a low quality match ending, but in a non-linear fashion. At the sample mean the net effect of u_t is actually negative; only for very low values does it turn positive. Thus, the duration of low quality matches (meeting the more stringent requirements of a recession) rises as conditions deteriorate, reflecting the fall in the job-to-job transition rate. Lastly, the positive interaction term between initial and current conditions implies that the cleansing effect of recessions (the negative coefficient on u_0) increases as the current unemployment rate rises; i.e. the endogenous job destruction threshold rises as aggregate productivity continues to fall.

4 Conclusion

This paper evaluates empirically and theoretically how match quality varies over the business cycle. Panel data from the NLSY confirms previous observations that expansions are associated with higher match quality. The current findings are robust to controlling for worker composition and suggest that quality rises more for matches where the worker was previously employed. Viewing these findings through the lens of Menzio and Shi's (2011) model of directed on-the-job search confirms the empirical evidence: more job-to-job transitions in expansions increase the frequency by which workers move into better matches. On net, however, the calibrated model implies expansions are associated with lower average quality. This occurs because of lower endogenous job destruction at the bottom of the quality distribution, which offsets quality improvements at the top. Overall, these findings underscore the importance of using a model to understand the empirical observations.

An important component missing from the present analysis is the firm dimension. The empirical work of **?** highlights the importance of firm heterogeneity for understanding worker flows over the business cycle. Future research could extend the present analysis with matched employer-employee data and a model along the lines of **?** to provide a richer description of cyclical variation in match quality and how it manifests itself in the data.

Online Appendix–*not for publication*

A Data

The data used in the paper comes from the NLSY survey years 1979-2010. The Employer History Roster is used to compile all variables of interest.

- ***Wages.*** Wages are calculated from the Employer roster variables *EMPLOYERS ALL TIMERATE* , *EMPLOYERS ALL PAYRATE*, and *EMPLOYERS ALL HRLY WAGE*. All wages are deflated by the CPI-U all urban consumers index.
- ***Hours.*** Hours are calculated from the Employer roster variables *EMPLOYERS ALL HOURSWEEK* and *EMPLOYERS ALL HOURSDAY*.
- ***Start and stop dates.*** Dates are calculated from the Employer roster variables *EMPLOYERS ALL STARTDATE ORIGINAL*, *EMPLOYERS ALL STOPDATE*, and *EMPLOYERS ALL STARTWEEK*.
- ***Layoffs and quits.*** Layoffs and quits are identified using the variable *EMPLOYERS ALL WHYLEFT*.

B Additional results

Table 7 presents results of estimating Equation 1 on the less restricted NLSY sample where the reason for leaving the prior job is not required. Qualitatively, the results are similar and still suggest the importance of initial conditions for understanding job duration.

Table 7. Hazard estimates: extended NLSY sample

	(1)	(2)	(3)	(4)
	all	all	from NE	from E
u_0	0.03737**	0.08124***	0.08585**	0.06860**
	(0.017)	(0.020)	(0.036)	(0.029)
u_t	-0.02569	-0.01349	-0.00668	0.01698
	(0.047)	(0.047)	(0.072)	(0.071)
u_t^2	0.00299	0.00871	0.00204	0.00750
	(0.010)	(0.011)	(0.017)	(0.017)
$u_0 \times u_t$	-0.00014	-0.01009	-0.01398	-0.00893
	(0.009)	(0.009)	(0.015)	(0.015)
Worker fixed-effects	–	yes	yes	yes
No. of obs.	485,562	485,562	218,051	267,511

Note: u_0 denotes the unemployment rate at the time when the match begins. u_t denotes the time-varying current unemployment rate. $u_0 \times u_t$ denotes the interaction between the initial and current unemployment rate. Columns 1 and 2 are estimated with all spells. Column 3 restricts the estimation to spells where the individual was previously not employed. Column 4 only uses spells where the individual was previously employed. Standard errors are clustered by time and appear in parentheses. Regressors not reported: cubic in experience, year fixed-effects, and indicators for race, less than high school education, some college, and college graduate (or more). $+, *, **, ***$ indicate statistical significance at 10%, 5%, 1%, and 0.1% levels.

C Simulation details

To construct the simulated sample of worker histories from the **?** model, the following procedure was used. From the steady state of the model, I draws (workers) were made from the distribution across match quality and employment states. Then, aggregate and individual histories were simulated for T periods using the equilibrium policy functions. Along the simulated path, employment status, duration, aggregate unemployment, and the state of aggregate technology were recorded. In practice, $I = 3,000$ while $T = 9,000$. To construct the main sample only the last 30 years of the simulation were used.