Web Appendix (not intended for publication)

Appendix A. Robustness of Spell Length Identification

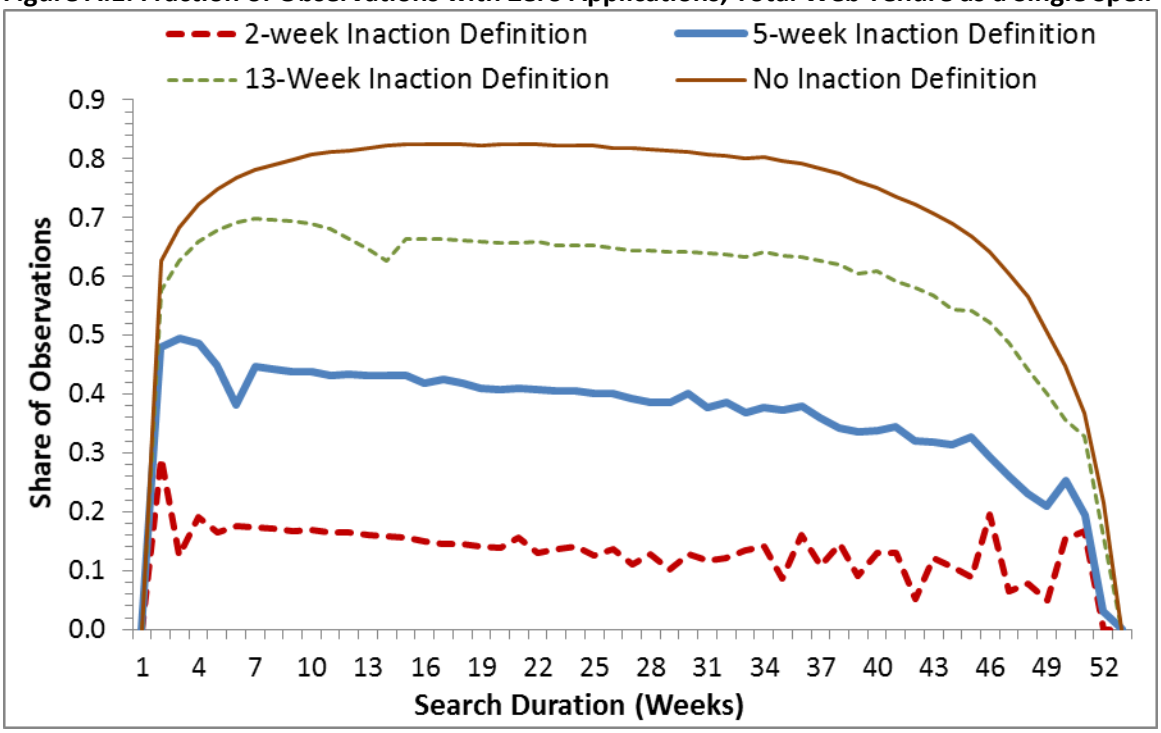
Identifying when job seekers are actually searching as opposed to being idle on the website is an important part of our analysis. Many job seekers send applications over a short period of time, take a break from searching, and return later, sometimes many weeks later, to send applications again. We use a strategy of counting more than five weeks of inactivity as the end of one search spell and the start of a new spell. 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.

In this section, we show how our results change under different assumptions about how much inactivity is required to identify the end of a search spell. In addition to the five-week cutoff, we also replicate results using a two-week and 13-week cutoff. Figure A.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, the share never rises above 50 percent and declines steadily thereafter.

Figure A.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 A.1. Finally, Figure A.3 replicates our regression analysis for four specifications taken from equation (1) in the text using the three different cutoffs (two weeks, five weeks, and 13 weeks, with five weeks being the cutoffs used in the main analysis). Prior to all controls being added, there are large quantitative differences across the cutoff rules. When we

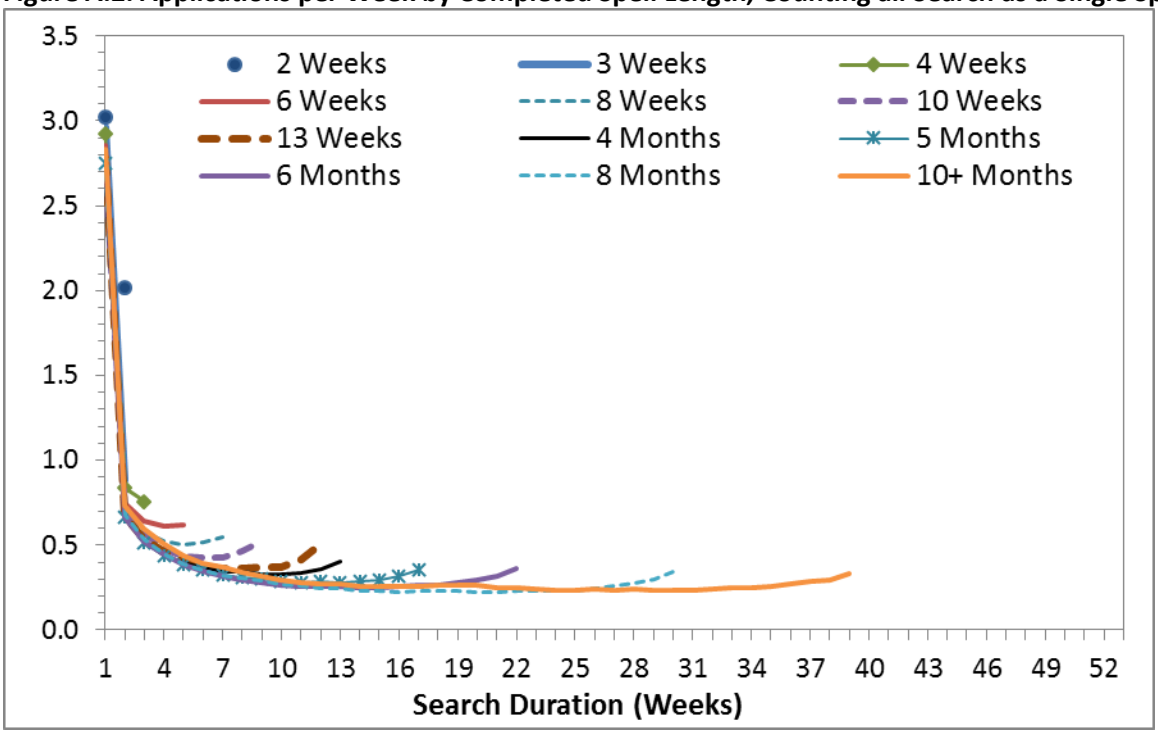
add either spell length or job seeker fixed effects, however, the results are similar regardless of the cutoff used. In all cases, applications per week decline with search duration.

Figure A.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.

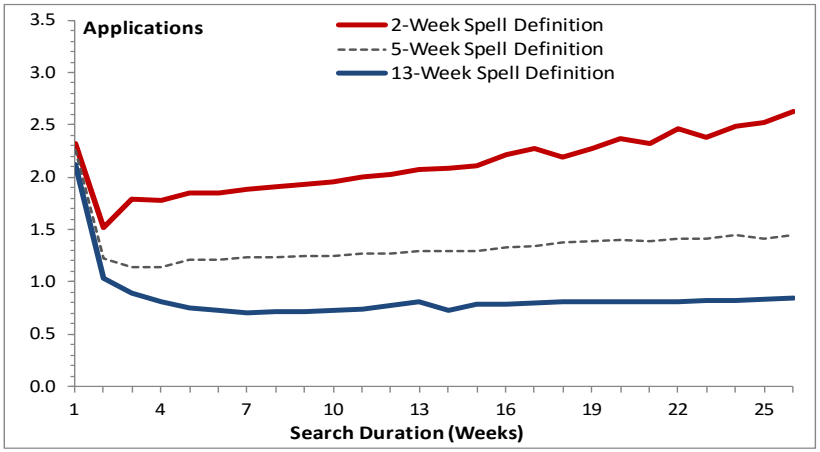
Figure A.2. Applications per Week by Completed Spell Length, Counting all Search as a Single Spell



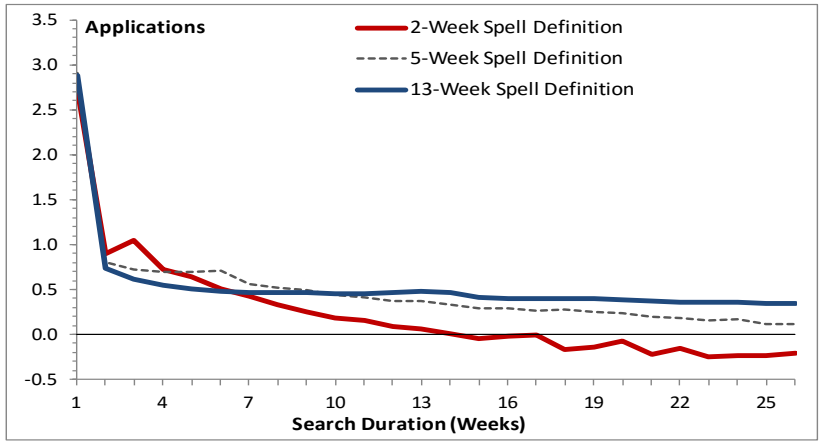
Note: Figure plots applications per week for job seekers of differing completed spell lengths, based on the assumption that all search on the website is contained within a single spell.

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

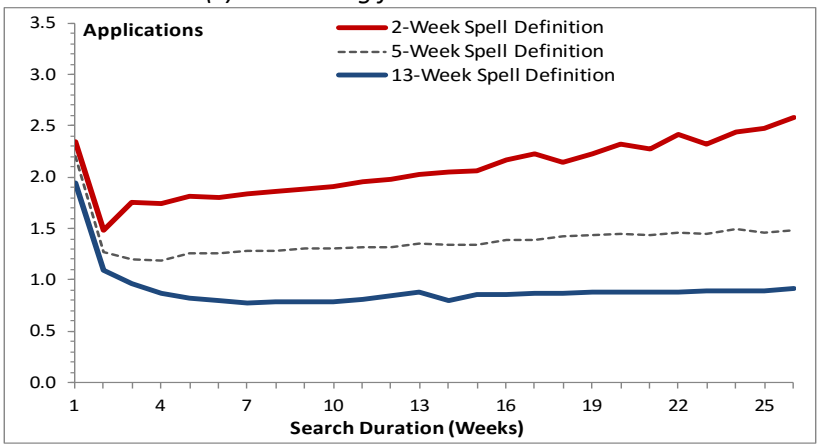
(a) Baseline Model



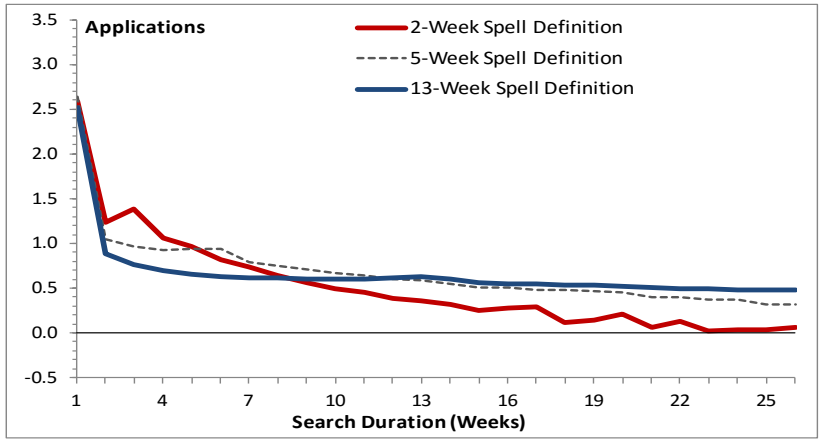
(b) Controlling for Demographics & Spell Length



(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.

Appendix B. Model of Job Search with Heterogeneity in Website Preference

We examine the robustness of our main results using a sample of "potential matches" in Section 6 of the paper. One may worry, however, that the potential match sample may still suffer from a spurious correlation between applications and search duration. Specifically, if our potential match sample contains a large number of individuals who are marginally interested in finding a job on the website (for example, because they have found a job on the website through a pure luck), one might worry that we obtain similar results in our full sample and potential match sample purely through this "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.

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 n applications per week. To keep the exercise simple, we assume that the number of applications per week remains constant over the duration of search.[2] Job seekers who search exclusively on the website send all n applications on the website. Those who are marginally attached send a fraction α of their applications through the website and the remaining $(1-\alpha)n$ 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. We also perform the simulation under the assumption of a constant

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

[2] Constant search over time is assumed for simplicity given that this exercise focuses on behavior across individuals with differing completed spell lengths, and not differences within search spells.

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, $\{n, \alpha, f\}$ and one function, $\rho(t)$, that we calibrate to the data. We assume that n equals the mean number of applications sent in their first week of search by those 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 n represents the total number of applications sent using all methods in the model. We calibrate α using n and the model's expression for the expected total number of applications sent in the first week, $\theta\alpha n + (1-\theta)n$. 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 $1-(1-f)^n$. For the website quit probability, we assume that $\rho(t) = \rho_0/t^{\rho_0+1}$, which allows it to decline with duration analogously to a p.d.f. of a Pareto distribution. We calibrate ρ_0 by equating the probability of exit after the first week to $1-(1-f)^n + \theta\rho(1)(1-f)^n$. 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.

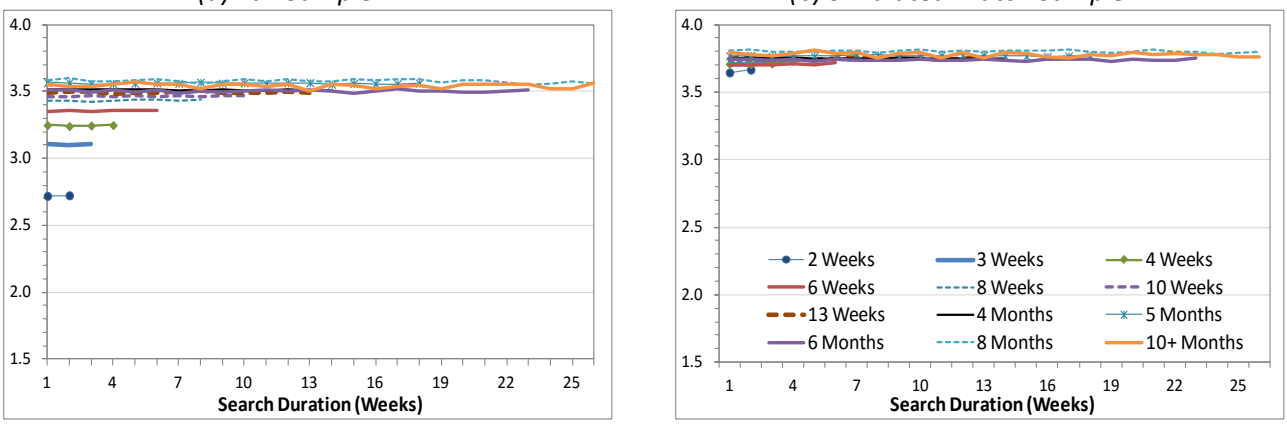
The results of the exercise are in Figure B.1. The left panel shows the simulated applications per week for the full sample of simulated job seekers (analogous to Figure 5 in the main text) and the right panel shows the simulated applications for those who found a job on the website (analogous to Figure 9 in the main text). 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 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 B.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 $1 - (1 - f)^n + \theta\rho_0(1 - f)^n$. 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 B.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 (full panel). If our results were driven only by individuals dropping out of searching on 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 suggesting that our main results are not driven by exits that are based primarily on tastes for search on the SnagAJob.com website.

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

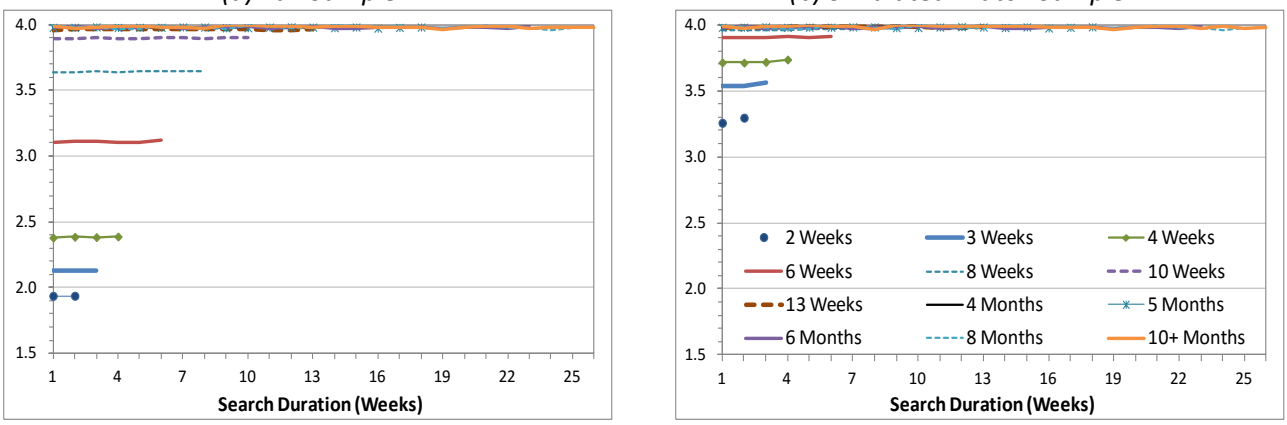
(a) Full Sample *(b) Simulated Match Sample*



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 B.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 C. Additional Results

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 C.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. A 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.

Table C.2 compares the resulting distribution of search durations in our sample with the distribution of unemployment durations within the Current Population Survey (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 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 C.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.

Additional results

Figure C.1 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 C.2 and C.3 examine the robustness of our main results. In Figure C.2, 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 C.2 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 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 use the website over time.

Finally, Figure C.3 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. As we discuss in the main text, there are reasons to believe that this cannot be the case. Nevertheless, 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 C.3 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.

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

Table C.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 website job seekers 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.

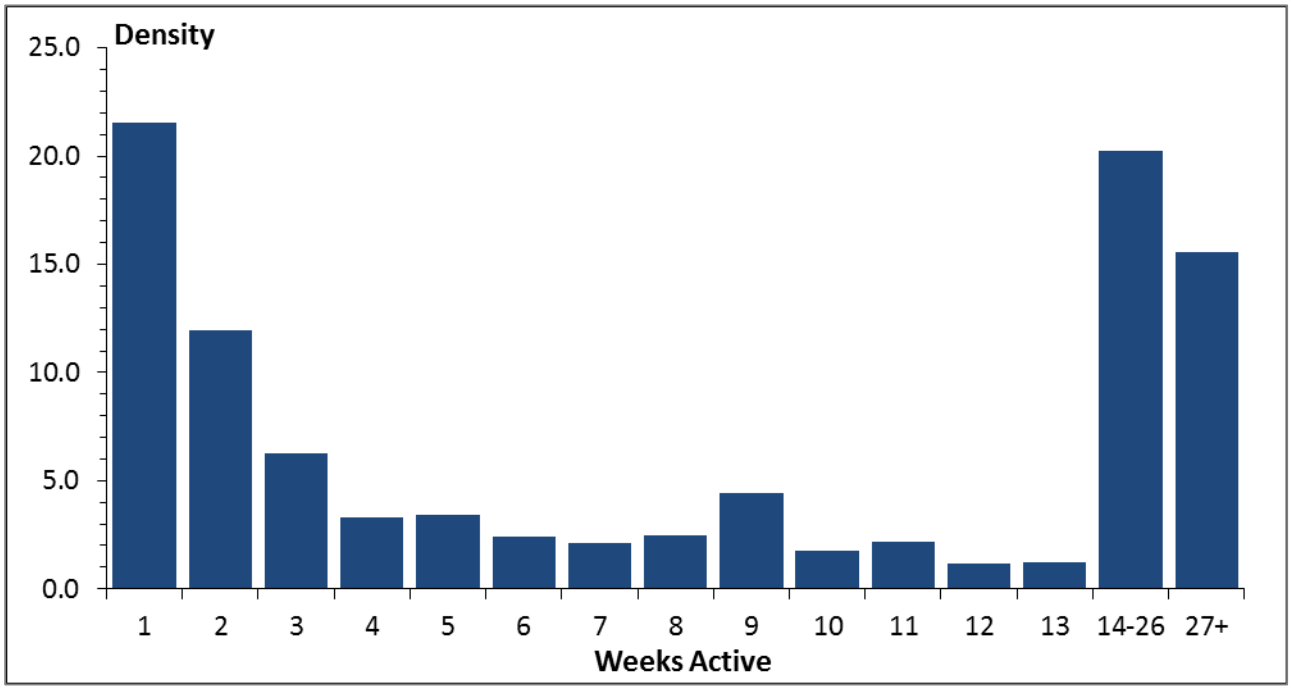
Table C.2. Differences in Duration, Website Sample and Current Population Survey, July 2011

	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.

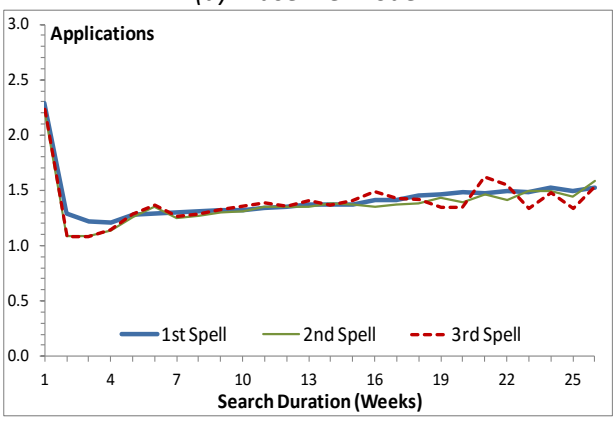
Figure C.1. Distribution of Vacancy Durations, July 2011



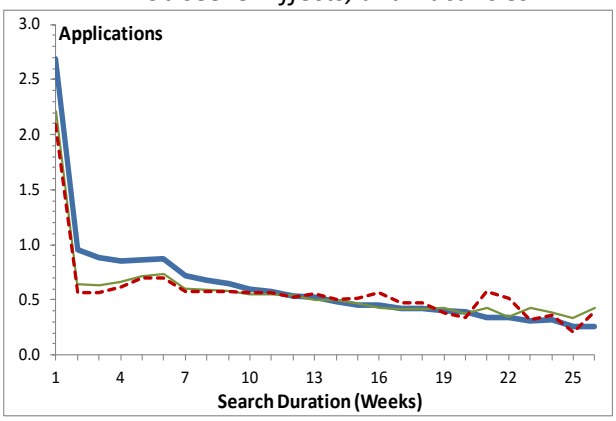
Note: The figure reports the fraction of vacancies active for each category. The sample excludes vacancies that begin before start of the sample period.

Figure C.2. Applications over the Duration of Search, Estimated with Multiple Spells per Job seeker

(a) Baseline Model



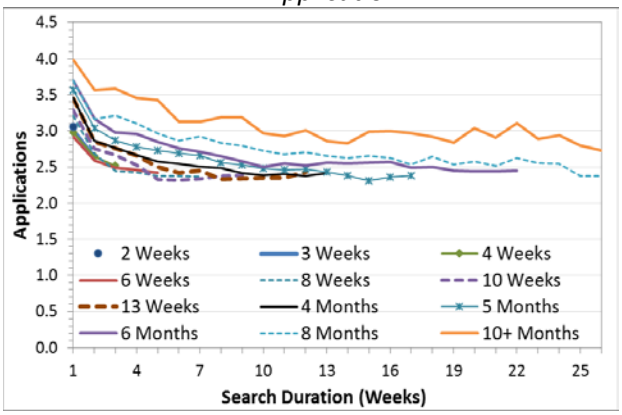
(b) Controlling for Spell Length, Job seeker Effects, and Vacancies



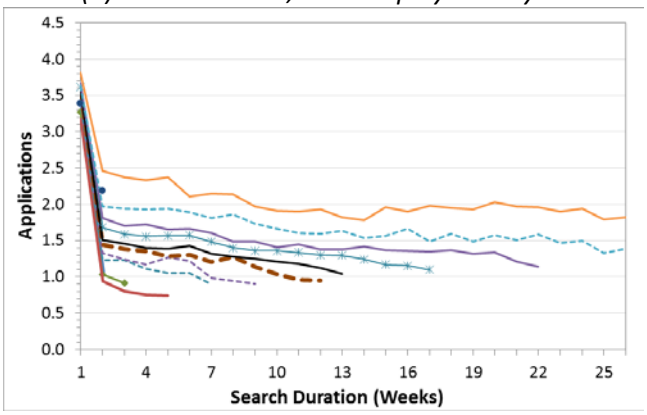
Notes: Figure shows the estimated relationship between applications per week and duration of search for our baseline model (left panel) and a model that additionally controls for active vacancies, fixed job seeker characteristics, and completed spell length (right panel). The model is estimated across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

Figure C.3. 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 who sent at least one application in a given week. In the right panel, mean applications are calculated for all individuals, but only after conditioning out demographic and local labor market characteristics. See text for details. Only selected spell lengths are reported.