"The Intensity of Job Search and Search Duration"

R. Jason Faberman and Marianna Kudlyak

Online Appendix

Appendix A. Sample Construction

Our dataset contains daily proprietary confidential micro data on all applications sent on the SnagAJob website between September 2010 and September 2011. The dataset contains each applicant's demographic information on age, education, race, and gender, as well as geographic location at the zip code level. All individual data are anonymized. For vacancies, the dataset contains geographic location at the zip code level and the website's classification of each vacancy's occupation, and as noted in the main text, the SnagAJob "occupation" classification is an amalgamation of industry and occupation. We also have data on employment histories for nearly three-quarters of the job seekers in the sample.

A.1. Sample Restrictions

The raw data in our sample period contain over 46 million applicant-vacancy observations for 8.0 million job seekers and over 1.8 million vacancies. We remove observations with missing geography or occupation data (a negligible amount of observations) as well as applications to postings that are to "work at home" jobs or similar postings that advertise positions that are self-employment opportunities rather than a job opportunity with a particular employer. This procedure reduces the number of job seekers and vacancies in the data by a negligible amount, but it reduces the number of application observations by 14 percent. In some cases, we observe repeat applications by the same individual to the same job posting. We delete repeat applications that are within one week of each other on the presumption that these are

the result of an applicant error. We retain repeat applications further than one week apart on the presumption that applicants may find it worthwhile to apply again if they observe that the job opening is still posted (and, presumably, unfilled). We also exclude applicants with missing education and gender information, and restrict our attention to applicants aged 16 to 75 years. These restrictions remove a negligible amount of individuals. We focus our analysis on the behavior of job seekers who register on the website after the beginning of our sample, September 1,' 2010, to ensure that we can follow applicants from the start of their search spell on the website. This reduces the number of applications in our sample by about 31 percent. As noted in the main text, though, we use information on the behavior of all job seekers in our sample to generate estimates of vacancy characteristics, such as total applications received and the start and end dates of each vacancy. Finally, we drop observations during the first week of our sample from our analysis because we cannot differentiate between new and incumbent vacancies during that week (a necessary criterion for one of our tests of stock-flow search behavior). We do, however, use the first week of data when calculating the vacancy and job seeker statistics. When we aggregate the sample to the job seeker-week level, we obtain 10.11 million observations, and observe job seekers sending at least one application for 7.65 million of these observations.

A.2. Employment History Identification

Our dataset contains information from self-reported employment histories for a majority of job seekers on the website. The website allows job seekers to provide information about their current and previous jobs. The information typically includes occupation (based on the SnagAJob categorization), a more-detailed job title, the job's start and end dates, an indicator for whether the job seeker currently holds the reported job, and additional information on the type of job (i.e., part-time, temporary or seasonal, etc.) The employment history records have typical measurement issues associated with self-reported data and the potential for non-random non-response (e.g., non-reporting of a recent job that one was fired from). To ensure that the timing of each reported job is consistent with the start of each

search spell, we use the start and end date information for each employment record to identify its status as a current or previous job.

Our main use of the employment history data is to identify a labor force status (i.e., employed or non-employed) for each job seeker reporting at least one employment record. We also extract any additional information on the type of employment and the job tenure of the job seeker's current or most recent job, with the latter measured using the reported starting and ending dates (or the date when the job seeker registered with the website, for jobs that are reported as current). We can identify a labor force status for about 72 percent of our sample. Of these, 77 percent report only one employment record. Another 20 percent report multiple employment records, but do so in a way that the timing across jobs allows a clean identification of labor force status. The remainder have multiple employment records either active at the start of the search spell (i.e., a multiple job-holder) or that terminate within the same month (multiple previous jobs). For these individuals, we identify a "main" job using a priority ranking based on full-time vs. part-time status and how recently each job started. We measure the tenure of their current or most recent job based on their "main" job.

We count an individual as employed if we identify one or more records as active at the start of search. Non-employed job seekers are those whose employment records identify a final employment record that ends before their search spell begins. About 3 percent of job seekers are identified as non-employed using information about their current "job" that suggests non-employment (i.e., job titles listing "volunteer," "student," "homemaker," or even "never had a job" in their employment record). Of those with a reported employment history, we identify roughly 33 percent of them as employed at the start of their job search spell.

Appendix B. Robustness of Spell Length Identification

Identifying the current and completed duration of a job seeker's search spell is crucial to our analysis. The micro data contain information about when a job seeker registers with the search engine as well as information on all applications sent by the job seeker during the sample period, which allows a clean identification to the start of a job seeker's search on the website. The data, however, do not contain information on why a job seeker stops applying to jobs on the website. They may have found a job (either through the website or through other job search methods), stopped searching on the website but continued search elsewhere, or stopped searching for a job altogether. A job seeker may also continue searching even after finding employment in cases where the new job reflects "underemployment" relative to the type of work a job seeker was seeking. In addition, job seekers often exhibit long spells of inactivity on the website and then resume sending applications again. During this period, a job seeker may have found work and become unemployed again, they may have become discouraged with the website and decided to use other methods for a while, or they may have temporarily stopped searching for work altogether. For example, Clark and Summers (1979) find that nearly half of all completed unemployment spells end with an exit from the labor force. They also may have been employed or out of the labor force, with only a marginal desire for (new) employment, and therefore searching intermittently.

Taking these considerations into account, we define an individual's first search spell on the website as the period between their first application date and any application date that is followed by more than five weeks of inactivity. The first application occurs within a week of registering on the website for 95 percent of our sample. For reasons we note in the main text, we consider an application after the five-week inactivity cutoff to be the beginning of a new search spell. In this appendix, we experiment with alternate cutoffs for identifying spell length. In the absence of this identification, results tend to be dominated by the fraction of job seekers who send zero applications, as the following figures show.

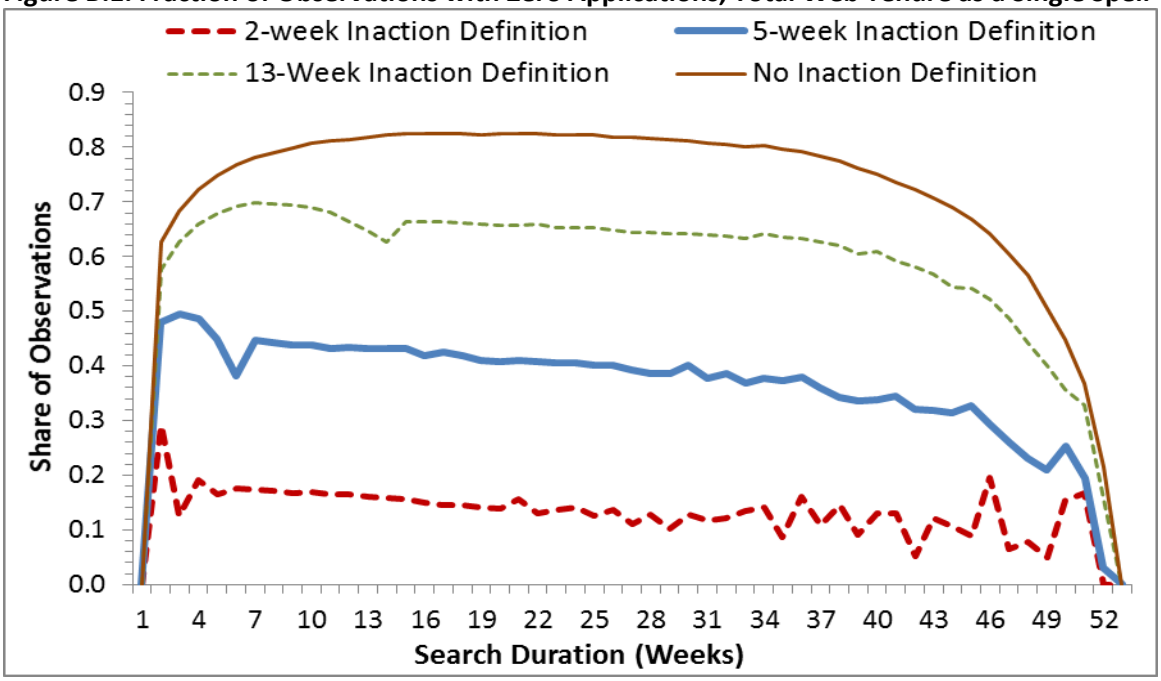
We find that, once we add the controls used in our main analysis, the differences due to different cutoff lengths essentially disappear. We replicate results from the main text using a two-week and 13-week cutoffs for spell identification. Figure B.1 shows the importance of using a plausible cutoff period. It plots the fraction of job seekers that sent no applications during each week of the search spell. It also includes the fraction estimated if we impose no cutoff at all. As one can see, the share of each week dominated by inactivity rises the longer the cutoff, and when there is no cutoff, nearly 80 percent of job seeker-week observations have no applications sent between 2 and 11 months of search. Under the five-week cutoff used in our analysis, the share never rises above 50 percent and declines steadily thereafter.

Figure B.2 replicates Figure 5 in the text for the case where we use no cutoff to identify search spells. In this extreme case, completed spell length has essentially no relationship to applications sent per week because of the dominance of inactivity observed in Figure B.1. Finally, Figure B.3 replicates our regression analysis for four specifications taken from equation (1) in the main text using the three different cutoffs (two weeks, five weeks, and 13 weeks). Prior to adding all controls, there are large quantitative differences across the cutoff rules. When we add either spell length or job seeker fixed effects, however, the results are similar regardless of the cutoff used. In all cases, applications per week decline with search duration.

References

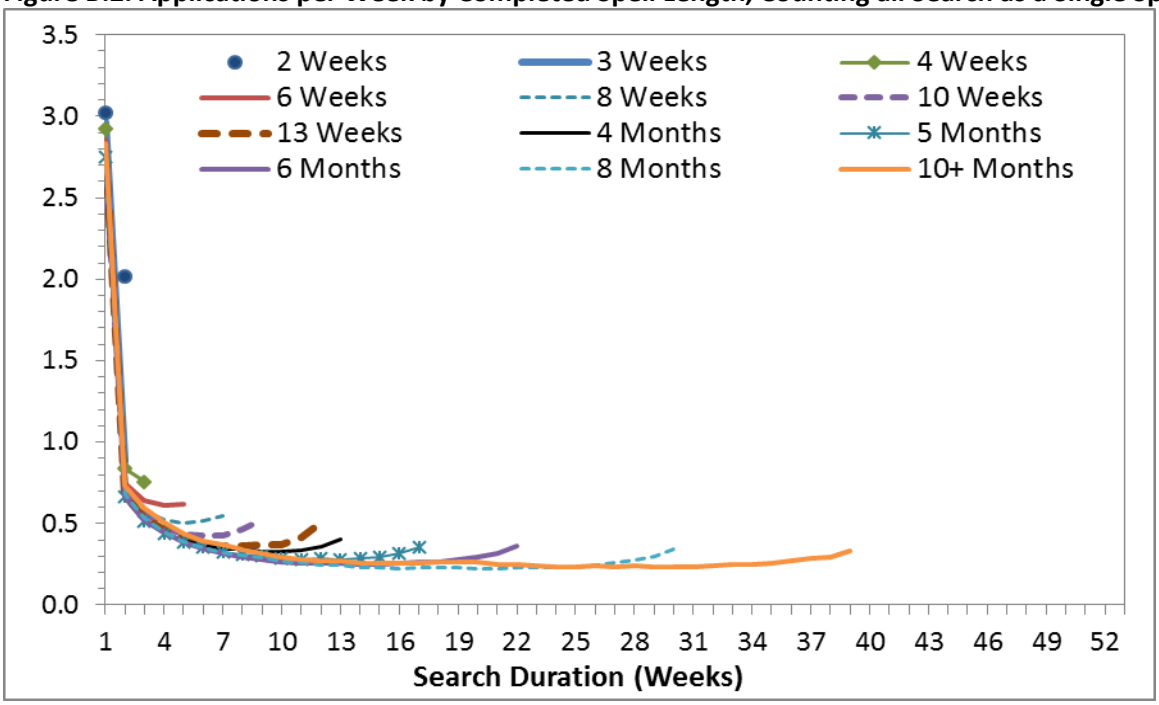
Clark, Kim B., and Lawrence H. Summers, 1979. "Labor Market Dynamics and Unemployment: A Reconsideration," Brookings Papers on Economic Activity, (1): 13-72.

Figure B.1. Fraction of Observations with Zero Applications, Total Web Tenure as a Single Spell



Note: Figure plots the fraction of jobseeker-week observations with zero applications sent that week, based on different assumptions on the end of a job seeker's search spell. Longer weeks of inaction reported refer to longer periods of continuous inactivity required to identify the end of a job seeker's search spell. Estimates use all job seeker-week observations for job seekers in our SnagAJob website sample.

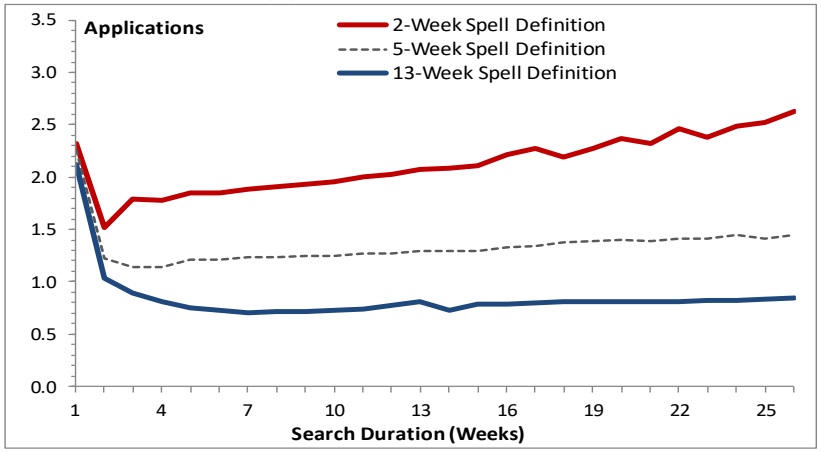
Figure B.2. Applications per Week by Completed Spell Length, Counting all Search as a Single Spell



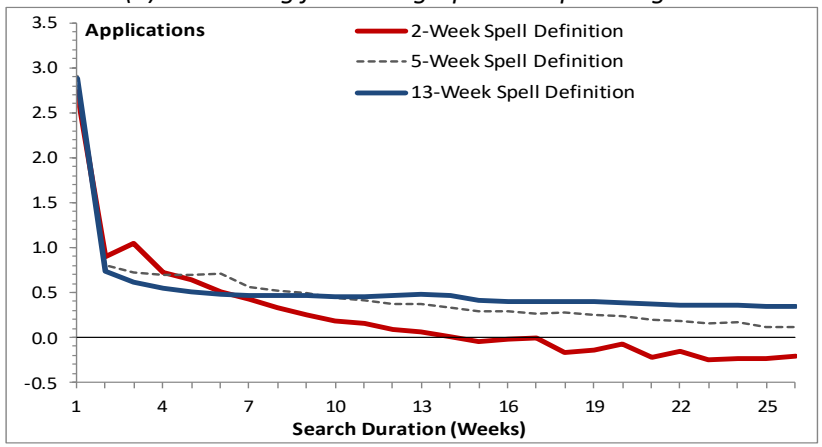
Note: Figure plots applications per week for job seekers of differing completed spell lengths, based on the assumption that all search on the website is contained within a single spell. Estimates use all job seeker-week observations for job seekers in our SnagAJob website sample.

Figure B.3. Estimated Relation between Applications and Search Duration under Alternate Spell Length Identification Criteria

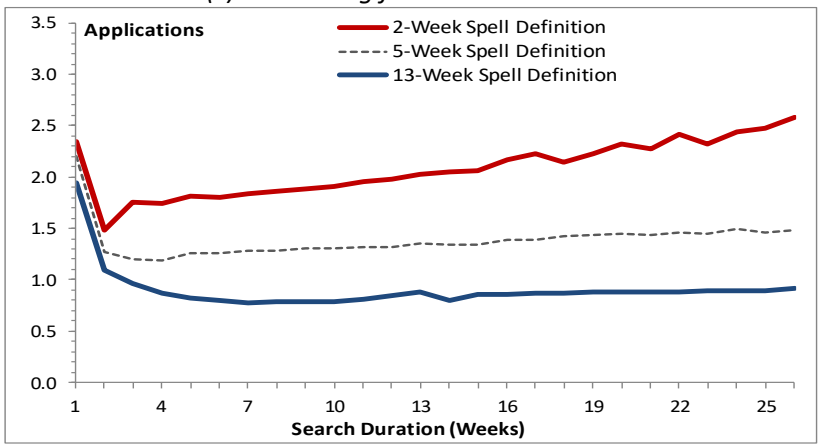
(a) Model without Controls



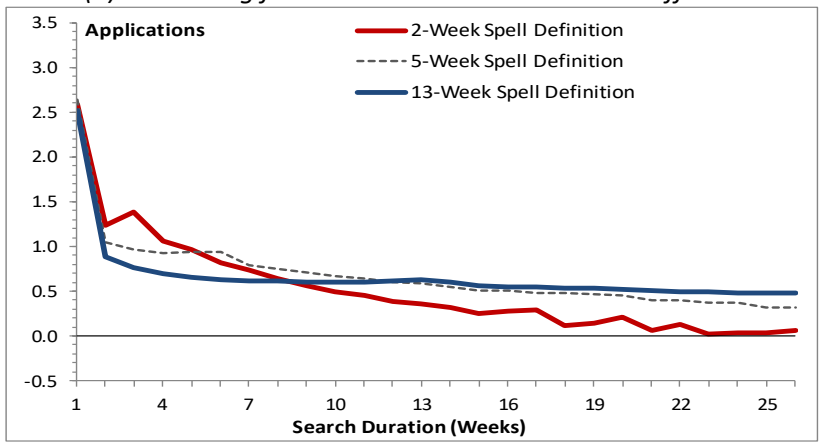
(b) Controlling for Demographics & Spell Length



(c) Controlling for Local Vacancies



(d) Controlling for Local Vacancies & Job seeker Effects



Notes: Panels depict the estimated relationship between applications sent per week and search duration under the four different regression specifications used in the analysis of the main text, using three different cutoff criteria to identify the end of a search spell: two weeks, five weeks, and 13 weeks of inactivity on the website. Estimates use all job seeker-week observations for job seekers in our SnagAJob website sample.

Appendix C. Simulation of Search Effort Behavior with Heterogeneity in Website Preferences

In Appendix D, we examine the robustness of our main results using a sample of "potential matches." We show that we obtain nearly identical results using the potential match sample as we do using the full sample in the main text. One may worry, however, that the potential match sample may still suffer from a spurious correlation between applications and search duration because of individuals who are only marginally interested in finding a job on the SnagAJob website. For example, we may obtain similar results in our full sample and potential match sample purely through the "luck" effect of stochastic job finding. Using a counterfactual simulation, we quantitatively evaluate such a possibility and show that pure luck in job finding cannot drive the results obtained from the potential match sample. Specifically, we show that if individuals only differed in their interested in using the SnagAJob website for job search, we should obtain very different results than what we do in either Appendix D or the main text.

We do our simulation using a model of job seekers who only differ in their preference for search on the website. We assume that there are N total job seekers registered on the website. A fraction θ of these job seekers is what we refer to as "marginally attached" to the website. That is, they search both on the website and through other methods (including, potentially, other job search websites). We set $\theta = 0.8$, which is roughly calibrated to the large amount of attrition we see within the first week of search.[1] Each job seeker sends s applications per week. To keep the exercise simple, we assume that the number of applications per week remains constant over the duration of search. We abstract from any heterogeneity or duration dependence that could lead to search intensity declining over time since we are focused on the effects of selection on our second empirical result (i.e., higher search effort by those

[1] The exit hazard after one week of search is 74.3 percent.

with longer completed search spells throughout). Job seekers who search exclusively on the website send all s applications on the website. Those who are marginally attached send a fraction α of their applications through the website and the remaining $(1-\alpha)s$ applications to job openings found outside of the website. In addition, marginally attached job seekers may quit the website entirely with probability $\rho(t)$, which we assume declines with search duration, t, given the sharp decline in job seekers observed in the data during the first few weeks of search. We also perform the simulation under the assumption of a constant quit rate, and report these results as well. All job seekers have the same probability f of having an application lead to a hire each period, regardless of whether the application was made on the website or elsewhere. Thus, the only heterogeneity among job seekers in the model is their preference for search on the website.

Given the model setup, job seekers can exit search on the website in one of three ways: 1) they can find a job on the website, 2) they can find a job through other means, or 3) they can quit searching on the website entirely. Those who are marginally attached to the website can exit through any of the three methods, but those who are committed to the website can only exit through the first method. We do not allow job seekers to quit search entirely, however. They can only change their method of search over time.

The model has three parameters, $\{s, \alpha, f\}$ and one function, $\rho(t)$, that we calibrate to the data. We assume that s equals the mean number of applications sent in the first week of search by applicants who completed spell lengths of at least 10 months. This is the highest amount of applications sent per week observed, on average, in the data, and is used since s represents the total number of applications sent using all methods in the model. We calibrate α using s and the model's expression for the expected total number of applications sent in the first week, $\theta\alpha s + (1-\theta)s$. We calibrate the job finding rate f to match the exit hazard of job seekers with completed spell lengths of six months or more. Given our assumption on the marginally attached, this exit hazard equals $1-(1-f)^s$. For the website quit

probability, we assume that $\rho(t) = \rho_0/t^{\rho_0+1}$, which allows it to decline with duration exponentially. We calibrate ρ_0 by equating the probability of exit after the first week to $1-(1-f)^s + \theta\rho(1)(1-f)^s$. We then run the model on 240,000 job seekers (roughly equivalent to 5 percent of our data sample), and use the results to generate the simulated versions of Figures 5 and 9 (i.e., search effort and duration by completed spell length). In the simulated data, the potential match sample is the subset of job seekers who find a job through an application on the website. This sample will include those who were committed to search on the website and those who were marginally attached but managed to find a job through the website anyway.

The results of the exercise are in Figure C.1. The left panel shows the simulated applications per week for the full sample of simulated job seekers (analogous to Figure 5 in the main text) and the right panel shows the simulated applications for those who found a job on the website (analogous to Figure D.3 in the following Appendix). The simulation shows clear differences in applications per week by spell length between the full sample and the simulated sample. These differences are concentrated among the short-duration job seekers. These job seekers send much fewer applications per week than long-duration job seekers in the full sample, but essentially the same amount of applications per week in the potential match sample.

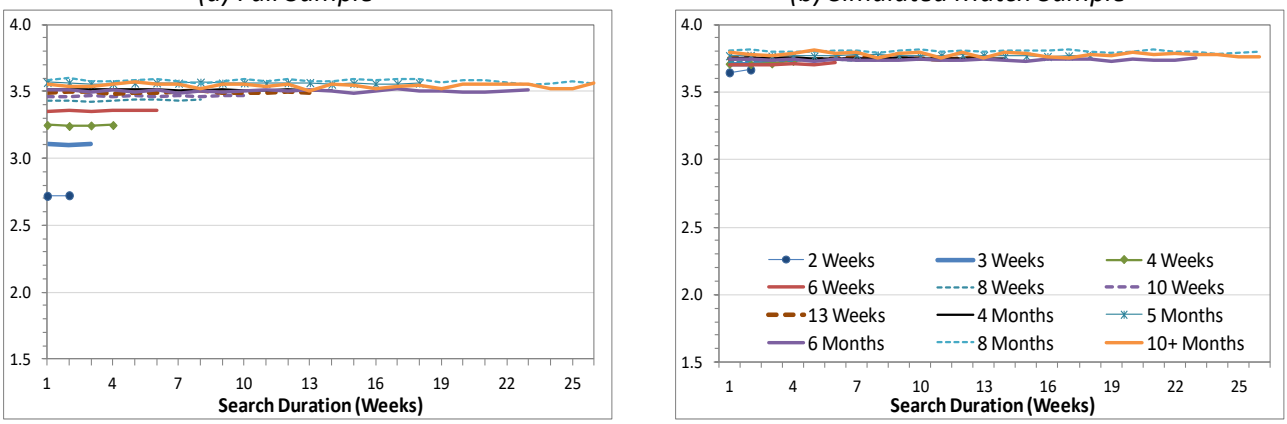
Intuitively, the marginally attached do not make up enough of the potential match sample to create much in the way of differences in application behavior (on the website) by completed spell length. Given the assumptions necessary for the initial fraction of the marginally attached to be consistent with the declining application rates observed in the data and an exponentially declining website quit rate, the marginally attached exit the website without finding a job and do so fairly quickly. This has two implications. First, relatively few of them find work on the website, leading to a small representation in the potential match sample. Second, many of them exit the website within the first few weeks (either

through attrition or job finding elsewhere). Thus, they are concentrated within the short-duration job seekers. As a result, there is only a small difference in application behavior between the long-duration and short-duration job seekers within the potential match sample when the only thing that differentiates job seekers is their preference for search on the website. We can relax the assumption that the quit rate declines exponentially with duration, which we do in Figure C.2. That is, we assume that $\rho(t) = \rho_0$. Under the assumption of a constant quit rate, the inverse of average spell length equals $1 - (1 - f)^s + \theta\rho_0(1 - f)^s$. All other calibrated parameters remain the same as in the text.

As one can see in the figure, the results are qualitatively similar to those in Figure C.1, though there is greater spread in the average number of applications by completed spell length. This is because there are relatively more of the marginally attached that remain on the website initially, but they are also relatively more likely to quit the website later in their search spell. Despite this, the subsample of potential matches (right panel) still shows considerably less dispersion across spell lengths than the full sample (left panel). If our results were driven only by individuals quitting the website, both versions of our simulation suggest that we should see less dispersion in the potential match sample when compared with the full sample. In the data, however, we find essentially the same patterns in both samples, which we interpret as rejecting the notion that our main results are driven by exits that are based primarily on tastes for search on the SnagAJob website.

Figure C.1. Simulated Application Behavior by Completed Spell Length, Heterogeneous Tastes for Website Search and Duration-Dependent Exit Rate

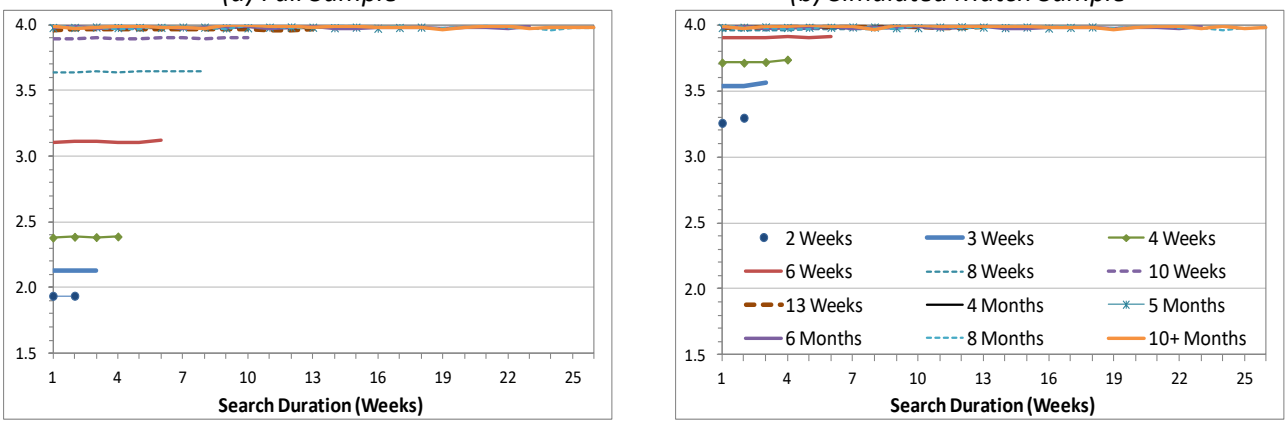
(a) Full Sample *(b) Simulated Match Sample*



Notes: Figure shows the estimated (unconditional) relationship between applications per week and duration of search separately for job seekers based on the total length of their search spell using a simulated sample of job seekers calibrated to the empirical distribution of job seekers in our website sample. The left panel reports the estimates for all simulated job seekers, while the right panel reports the estimates for simulated job seekers who found employment through the website. Only selected spell lengths are reported.

Figure C.2. Simulated Application Behavior by Completed Spell Length, Heterogeneous Tastes for Website Search and Constant Exit Rate

(a) Full Sample *(b) Simulated Match Sample*



Notes: Figure shows the estimated (unconditional) relationship between applications per week and duration of search separately for job seekers based on the total length of their search spell using a simulated sample of job seekers calibrated to the empirical distribution of job seekers in our website sample. The left panel reports the estimates for all simulated job seekers, while the right panel reports the estimates for simulated job seekers who found employment through the website. Only selected spell lengths are reported.

Appendix D. Additional Results

D.1. Comparability to Published Data

In this section, we examine how comparable the SnagAJob.com sample of job seekers is to the unemployed, and those in the labor force more broadly, as measured by the Current Population Survey (CPS). Much of our analysis is related to a companion review article (Faberman and Kudlyak, 2016). Table D.1 compares our job seeker sample to the CPS unemployed and labor force samples for respondents pooled between September 2010 and September 2011. Our sample has a disproportional number of younger, minority, and less-educated job seekers relative to the labor force in the CPS. The demographic composition of our sample is closer to the demographic composition of the pool of unemployed, though it still over-represents the young and those with at least a college degree. One key difference between our sample and the pool of unemployed in the CPS is that our sample has a majority of female job seekers (56.9 percent) while the unemployed in the CPS are mostly male (females constitute 43.7 percent).

Table D.2 compares the resulting distribution of search durations in our sample with the distribution of unemployment durations within the CPS. We use a cross section of job seekers during the CPS reference week of July 2011 for consistency with the CPS sample timing. As can be seen from the table, the average duration of the first search spell on the website is shorter than the duration of unemployment from the CPS. This occurs because the duration of the search on the website does not correspond to the notion of the duration of unemployment from the CPS. First, the job seekers in the SnagAJob sample include not only unemployed but also the employed and those who could have reported themselves as out of the labor force but still searched for work (e.g., retired individuals). Second, the unemployed job seekers might begin searching on the website a few weeks into their unemployment spell. Finally, the CPS unemployment duration measure faces issues with individuals transitioning

between being unemployed and out of the labor force, i.e., unemployed respondents may report their total time of non-employment as their unemployment duration, despite periods when search did not occur. Nevertheless, it is useful to understand how our measure of job seeker search spells compares with the search spells of the unemployed. From Table D.2, it is clear that the website has many more short-duration job seekers and much fewer long-duration job seekers relative to the unemployed in the CPS.

Table D.3 presents evidence comparing the demographics of non-employed job seekers in the SnagAJob sample to the demographics of the unemployed in the CPS by search duration. The top panel is analogous to the top panel of Table 3 in the main text, with the only difference being that the sample is restricted to the non-employed. The bottom panel reports the same demographic statistics for the unemployed in the CPS during the same sample period (between September 2010 and September 2011). We use the last reported duration to determine spell length in the CPS since individuals may appear in the CPS up to eight times during our sample period. The table shows that, conditional on search duration, the demographics of job seekers in both samples are remarkably similar, with the caveat that the SnagAJob data over-represent women and younger job seekers overall. Nevertheless, the shares of individuals by race and education are very similar. Both samples also show similar patterns across spell duration categories. In both samples, those with longer durations are more likely to be older, male, and non-white. Both samples exhibit a slightly higher likelihood of having some college (including an associate's degree or vocational training) and show little to no differences in college attainment by search duration.

Finally, Figure D.1 compares the spell exit hazards from the SnagAJob data to the unemployment exit hazards from the CPS. The SnagAJob exits are identified as they are in our main analysis—i.e., the last application followed by five or more weeks of inactivity. The CPS exits are identified based on labor market transitions either to employment or out of the labor force, and an individual's unemployment duration is measured using their self-reported value. We condition out calendar effects from both sets of exit hazard

estimates. Figure 1 shows that the exit hazards are comparable across the two surveys in the early weeks of the search spell, but that the exit rate falls much more, and to a considerably lower rate in later weeks, in the CPS data. Part of the difference is due to differences in measurement. The SnagAJob data measure exit using observed behavior, while the CPS measures exit based on self-reported duration. It is well-known that when individuals transition back and forth between unemployment and being out of the labor force, they tend to report their unemployment duration as their total time out of work (see, for example, Elsby, Hobijn, and Sahin, 2015, and Kudlyak and Lange, 2017). This lowers the exit hazard rate for the CPS relative to the SnagAJob data for longer durations because, for the latter, we treat search after prolonged periods of inactivity as an entirely new spell. Of course, the remainder is due to the relatively short durations observed in Table D.2 for the SnagAJob data overall.

D.2. Results with the Potential Match Sample

We now turn to a replication of our main results using the potential match sample described in Section 3.C of the main text. We start with a replication of Figure 6 in the main text that parametrically estimates the relationship between the probability of applying to an expiring vacancy and the time remaining in a job seeker's search spell. To do so, we regress the fraction of applications a job seeker sent to an expiring vacancy (conditional on sending at least one application that week) on a set of dummy variables for the number of weeks remaining in her search spell and a set of controls. These controls are the same as those used for estimating application behavior in equation (1) of the main text. To control for job seeker characteristics, we use the vector observable characteristics defined in the main text and a set of fixed effects for completed spell length. To control for the local stock of vacancies, we use the specification that differentiates between the (log) number of new and existing vacancies described in the main text. We report the coefficients on the dummies for remaining spell length in Figure D.2. For reference, we also report coefficient estimates for a specification that does not include the additional

controls. Both specifications show that the fraction of applications to an expiring vacancy is relatively flat in the weeks and months prior to the end of the search spell. The one caveat is that the estimates that control for observable characteristics and spell length initially show a gradual decline in the fraction for application behavior more than three months prior to the end of the search spell. Nevertheless, just as in Figure 6 of the main text, the fraction spikes up in the last two weeks of search. Without controls, the fraction to an expiring vacancy rises 2.4 percentage points (18 percent), and with controls, the fraction to an expiring vacancy rises 1.6 percentage points (12 percent). Both increases are statistically significant. We interpret this evidence as reinforcing the belief that our selection criterion provides a good identification of hiring through the website.

Figure D.3 plots the applications per week by completed spell length for the potential match sample. It is a replication of Figure 5. As one can see, the main patterns from Figure 5 are also present in Figure D.3. In fact, they are nearly identical. Figure D.4 shows the estimates from our baseline model and the baseline model extended to include controls for demographics and completed spell length, replicating our results reported in Figure 2 of the main text. For comparison, we report the estimates using our potential match sample and using the full sample of job seekers. The figure shows that the potential match sample yields qualitatively similar results. Quantitatively, it shows a somewhat smaller effect of including completed spell length as a control, but there is still a monotonically declining relationship between applications per week and search duration. [2]

D.3. Additional Results

[2] In results not reported here, we also experiment with a version of the potential match sample where we further restrict the sample to job seekers who are the *only* ones who send an application to an expiring vacancy in their last week of search. This more restrictive sample represents only 6 percent of all job seekers but nevertheless produces qualitatively similar results to those in Figures D.3 and D.4.

Figure D.5 shows the distribution of search duration for our sample of website job seekers in July 2011. Mean vacancy duration is 6.5 weeks (Table 2 in the main text), but over 21 percent of job seekers are on the website for only one week, with 43 percent on the website for one month or less. Nearly two-thirds of all vacancies are filled within three months, with only 15 percent of vacancies lasting six months or more.

Figures D.6 through D.8 examine the robustness of our main results. In Figure D.6, we examine whether the second and subsequent search spells on the website, identified using the five-week cutoff, exhibit qualitatively similar application behavior as the one documented for the first search spell after registration on the website. In doing so, we identify job seekers with two or more spells and stack the job seeker-week observations of these spells with the first-spell observations of our main sample. We then replicate our regression analyses based on equation (1) from the main text on the stacked panel, including dummy variables for the spell number and interactions between the spell number and the current duration of the spell. We identify a second spell for about 17.3 percent of job seekers, a third spell for 4.0 percent of job seekers, and a fourth or higher spell for about 0.9 percent of job seekers. In the regression analysis, we use a single dummy variable for the fourth and subsequent spells because of the relatively small sample size for this group of job seekers and the fact that later spells are increasingly right-censored given the one-year length of our sample period.

Figure D.6 shows the results using our baseline specification and the full specification that includes additional controls for jobseeker fixed effects and the number of incumbent and newly-posted vacancies active in the metropolitan area.[3] The figure shows that the later search spells all exhibit a declining number of applications per week over their duration. In fact, their patterns are nearly identical to those

[3] We report the estimates for the first three spells given the noisy nature of the estimates for the fourth and subsequent spells.

one observes for the first spell. The evidence confirms the robustness of our results, and rejects a hypothesis that the observed decline in applications per week in our main results is the consequence of increasingly efficient search by job seekers that learn how to better use the website over time.

Figure D.7 replicates our test of the effects of stock-flow matching on the effort-duration relationship using a narrower definition of the stock and flow of vacancies. We define these vacancies based on the modal occupation of the job seeker (i.e., the occupation that they applied to most often over their full search spell) and the skill relevance of the vacancy. The skill relevance is based on whether the job seeker's education roughly matches the average education of all job seekers who applied to the vacancy in question. This results in 25,769 unique CBSA-Occupation-Skill categories that define the relevant stock and flow of vacancies. Figure D.7 reports the results for all weeks of the search spell, rather than just the weeks where at least one application was sent, but we obtain nearly identical results for both samples. Furthermore, we only report the results for the stricter specification of stock-flow matching that uses the log number of new and existing vacancies within each category during each week. The broader measure of average vacancy duration is based only on vacancies that the job seeker applied to so it does not suffer from a mismeasurement of the relevant stock of vacancies. In short, the results for the narrower definition of the relevant stock and flow of vacancies are nearly identical to those observed for the CBSA-wide measure used in Figure 3.

Finally, Figure D.8 replicates the exercise from Figure 5 of the main text using different subsets of the data. One may be concerned that our results are an artifact of how we define search spells. Therefore, we replicate the exercise from Figure 5 using only the job seeker-week observations where at least one application was sent. The results in the left panel of Figure D.8 show that our main result—that longer-duration job seekers exert more effort throughout the duration of search—holds. We also restrict our

sample to those that we identify as non-employed. In this case, the results are nearly identical to those observed in Figure 5.

References

Elsby, Michael W.L., Bart Hobijn, and Ayşegül Şahin, 2015. "On the Importance of the Participation Margin for Labor Market Fluctuations." *Journal of Monetary Economics* 72(1): 64-82.

Faberman, R. Jason, and Marianna Kudlyak, 2016. "What Does Online Job Search Tell Us about the Labor Market?" *Federal Reserve Bank of Chicago Economic Perspectives* 40: 2016-1.

Kudlyak, Marianna, and Fabian Lange, 2017. "Measuring Heterogeneity in Job-Finding Rates among the Nonemployed Using Labor Force Status Histories." Federal Reserve Bank of San Francisco Working Paper 17-20.

Table D.1. Demographic Characteristics, Website Sample and the Current Population Survey

	Share of Website Job seekers				
	All	Spell Length > 1 week	Spell Length ≤ 1week	Share of Unemployed (CPS)	Share of Labor Force (CPS)
Gender					
Male	43.1	43.2	43.1	56.3	53.3
Female	56.9	56.8	56.9	43.7	46.7
Age					
16-24 Years Old	52.8	48.5	54.2	26.3	13.6
25-39 Years Old	26.9	26.4	27.1	31.6	32.2
40-54 Years Old	15.2	18.1	14.2	27.4	34.2
55+ Years Old	5.1	6.9	4.5	14.7	19.9
Education					
High School or Less	62.5	58.4	63.9	51.0	37.1
Certification or Some College	10.1	11.0	9.8	19.5	17.1
Associates Degree	12.0	14.3	11.3	20.0	10.6
Bachelor's Degree or More	15.3	16.3	15.0	9.4	35.1
Race					
White	50.3	50.0	50.4	54.4	67.2
Black	25.4	26.1	25.2	19.4	11.0
Hispanic	14.6	14.2	14.7	19.2	14.8
Other	9.7	9.7	9.7	6.9	6.9
*Modal Occupation Applied To**					
Health & Education	2.7	1.8	3.0	NA	NA
Other Professional	3.2	2.7	3.7	NA	NA
Food & Hospitality	19.9	19.0	20.2	NA	NA
Retail	54.9	63.8	51.8	NA	NA
Customer Service	2.9	2.0	3.1	NA	NA

Notes: Table reports the share of individuals in each demographic category from our sample of job seekers on the SnagAJob website as well as the unemployed and those in the labor force, as reported in the Current Population Survey (CPS). CPS statistics are monthly averages over September 2010 to September 2011.

Table D.2. Differences in Duration, Website Sample and Current Population Survey, July 2011

	All Job Seekers	All Job Seekers with > 1 Application	Non-Employed Job Seekers with > 1 Application	CPS Unemployed
Unemployment Duration				
Less than 5 weeks	72.3	54.2	52.5	20.5
5-14 weeks	22.7	37.6	38.0	24.2
15-26 weeks	3.7	6.1	7.0	12.2
27 or more weeks	1.2	2.1	2.5	43.1
Mean duration, weeks	4.0	6.0	6.3	39.0
Median duration, weeks	1.0	4.0	4.0	19.7
N	185,891	112,293	67,824	*

Notes: Table reports the share of job seekers (or the unemployed, for the CPS) with an active search spell within the listed rage, with summary statistics on the duration of (incomplete) search spells included. Website data are from a cross-section of job seekers identified as actively searching during the CPS reference week of July 2011, and only include job seekers during their first identified search spell.

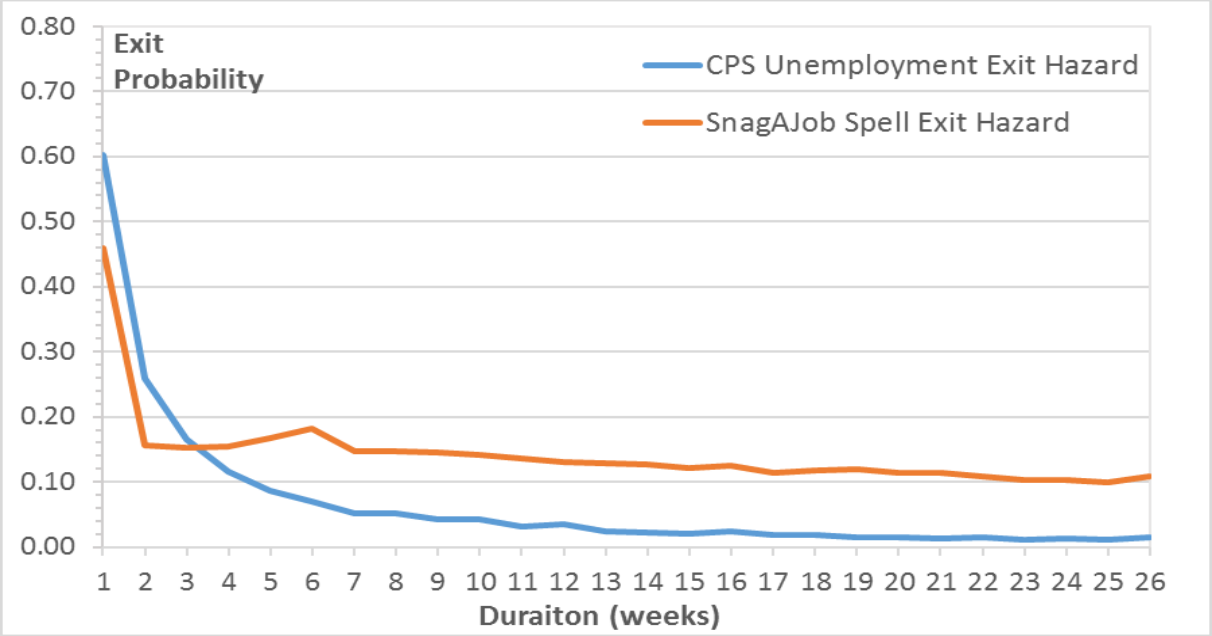
* CPS statistics are from published data, which typically come from a sample of about 100,000 individuals aged 16 and over.

Table D.3. Demographics by Spell Length, Website Sample and Current Population Survey

	Search Duration			
	0-6 Weeks	6 Weeks - 3 Months	3 - 6 Months	> 6 Months
SnagAJob, Non-Employed				
Mean Age	29.8	32.9	36.1	40.9
Pct. Female	56.7	56.9	56.3	53.5
Pct. Nonwhite	49.9	52.5	52.8	50.0
Pct. Associates/Vocational	23.8	28.0	29.5	28.6
Pct. College or More	15.7	15.9	15.3	16.1
Current Population Survey, Unemployed				
Mean Age	33.5	33.2	35.2	39.2
Pct. Female	45.3	46.8	44.0	44.6
Pct. Nonwhite	41.9	45.9	46.8	48.0
Pct. Associates/Vocational	27.8	26.7	27.2	28.8
Pct. College or More	15.1	15.3	15.3	16.0

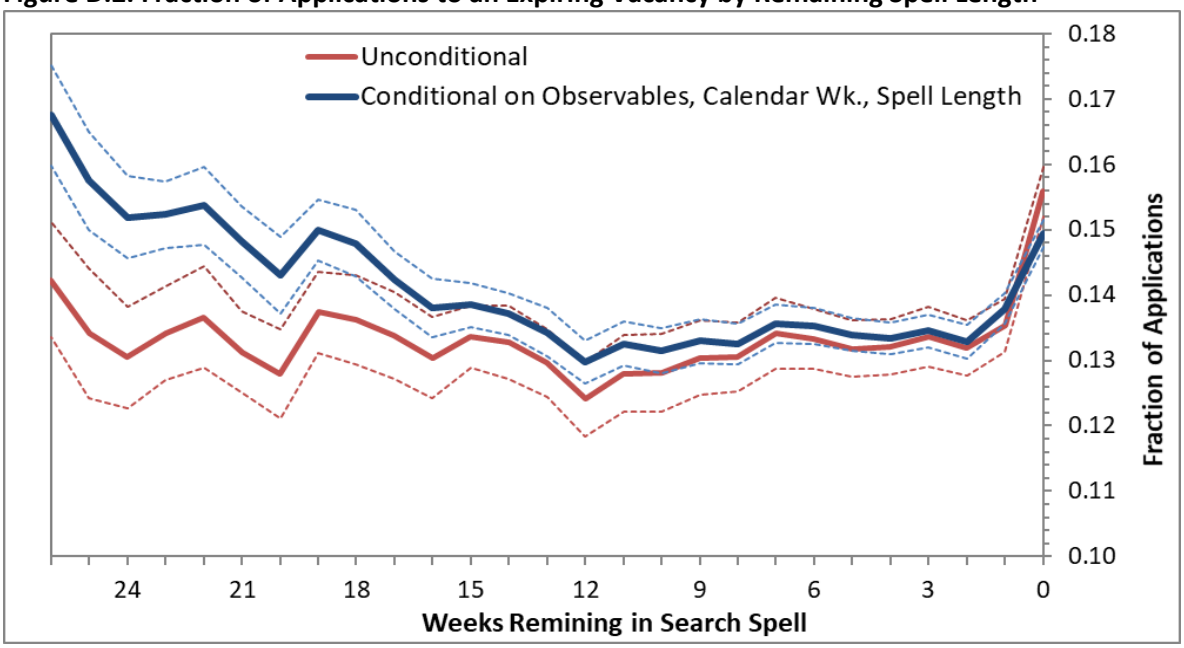
Notes: Table reports the demographic characteristics of non-employed (website) or unemployed (CPS) individuals by total observed search duration. For the website data, the total duration is their observed completed spell length. For the CPS, the duration is the longest duration that they report in the survey. Both samples include job seekers aged 16 to 75 who either registered on the website between September 2010 and September 2011 or were in the CPS during the same period.

Figure D.1. Unemployment and Search Spell Weekly Exit Hazards



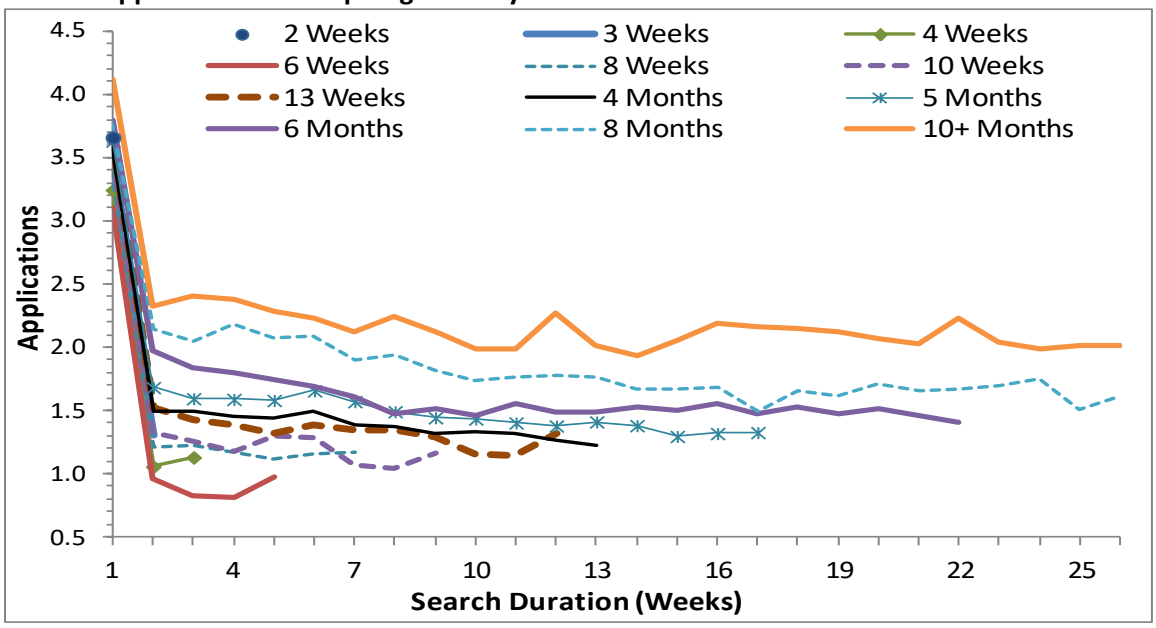
Note: The figure reports the exit hazard rate by week of duration for the SnagAJob and CPS data. The SnagAJob data measures the exit from the first search spell, identified by the last observed application prior to five or more weeks of inactivity. The CPS data measures the exit from unemployment to another labor market state as a function of self-reported unemployment duration. SnagAJob estimates use all job seeker-week observations from the first identified search spell of our full sample of job seekers. CPS estimates use all unemployed pooled over the September 2010 to September 2011 period, grouped by identified unemployment duration.

Figure D.2. Fraction of Applications to an Expiring Vacancy by Remaining Spell Length



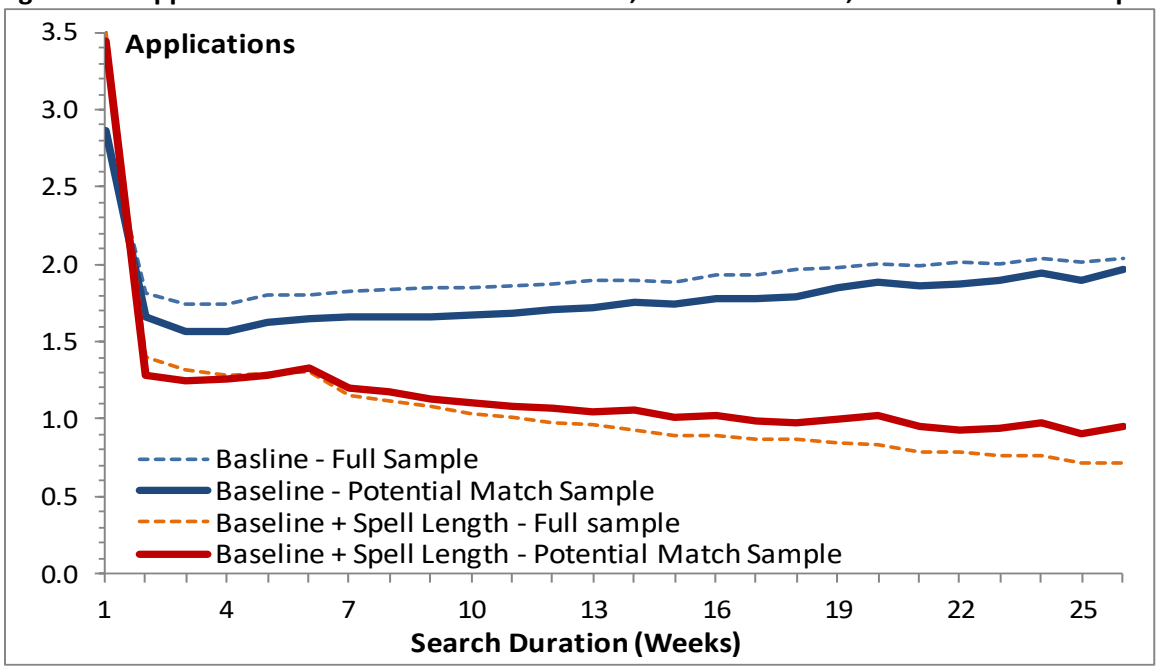
Note: The figure reports the fraction of job seeker applications sent to an expiring vacancy by weeks remaining in their (completed) search spell. Conditional estimates control for observable job seeker characteristics, calendar week, and completed spell length. Dashed lines represent 95 percent confidence intervals.

Figure D.3. Applications by Search Duration and Completed Spell Length, Conditional on Ending Search with an Application to an Expiring Vacancy



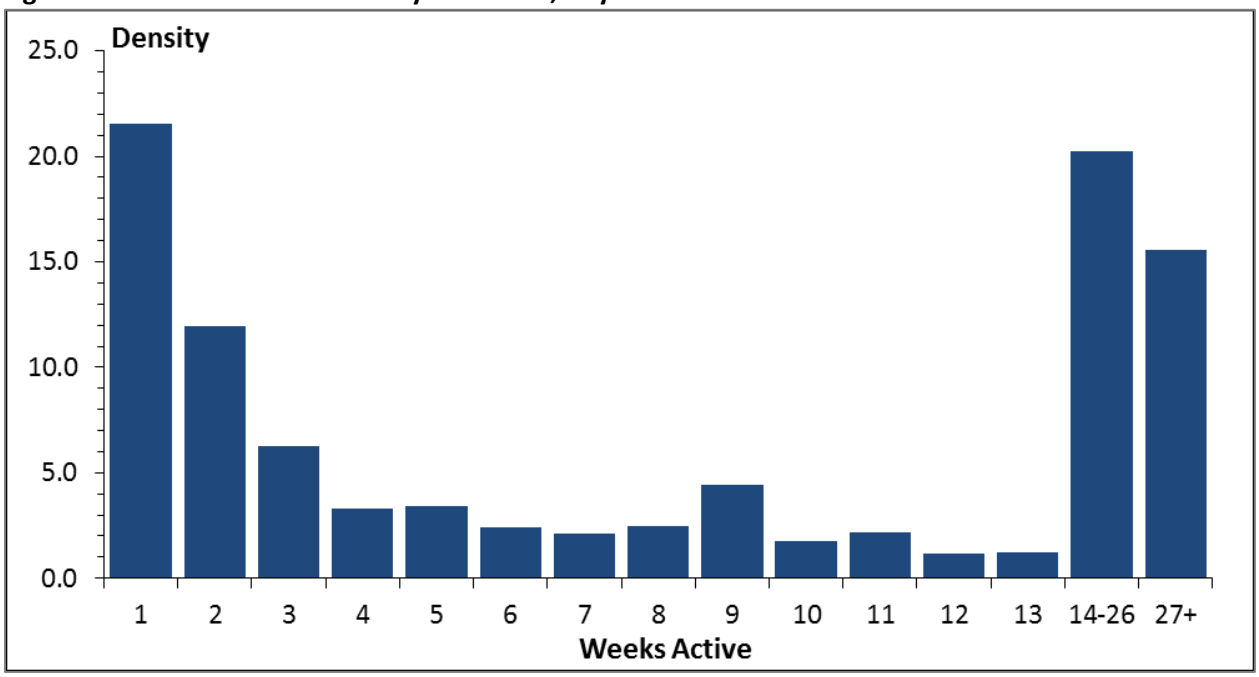
Notes: Figure shows the estimated (unconditional) relationship between applications per week and duration of search separately for job seekers based on the total length of their search spell. Mean application estimates are calculated using jobseeker-week observations of the first identified search spell in our sample who sent at least one application to a vacancy that expired during their last week of search. Only selected spell lengths are reported.

Figure D.4. Applications over the Duration of Search, Various Controls, Potential Match Sample



Notes: Figure shows the estimated relationship between applications per week and the duration of search for our baseline model and a model that additionally controls for total spell length using the full sample of job seekers in our SnagAJob website sample and a subsample that are identified as potentially matching with an expiring vacancy.

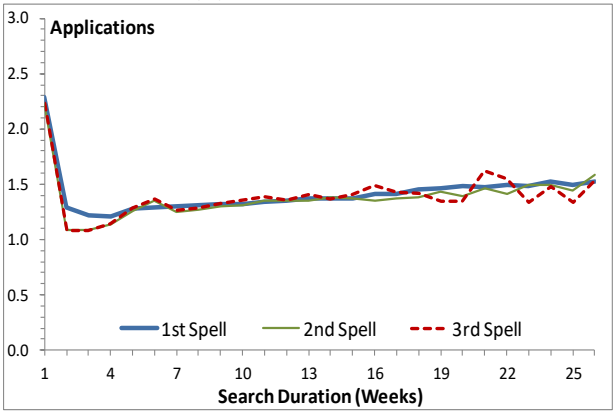
Figure D.5. Distribution of Vacancy Durations, July 2011



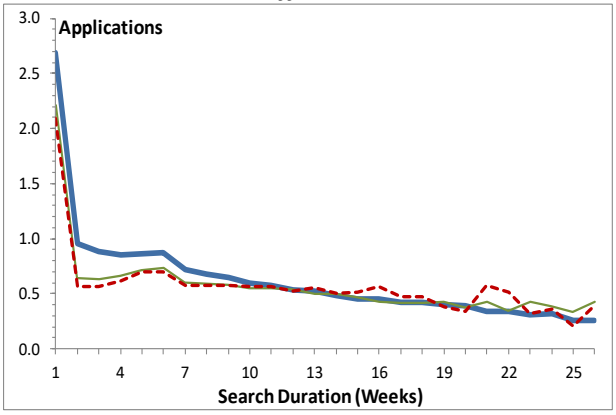
Note: The figure reports the fraction of vacancies active for each category on the SnagAJob website in July 2011. The sample excludes vacancies that begin before start of the sample period (September 2010).

Figure D.6. Applications over the Duration of Search, Estimated with Multiple Spells per Job seeker

(a) Baseline Model

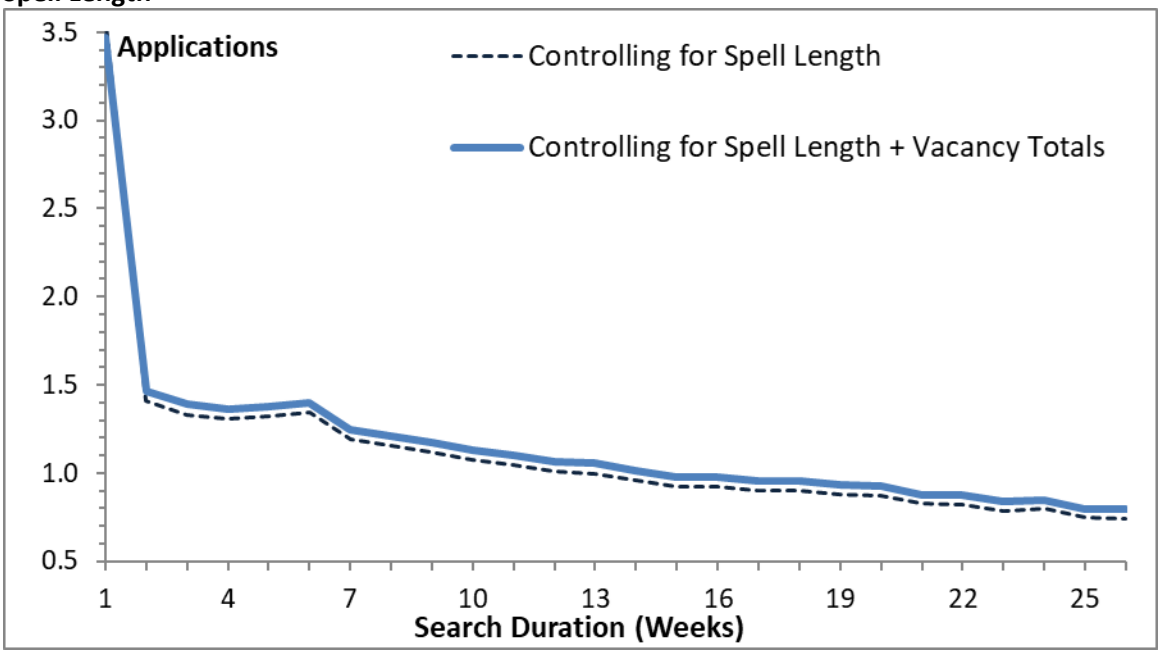


(b) Controlling for Spell Length, Job seeker Effects, and Vacancies



Notes: Figure shows the estimated relationship between applications per week and duration of search for our baseline model (left panel) and a model that additionally controls for active vacancies, fixed job seeker characteristics, and completed spell length (right panel). The model is estimated across all search spells for each job seeker in our SnagAJob website sample.

Figure D.7. Applications over the Duration of Search, Controlling for Active Vacancies and Completed Spell Length

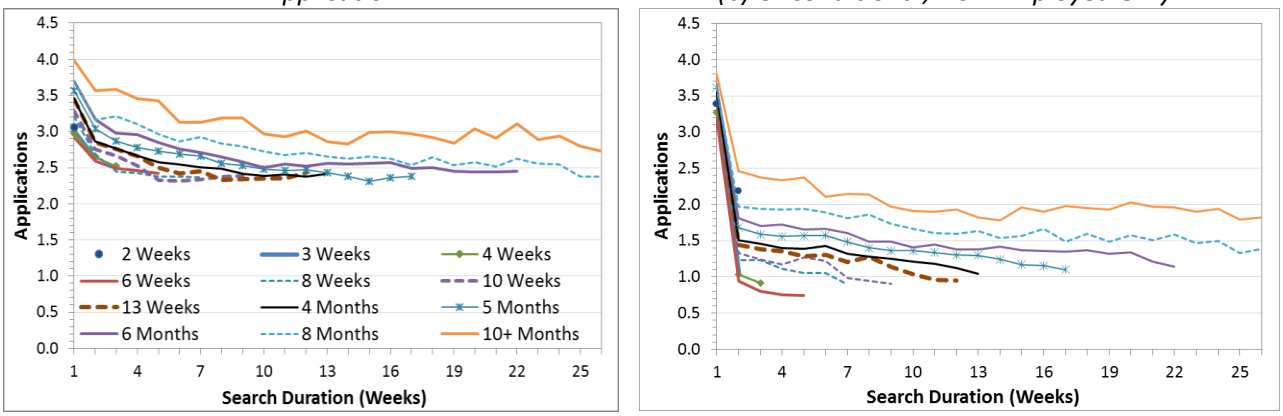


Notes: Figure shows the estimated relationship between applications per week and the duration of search for our baseline model with demographic controls and spell length fixed effects alone, and in a model that additionally controls for the number of newly-posted and previously active vacancies within a narrow definition of CBSA-Occupation-Skill categories.

Figure D.8. Applications over the Duration of Search by Completed Spell Length, Robustness

(a) Conditional on Sending at Least One Application

(b) Unconditional, Non-Employed Only



Notes: Figure shows the estimated (unconditional) relationship between applications per week and duration of search separately for job seekers based on the total length of their search spell. In the left panel, mean applications are only calculated for individuals in our SnagAJob website sample who sent at least one application in a given week. In the right panel, mean applications are calculated for all individuals in the sample, but only after conditioning out demographic and local labor market characteristics. See text for details. Only selected spell lengths are reported.