Firm Dynamics and Assortative Matching

by

Leland D. Crane*
U.S. Census Bureau

CES 14-25 May, 2014

The research program of the Center for Economic Studies (CES) produces a wide range of economic analyses to improve the statistical programs of the U.S. Census Bureau. Many of these analyses take the form of CES research papers. The papers have not undergone the review accorded Census Bureau publications and no endorsement should be inferred. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. Republication in whole or part must be cleared with the authors.

To obtain information about the series, see www.census.gov/ces or contact Fariha Kamal, Editor, Discussion Papers, U.S. Census Bureau, Center for Economic Studies 2K132B, 4600 Silver Hill Road, Washington, DC 20233, CES.Papers.List@census.gov.

Abstract

I study the relationship between firm growth and the characteristics of newly hired workers. Using Census microdata I obtain a novel empirical result: when a given firm grows faster it hires workers with higher past wages. These results suggest that productive, fast-growing firms tend to hire more productive workers, a form of positive assortative matching. This contrasts with prior research that has found negligible or negative sorting between workers and firms. I present evidence that this difference arises because previous studies have focused on cross-sectional comparisons across firms and industries, while my results condition on firm characteristics (e.g. size, industry, or firm fixed effects). Motivated by the empirical findings I develop a search model with heterogeneous workers and firms. The model is the first to study worker-firm sorting in an environment with worker heterogeneity, firm productivity shocks, multi-worker firms, and search frictions. Despite this richness the model is tractable, allowing me to characterize assortative matching, compositional dynamics and other properties analytically. I show that the model reproduces the positive firm growth-quality of hires correlation when worker and firm types are strong complements in production (i.e. the production function is strictly log-supermodular).

Keyword*:* Assortative Matching, Firm Growth, Wages, Unemployment, Vacancies, Search Theory, Microdata

JEL Classification: E24, J31, J63, J64

---

[*] University of Maryland, College Park and U.S. Census Bureau.

I would like to thank John Haltiwanger, Boragan Aruoba, and Sergio Urzua for extensive advice and support. I thank Shigeru Fujita, Brian Kovak, Andrew Levin, Kristin McCue, Toshihiko Mukoyama, Chris Nekarda, Illenin Kondo, Jeongmin Lee, John Shea, Luminita Stevens and seminar participants at the University Maryland, the U.S. Census Bureau, the 2013 Georgetown CER Conference, the 2013 SGE Conference, and the 2013 CAED Conference for many helpful comments. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. All errors are my own.

1 Introduction

Do fast growing firms hire different workers from slow growing or shrinking firms? How do firm recruitment strategies vary with employment growth? Do more productive workers match with more productive firms? I address these questions empirically and theoretically. Using linked Census microdata I document a new fact, that faster growing firms hire better workers. This suggests positive sorting of workers across firms, a result that has been elusive in the empirical literature. I extend the search and sorting model of Shimer and Smith (2000) to include multi-worker firms, productivity shocks, and firm growth. Once solved the extended model turns out to be highly tractable and yields analytical comparative statics. The model tightly links firm employment growth to productivity and the quality of hires, and I show that the empirical correlation can be reproduced if the production function exhibits strong complementarities between firm and worker types.

My focus is on how workers of differing productivity sort across firms in a non-random way. Textbook search models (e.g. Pissarides (2000) Chapter 1) assume homogenous workers and firms in order to focus on the number of workers employed, rather than the exact sorting of workers across jobs. Many authors have extended the bare bones model to capture heterogeneity in worker productivity or firm productivity, but less often both. This has been partially due to the difficulties of measuring sorting empirically and modeling it in a tractable way. However, significant dispersion in productivity across workers and firms has been well documented, and there is abundant anecdotal evidence of workers and firms sorting according to productivity. This evidence suggests that worker-firm sorting has potentially significant consequences both empirically and in terms of welfare.

Recently, the increased availability of matched employer-employee data and the seminal contribution of Abowd, Kramarz and Margolis (1999) (AKM hereafter) have brought renewed interest to measuring and modeling sorting in the labor market. AKM jointly estimate worker and firm fixed effects from linked wage data. The estimated worker and firm effects may be interpreted as estimates of worker and firm productivity. Attention has centered on the question of *positive assortative matching*. Under positive assortative matching, or assortative matching for short, the best workers match with the best firms and the worst workers match with the worst firms. Negative assortative matching obtains when the worst workers match with the best firms, and the best workers with the worst firms. Perhaps surprisingly, AKM and the large subsequent empirical literature found mixed evidence of positive assortative matching. Authors have proposed a number of explanations, including match effects, limited mobility bias, and non-monotonicity of wages in firm productivity. All three may be serious issues, and it appears infeasible to simultaneously correct the AKM estimator along all these dimensions.

In this paper I use a different methodology and reach different conclusions. I find that, conditional on observable firm characteristics, faster growing firms tend to hire more productive workers. This suggests a kind of positive assortative matching, where fast growing and more productive firms match with more productive workers. My results contrast with much of the previous literature. The different results are due to a difference in methodologies. In my baseline specifications I

condition on a variety of firm characteristics, including industry, size, and firm fixed effects. The positive correlation changes dramatically when firm characteristics are not controlled for. Without firm characteristics the relation between firm growth and worker quality is highly nonmonotonic with a maximum around zero growth. This result echoes the findings in the previous literature, which have generally not conditioned on firm characteristics, and found mixed evidence of positive sorting.

I attribute these differences to segmentation in the labor market. When I condition on firm characteristics I implicitly compare firms to very similar firms, which are likely to inhabit similar labor markets.[1] Thus the baseline specifications, controlling for industry and other characteristics, pick up positive sorting within each segment of the labor market. The best doctors work for the best hospitals, the best pilots work for the best airlines, and so on. When I do not control for firm characteristics I am comparing nurses to pilots and hospitals to airlines. Given the significant differences in technology and human capital between sectors it is possible that sorting patterns between sectors differ from those within sectors. This argument suggests that while assortative matching may be the norm within any given labor market, the distribution of workers and firms across markets need not be assortative. Informally, intuition about assortative matching is based more on examples of specific industries or occupations, rather than economy-wide matching patterns. Thus it need not be surprising that I find evidence of positive sorting within narrow markets, but neither I nor the related literature find clear evidence of positive sorting across the entire economy.[2]

A second important difference between my work and the previous literature is that I focus on firm employment growth. There is little, if any, previous research on how the quality of hires changes as a firm's employment growth varies. Firm growth is important to understand in its own right, but also serves as a plausible proxy for firm productivity. The measurement of firm productivity is critical for studying assortative matching. Recent papers have argued that the AKM fixed effects do not necessarily capture firm productivity. In my data I cannot measure productivity directly, so I propose firm growth as a proxy for firm productivity. This connection is supported by the extensive literature showing that more productive firms tend to grow relative to their less productive peers (See Bartelsman and Doms (2000), Foster et al. (2001), Foster et al. (2013) and the related literature). Firm growth also has the advantages of being trivial to compute and exhibiting

[1] Controlling for industry, size and location clearly narrow the scope of comparison towards a single labor market. Including firm fixed effects makes the scope of comparison a single firm (using only time variation). This may seem to be a different kind of approach from controlling for industry, size, etc. However, we can interpret firm fixed effects as controlling for unobserved firm characteristics. It is highly doubtful that firm characteristics observable in the data are sufficient to precisely identify the labor market the firm faces. Under the assumption that firms do not jump between markets often, firm fixed effects are an effective means to proxy the firm's market in the absence of more detailed firm characteristics.

[2] A closely related issue is whether we can meaningfully rank workers and firms in disparate markets according to productivity. It is straightforward to construct at least a partial order of doctors or of pilots, but more difficult to compare productivity across occupations. When skill is multi-dimensional the assortative matching model is not appropriate, since it presumes a scalar worker (or firm) productivity variable. A generalized Roy model may be more appropriate when comparing different sectors.

significant time variation. Previous studies have abstracted from firm productivity dynamics, and thus do not use exploit the time dimension when studying sorting. The presence of time variation also allows me to use firm fixed effects as controls in some specifications.[3]

In addition, I use lagged worker wages to measure worker quality rather than AKM worker fixed effects. This is again motivated by questions regarding the consistency of the AKM fixed effects. Using lagged worker wages raises a number of potentially serious measurement issues, which I address in Section 2. Chief among these concerns is that on the job search generates a correlation between lagged and current wages independent of any worker productivity effects. I discuss this issue in detail in Section 2 and argue that this problem can be overcome by focusing on workers hired from nonemployment, as I do.

Motivated by the empirical findings, I develop a theory of sorting with search frictions and productivity shocks. My focus is on characterizing equilibrium matching patterns within a single market, and finding conditions under which (1) better workers match with better firms, (2) better workers match with faster growing firms. The model incorporates costly search, wage bargaining, ex-ante heterogeneity in worker and firm productivity, productivity shocks, and multi-worker firms. To my knowledge this is the first paper to characterize sorting in such an environment. Despite its richness the model is tractable, and I analytically derive restrictions on the production function which ensure that in equilibrium faster growing firms hire better workers, reproducing the empirical findings. The argument requires two steps. First, I show that strong production complementarities between worker and firm type (i.e. strict log-supermodularity of the production function) ensure positive assortative matching: better workers match with better firms. Second, I show that under positive assortative matching better firms are also faster-growing. This point, while intuitive, is not trivial to prove when workers are heterogeneous. Taken together these results imply a positive correlation between firm growth and the quality of hires as well as assortative matching by types.

An important feature of my model is that firms can employ multiple workers, whereas Shimer and Smith (2000) and others focus on single worker firms, or firms that cannot freely post additional vacancies. This difference turns out to matter a great deal. The reason is that when firms can post additional vacancies they have no opportunity cost of hiring a given worker. A central feature of the Shimer and Smith (2000) model is that both parties face a tradeoff between matching immediately or waiting for a better partner. If firms can employ multiple workers, they can hire a given worker today *and continue searching* (posting vacancies) in subsequent periods. This, along with linear production, eliminates the opportunity cost of hiring of for firms that are in contact with workers, greatly simplifying the analysis. This tractability allows me to explore other aspects of the model. I show

[3] A potential concern with using employment growth to proxy productivity is that productivity growth may be caused by or concurrent with *downsizing* of the workforce. In such a case employment growth and productivity would be negatively correlated. While there is no doubt that there are examples of this phenomenon, it seems unlikely that productivity-enhancing downsizing is more common than newly-productive firms adding to their workforce. This is one area for future research.

analytically how quality of newly hired workers relates to the quality of incumbent workers at a firm, and how changing the frequency of productivity shocks changes matching patterns. These implications have not been well studied in the literature, and represent avenues for future research.

This paper also addresses the question of *recruiting intensity.* The recent work of Davis et al. (2013) has called into question some of the basic assumptions of the workhorse search model. They find that fast growing establishments tend to fill their job openings more quickly. This is at odds with the textbook search theory, which assumes the vacancy fill rate is exogenous to the firm and consequently employment growth is linear in vacancy postings at the firm level. The conclusion of Davis et al. (2013) is that firms are using other margins to attract workers, besides simply posting more vacancies. These margins may include advertising intensity, selectivity, or wage offers. The results in this paper show that selectivity does indeed change with firm growth, but not in the direction one might expect: faster growing firms hire *better* workers on average. The theory reconciles the two empirical results through the wage margin. According to the theoretical model, a faster growing firm is willing to pay higher wages to all workers. Under positive assortative matching this allows the firm to attract high-type workers, while they continue to hire low type workers at the same time. Thus, faster growing firms fill vacancies faster by offering better wages and hiring more of the workers they meet.

1.1 Literature Review

The empirical literature on assortative matching in the labor market has developed quickly since Abowd et al. (1999). They used linked employer-employee data, similar to mine, to estimate worker and firm fixed effects. For example, if w_{ijt} is the log wage of worker i at firm j in period t and x_{it} is a vector of time-varying worker characteristics, an AKM regression has the form

$$w_{ijt} = x_{it}\beta + \theta_i + \psi_j + \varepsilon_{ijt}$$

where ε_{ijt} is an error term and θ_i and ψ_j are worker and firm fixed effects that can be estimated under some mild identification assumptions. The estimated worker fixed effect $\hat{\theta}_i$ measures how much more or less the worker is paid relative to his coworkers (holding x_{it} constant). The firm fixed effect $\hat{\psi}_j$ measures the average wage premium the firm pays, compared to other firms. If θ_i reflects worker productivity and ψ_j firm productivity then the empirical correlation $\text{Corr}\left[\hat{\theta}_i, \hat{\psi}_j\right]$ measures whether good workers match with good firms. If $\text{Corr}\left[\hat{\theta}_i, \hat{\psi}_j\right]$ is large and positive, then productive workers tend to end up at productive firms. If is it close to zero then workers are not strongly sorted across firms. AKM found $\text{Corr}\left[\hat{\theta}_i, \hat{\psi}_j\right]$ to be small and *negative.* Subsequent work (see Abowd et al. (2002), Susana Iranzo and Tosetti (2008), Andrews et al. (2008), Woodcock (2008), Lopes de Melo (2013)) largely confirmed AKM's finding in other data sets, finding negative or small positive correlations between worker and firm effects. These results have been especially puzzling given our tacit intuition that positive sorting is the norm, as least in some markets.

Many criticisms have been made of the AKM methodology as applied to sorting. Researchers have corrected the AKM methodology to address three major issues: Andrews et al. (2008) and Andrews et al. (2012) study limited mobility bias, Woodcock (2008) takes account of match effects, while Lopes de Melo (2013) addresses nonmonotonicity of wages in firm type.[4] See also Abowd et al. (2012), who interpret the AKM fixed effects in the context of Shimer (2005)'s directed search model. Each of these papers uses a variant of the AKM methodology corrected for some possible biases, while abstracting from the other issues. It appears infeasible to correct the AKM estimator for all three issues simultaneously. Nonmonotonicity of wages has received the most attention in the literature, perhaps reflecting a belief that it is a more serious problem than limited mobility bias and match effects. In an important contribution, Lopes de Melo (2013) shows that when the data is generated from a search model such as Shimer and Smith (2000), the AKM firm fixed effects are poor proxies for firm productivity. However, estimated worker effects appear to accurately reflect worker productivity. Lopes de Melo (2013) shows that in Brazilian data workers with high fixed effects tend to work in firms with other high fixed effect workers. In other words, workers are clustered with similar coworkers. This implies a kind of sorting, but without accurate firm productivity data we cannot tell how the workers are sorting. It might be that the best workers are clustered at the worst firms, or workers may be clustered according firm characteristics unrelated to firm productivity. To summarize, there are several potentially serious problems with using the AKM fixed effect methodology to study sorting. These issues have been addressed separately, but their interactions have not been explored. Correcting for the most serious issue, nonmonotonicity of wages, comes at the cost of only being able to study clustering of workers with similar coworkers, instead of studying sorting of workers across firms of different productivity.

Here I contribute to the literature by proposing an alternative way to identify sorting. I use firm growth to proxy firm productivity and lagged worker wages to proxy worker productivity. These measures have the advantages of being directly observable in the data. In my specifications I find evidence of positive sorting that is independent from the criticisms of the AKM fixed effect methodology. Of course, I face other measurement issues of my own, which I address in Section 2.

Other researchers have departed from the AKM methodology and used other data to identify sorting. Bartolucci and Devicienti (2012) use firm balance sheet data to measure firm type directly. Lise et al. (2013) present an original strategy, using a method of simulated moments estimator to measure sorting from longitudinal worker data without links to firm data. They argue that assortative matching can be identified from the profile of worker wages after a spell of unemployment. Intuitively, workers exiting unemployment tend to be mismatched, and the cross sectional distribution of wages will be compressed. As time passes job to job transitions should result in better wages for all workers. If sorting is important then then productive workers should experience more wage growth, as the degree of mismatch is reduced. Then assortative matching may be identified by the

[4] Eeckhout and Kircher (2011), Lise et al. (2013), Hagedorn et al. (2012) and Bagger and Lentz (2012) also discuss nonmonotonicity of wages, but do not propose any correction to the AKM methodology.

spreading out of wages across workers after an unemployment spell. Since they rely on worker side data only, they use the structure of their model to infer the sign and strength of assortative matching. To keep the model tractable they are forced to abstract from human capital accumulation, which may also produce a spreading out of wages. In contrast, I exploit the matched nature of my data to measure sorting in an arguably more direct fashion. My proxy for worker productivity is a lagged moving average of wages, which allows human capital to evolve over time. In addition, Lise et al. (2013) abstract from firm growth, which is part of my focus.

In closely related work, Bagger and Lentz (2012) propose a structural search model and estimate it on linked Danish data. The model is based on Lentz (2010), where worker search intensity is endogenous. With complementarities in production, high productivity workers gain more from searching on the job, and consequently search more intensely and tend to match with better firms faster. Firm dynamics and productivity shocks are not modeled. Their estimation of the model requires a large number of auxiliary models and moments. My work complements theirs by focusing on directly observable moments in the data and presenting an alternative theoretical model. While Bagger and Lentz (2012) focus on search intensity and search while employed, I focus on transitions from unemployment and the role of firm dynamics.

The theoretical literature on assortative matching has grown along with the applied literature. In his influential study, Becker (1973) examined assortative matching in the context of a frictionless marriage model with transferable utility. Shimer and Smith (2000), Shi (2001), Shimer (2005), Atakan (2006), and Eeckhout and Kircher (2010) reexamine Becker's results under various kinds of search frictions, while maintaining the assumption of one-to-one matching and no productivity dynamics. Usually stronger forms of complementarities are needed to ensure sorting when agents face search frictions. In my theoretical work I extend the widely used model of Shimer and Smith (2000) to incorporate multi-worker firms and firm productivity shocks. I find that introducing multi-worker firms considerably simplifies some of the technical complications in Shimer and Smith (2000), allowing me to characterize matching patterns and derive comparative statics in a straightforward fashion. Lentz (2010) and Bagger and Lentz (2012) discuss a similar tractability result in a different model, but do not address the firm dynamics that are my focus. Lise and Robin (2013) and Eeckhout and Kircher (2012) also allow for multi-worker firms, but again abstract from idiosyncratic firm dynamics. Lise et al. (2013) allow for match productivity shocks in a sorting model, but do not present analytical results and do not use data that (directly) measures firm dynamics. To summarize, I contribute to this fast-growing literature by modeling sorting with multi-worker firms that experience productivity shocks, and deriving analytical results.

A relatively new literature studies how firms fill job openings. Davis et al. (2013) document that faster-growing firms fill their vacancies more quickly. This suggests that growing firms are using other margins to attract workers, besides just posting more job openings. Kaas and Kircher (2011) develop a theoretical directed search model, where growing firms post higher wages to attract homogeneous workers more quickly. Belzil (2000) documents that faster growing firms do indeed seem

to pay identical workers higher wages.[5] Brown and Matsa (2012) show that job postings of financially distressed firms attract fewer and less qualified applicants, suggesting that a shrinking, distressed firm will have to wait longer to fill its vacancies. In this paper I develop a related point, that fast growing firms hire better workers. By itself, this does not address vacancy fill rates. But my theory implies that fast growing firms pay better wages (similar to Kaas and Kircher (2011), albeit in a random search context) and are thus able to hire more of the workers they meet, increasing the speed with which vacancies fill.

Firm employment growth is the first difference of firm size, so my work on firm growth is related to the extensive literature on firm size and wages (see Brown and Medoff (1989) and Abowd et al. (1999), Oi and Idson (1999) and Gordanier (2012) among others). This literature has typically found that larger firms pay more and employ better workers, according to a variety of measures. This is consistent with my finding that faster growing firms (i.e. firms that are getting larger) hire better workers. However I find that the firm size and firm growth effects are distinct, since the growth effect survives the inclusion of firm fixed effects and controls for (beginning of period) firm size. In contrast, firm size effects are imprecisely estimated and ambiguously signed in this specification. This pattern suggests that part of the positive worker quality-firm size correlation found in the literature is due to unobserved time-invariant firm characteristics, rather than a firm's size at any particular point in time. This is consistent with models in the literature where firm size proxies productivity, opportunities for workers to specialize, or other unobserved characteristics.

Finally, there has been some attention devoted to the hiring policies of start-ups (see Dahl and Klepper (2007)) and high-growth firms (see Coad et al. (2011)). These papers study the correlation between indicators of a firm's long run prospects (long run growth, for example) and the characteristics of their hires. In other words they study primarily cross-sectional differences in hiring. In contrast, my focus is on within firm variation in growth and worker productivity. In addition, I emphasize the need for nonparametric methods to accurately identify the relationship between growth and quality of hires. Some of the most interesting findings in Section 2.4 are highly nonlinear relations that would not be detected in a linear specification.

2 Empirics

My interest is in the relationship between firm growth and the productivity of new hires. I will show that when a given firm grows faster, it hires more productive workers. The productivity of the new hires is a characteristic of the workers themselves, not the firm they are hired into. To identify this phenomenon I need to measure firm growth and the productivity of newly hired workers. In addition, the data allow me to control for a number of observable and unobservable firm characteristics. Controlling for firm characteristics is critical to properly measuring the ceteris paribus impact of a change in firm growth on the quality of hires.

[5] Belzil uses Danish administrative data on matches and wages. The LEHD, which I use, only has quarterly earnings (and no hours data) so I cannot replicate Belzil's wage results.

I take the quality of new hires to be the dependent variable, and regress it on firm growth and a vector of controls. The empirical specification has the form

$$Q_{j,t} = m(g_{j,t}) + \beta \mathbf{Z}_{j,t} + \varepsilon_{j,t}$$

where $Q_{j,t}$ is a measure of the productivity of new hires at firm j in period t, $g_{j,t}$ is the employment growth rate of firm j in period t, $\mathbf{Z}_{j,t}$ is a vector of control variables, and $\varepsilon_{j,t}$ is the error term. The vector β are the coefficients on the controls and $m(\cdot)$ is an unknown function to be estimated. The function $m(\cdot)$ summarizes the relationship between firm growth and the quality of hires, so it is the focus of what follows.

The main research question is what happens to $Q_{j,t}$ when $g_{j,t}$ changes, holding all else constant. Thus it is important that $\mathbf{Z}_{j,t}$ include extensive controls for firm characteristics. The data I use includes industry, location, size and some other firm information. However, there is likely to be a large amount of unobserved firm heterogeneity imperfectly correlated with these characteristics. This cross-firm heterogeneity in growth rates and $Q_{j,t}$ will bias any estimate of $m(\cdot)$. Motivated by this concern I include firm fixed effects in my main specifications. These specifications use only within firm variation to identify $m(\cdot)$, and are thus robust to unobserved cross sectional heterogeneity. Inclusion of firm fixed effects brings the empirical specification closer to the thought experiment of the ceteris paribus impact of a change $g_{j,t}$ on $Q_{j,t}$.

It is straightforward to construct a firm employment growth variable, $g_{j,t}$, from the data. More complicated is the construction of $Q_{j,t}$, the productivity of new hires. In this context worker productivity are those characteristics that make the worker more productive regardless of job. Thus $Q_{j,t}$ does not include match-specific productivity, or characteristics that are only valued by a narrow range of jobs. In the theoretical part of this paper I capture productivity by assuming that a higher "type" worker produces more on any given job than a low type worker. Thus, high type (high productivity) workers have an *absolute advantage* in all jobs. This simple setup ensures that the ranking of workers is unambiguous, and is widely used in the literature.

Intuition and the models discussed in Section 1.1 suggest that high productivity workers should earn more on average. If a worker produces more in all jobs, he ought to earn at least as much as a lower productivity worker. Then a worker's average earnings are a noisy proxy for his productivity. My measure of the quality of hires is based on this idea. I build the quality of hires variable in two steps. First, I construct a productivity measure for each worker in each period. Since my data are quarterly this gives me a measure of a worker's productivity in each quarter. This measure is essentially a backward looking moving average of earnings. Second, I find the set of workers hired by firm j in quarter t, and average their productivity measure. This gives me the average productivity of workers hired by the firm in that quarter.

Formally, let

$$q_{i,t} = \frac{1}{8} \sum_{k=1}^{8} earn_{i,t-k} \tag{1}$$

where $q_{i,t}$ is the worker's measured quality, and $earn_{i,t-k}$ is the worker's total earnings (in all jobs) in quarter $t-k$. $q_{i,t}$ is thus a two year lagged moving

average of earnings. I have experimented with longer and shorter moving averages and found no significant effect on my results. Longer horizons for the moving average reduce the noise from search frictions, unpaid leave, and firm-level shocks. However, when there is human capital accumulation (and depreciation) the most recent earnings are the best indicators of current productivity. The eight quarter horizon was chosen to balance these two concerns.

Let $H_{j,t}$ be the set of workers *hired* by firm j in period t. Then my quality of hires variable can be written as

$$Q_{j,t} = \ln\left(\frac{1}{|H_{j,t}|} \times \sum_{i \in H_{j,t}} q_{i,t}\right).$$

This is simply the log average of $q_{i,t}$ over the set of hires. In practice I will use several variants of $Q_{j,t}$ to address concerns about measurement and selection. These variants are denoted by superscripts.

2.1 Nonparametric Estimation

Given $g_{j,t}$, $Q_{j,t}$ and $\mathbf{Z}_{j,t}$ I am prepared to estimate $m(\cdot)$. As will be made clear in the results section $m(\cdot)$ is highly nonlinear, so a linear regression is inappropriate. Here I follow Davis et al. (2013) and reformulate the problem as a dummy variable regression. Intuitively, the idea is to partition the domain of $m(\cdot)$ into narrow intervals (one percentage point wide in my application) and create indicator variables for each interval, or bin. An indicator variable is set to 1 if $g_{j,t}$ falls into the bin and zero otherwise. Then I regress $Q_{j,t}$ on the full set of indicator variables and $\mathbf{Z}_{j,t}$. The coefficient on a given indicator is then the sample average of $m(\cdot)$ within that bin conditional on $\mathbf{Z}_{j,t}$.[6] Plotting the bin coefficients yields an estimate of $m(\cdot)$ on a discretized domain. where f_i are dummy coefficients to be estimated. All reported regressions are weighted by the number of hires. Unweighted and employment weighted results are similar.

2.2 Measurement and Selection Issues

Before turning to the data and results, there are several potential measurement problems that must be addressed. I show how alternative samples and dependent variables can be used to avoid some measurement problems. Where there are not good alternatives, I present evidence that the problems are not too severe in the data.

Recall that $Q_{j,t}$ is the log average of worker's earnings in all quarters, regardless of whether they were employed. In other words, it captures average earnings, not

[6] This formulation is similar to but distinct from more standard nonparametric methods, such as kernel regression. I am using comprehensive administrative data and include firm fixed effects in my regressions, which can result in millions of fixed effects. Under these circumstances estimating ordinary least squares requires special procedures. Estimating a kernel regression in the presence of this many fixed effects would not be feasible without additional specialized software. Ultimately, the dummy variable regressions are a convenient way to capture the shape of the relationship.

average earnings while employed. This choice is motivated by the fact that the extensive margin contains information about a worker's productivity. If a worker is unable to find a job, it tells us he may not be a high quality employee. In addition, in theoretical models average lifetime earnings are monotone increasing in worker productivity, but earnings while employed may not be. However, if shocks unrelated to worker productivity (e.g. health shocks or fertility) move the worker in and out of employment, then average earnings may not be an accurate proxy for productivity. To examine this possibility I use a measure of worker quality that excludes nonemployment. As an alternative to $q_{i,t}$ I define

$$q_{i,t}^{FQ} = \frac{\sum_{k=1}^{8} earn_{i,t-k}^{FQ}}{\sum_{k=1}^{8} \mathbf{1}\left\{earn_{i,t-k}^{FQ} > 0\right\}}$$

where $earn_{i,t}^{FQ}$ is the earnings of the worker in jobs that lasted all of quarter t. Thus, periods of nonemployment and partial employment are excluded. The corresponding quality of hires measure is

$$Q_{j,t}^{FQ} = \ln\left(\frac{1}{|H_{j,t}^{FQ}|} \sum_{i \in H_{j,t}^{FQ}} q_{i,t}^{FQ}\right) \tag{2}$$

where $H_{j,t}^{FQ}$ is the subset of hires that had at least one quarter of full quarter employment in the previous two years. I report regressions using both all quarter earnings and full quarter earnings.[7] While the magnitudes of the effects differ, the qualitative patterns are insensitive to the choice of variables.

Relatedly, note that $q_{i,t}$ will be a poor measure of productivity for new entrants to the full time labor force. Recent high school or college graduates will have zero or low lagged earnings for reasons unrelated to their desirability as workers. To mitigate this problem I only consider hires of workers aged 26 or older. The results are not sensitive to dropping this restriction, or replacing it with a requirement that the worker have positive earnings two or more years before being hired.

Next, there is a clear endogeneity issue when the worker's wages depend on the firm they work for. For example, if fast growing firms pay more, then wages of current employees will be positively correlated with firm growth, even if there are no differences in worker productivity. This concern motivates my use of a *lagged* moving average to measure the quality of new hires. $Q_{j,t}$ contains only wages from periods $t-1$ and earlier, before the workers were hired by firm j. Thus fluctuations in firm j's wages do not directly contaminate $Q_{j,t}$.

However, we might be concerned about the workers' previous employers. $q_{i,t}$ is a function both of a worker's inherent productivity as well as their luck in the labor market. As emphasized by the voluminous wage dispersion literature,[8] workers of

[7] Both $Q_{j,t}^{FQ}$ and $Q_{j,t}$ involve taking a logarithm of earnings, so observations where the firm only hires zero-earnings workers are dropped. It might be wondered whether this creates a selection problem. Alternative specifications suggest that this is not a serious problem. In particular, the empirical results are largely unchanged when I use the raw average of worker earnings as the dependent variable. The same applies for the inverse-hyperbolic sine transformation.

[8] See Mortensen (2003) and the literature reviewed therein. Burdett and Mortensen (1998) and Moscarini and Postel-Vinay (2013) each present models where homogeneous workers are paid differently.

identical productivity may be paid different wages depending on which employers they find. Equally productive workers will have different $q_{i,t}$ because they had different previous employers, and have spent different durations in nonemployment. Thus we should think of $q_{i,t}$ as measuring the worker's true productivity plus some noise. Measurement error alone in $q_{i,t}$ will not likely cause serious bias, since $q_{i,t}$ goes into the dependent variable in my regressions.[9] But with on the job search there is reason to think that a worker's lagged wage, including the "luck" component, influences his choice of whether to be hired. In particular, if a worker has a continuing job in period $t-1$, his wage at that job can act as a reservation wage: workers will refuse any offer below what they currently earn. If firm wages are positively correlated with firm growth (as Belzil (2000) documents) then only a fast growing firm will be able to hire the worker away from the current employer. This means that the occurrence of a hire depends on the worker's $t-1$ wage, *including any luck component of the wage.* Thus, faster growing firms will hire more workers from high wage firms, regardless of worker productivity. This creates a positive correlation between firm growth and $Q_{j,t}$, independent of any true differences in worker productivity. To address this issue I restrict attention to hires from nonemployment.[10] Nonemployed workers do not have a current wage to act as a threat point, and so their choice of whether to be hired is a function of their productivity, not their (random) current employer. All the empirical results reported below calculate $Q_{j,t}$ and $Q^{FQ}_{j,t}$ on the subset of hires that came from nonemployment. Results including job-to-job hires (not reported) are qualitatively similar.

Variation in the reservation wage is another possible concern. It may be that workers with high lagged wages have higher reservation wages even after entering nonemployment. This could be because the worker falsely believes, based on their previous luck, that they can find another high wage job easily (see Hogan (2004)). Or it could simply be that some workers value leisure more, and so demand higher wages. These mechanisms will induce a positive correlation between lagged earnings while employed and firm growth, since only high growth, high wage firms can meet the reservation wages of these workers. But these mechanisms would also induce a weaker or negative correlation between lagged *all quarters* earnings and firm growth, since the more selective workers would spend more time unemployed. My empirical results run opposite to this prediction. I find that both all quarters earnings and full quarters earnings are positive functions of firm growth, and that all quarters earnings increase faster than full quarter earnings. Thus, while I cannot rule out variation in reservation wages, my results are inconsistent with reservation wages being the dominant force relating firm growth to measured

[9] $Q_{j,t}$ is a nonlinear transformation of $q_{i,t}$, so measurement error is not entirely innocuous. But note that the hire occurs at a large firm hiring multiple workers simultaneously. As more workers are hired, (uncorrelated) measurement error averages out of $\frac{1}{|H_{j,t}|} \times \sum_{i \in H_{j,t}} q_{i,t}$ before the log transformation.

[10] The LEHD data I use only include data on employment status, not search activity. Thus I cannot distinguish between unemployed workers and those out of the labor force. All workers without jobs are taken together as nonemployed. Information on search activity would be useful for my purposes, but not critical. Recent work has shown that while self-reported search activity does predict future employment status, the link is noisy and unstable over time. Workers who exit the labor force are more likely to have low productivity, which my quality measure $q_{i,t}$ will capture.

worker quality.

2.3 Data and Sample

I use the linked employer-employee data from the Longitudinal Employer-Household Dynamics (LEHD) program at the U.S. Census Bureau. These files are derived from state level unemployment insurance records, which require firms to report quarterly wage and salary earnings of each covered employee to the state unemployment agencies. The main LEHD files consist of these earnings records, along with worker identifiers and firm identifiers. Limited worker and firm characteristics are also included. The worker identifiers are consistent across firms, time and states, so it is possible to identify the same worker at different points in time, or at different firms. The firm identifier, called a state employer identification number (SEIN) is similarly consistent across time, allowing me to follow an SEIN over time. I will use the SEIN as my firm concept in this paper.[11]

Importantly, the LEHD files cover the vast majority of private non-farm jobs. Estimates put the coverage of such jobs at 96%. The main exclusions are federal workers, agricultural workers, and the self-employed. I draw firm (SEIN) data from six states: DE, IL, OH, IN, WI, VT. Importantly, when I construct worker earnings histories I draw on data from all states participating in the LEHD. This means that if a worker moves from New York and is hired by a firm in Chicago I include their New York earnings history in $q_{i,t}$. My sample period runs from 2000:Q1 through 2008:Q2, the last quarter of data available in the research snapshot I use. The results are not sensitive to the inclusion of the Great Recession.

I use as broad a sample as possible. In the reported results I only drop quarters where a given firm enters or exits, since there is good reason to think that selectivity and growth are determined by other factors in those quarters. The results are qualitatively unchanged when much more stringent restrictions are imposed.

The construction of firm growth, hires from nonemployment and other variables requires some data processing. The main issue is accounting for the fact that the LEHD variables are reported quarterly, so they do not necessarily approximate high frequency, point-in-time data. For example, the raw LEHD earnings records for a quarter include all workers employed *for any duration* at a given firm in a given quarter. This does not necessarily reflect the firm's actual employment at any point in time during the quarter. The solution is to define "beginning of quarter" and "end of quarter employment".[12] In quarter t the beginning of quarter employment is the number of workers observed to work at the firm both in $t-1$ and t. End of quarter employment is the number of workers employed both in

[11] The SEIN identifies the reporting entity for UI purposes. In single unit firms, the SEIN is synonymous with the firm. However, multi-unit firms may have several SEINs, and each SEIN may include several establishments (though the same establishment is never a part of two SEINs). Thus, a SEIN falls somewhere in between an establishment and a firm. My results are qualitatively unchanged if I restrict the sample to include only SEINs that did not report a change in the number of establishments. This sample restriction brings the data as close as possible to an establishment definition.

[12] See http://lehd.ces.census.gov/doc/QWI_101.pdf for further documentation, and the use of these concepts in public-use data.

Variable	Mean	Std. Dev.
Hires Rate	0.17	0.25
Separation Rate	0.17	0.24
Employment Growth Rate	0.00	0.19
NE share of hires	0.53	0.22
EE share of hires	0.22	0.20

Table 1: Worker Flows

t and $t+1$. These measures more or less reflect firm employment on the first and last day of the quarter. Additional data processing details are in Appendix Appendix A.

Table 1 presents employment weighted summary statistics on worker flows in my sample. As has been well documented, these flows are large: a firm typically hires and separates 17% of its workforce over the course of a quarter. The exact values are somewhat lower than those reported in Lazear and Spletzer (2013) (who use similar data), in part because I exclude entry and exit quarters. Average employment growth is indistinguishable from zero. This is partially a reflection of the states in the sample, and again partially due to my exclusion of entry and exit quarters. Importantly, the standard deviation of employment growth is 19 percentage points. To exclude outliers I will focus on the interval between -30% and 30% growth in my graphical results.

The last two rows of Table 1 report the share of hires that come from nonemployment (NE Hires) and from other jobs (EE Hires). NE hires account for 53% of total hires, while EE hires account for 22%.[13] The remainder are hires that could not clearly be assigned to either of these categories. The shares of hires from nonemployment and employment are roughly in line with those reported by Fallick and Fleischman (2004) and Bjelland et al. (2011).

Table 2 reports summary statistics for the characteristics of new hires. The top two rows summarize the main dependent variables in the regressions: log average earnings ($Q_{j,t}$) and log average full quarter earnings ($Q_{j,t}^{FQ}$). Recall that these are logged averages of newly hired workers' lagged wages. They are calculated using only hires from nonemployment, to mitigate the selection concerns raised in Section 2.2. The sample average of $Q_{j,t}$ is 8.06, while the average of $Q_{j,t}^{FQ}$ is 8.82. The difference between the two implies that, roughly speaking, average earnings while employed are 80% greater than average total earnings. Turning to the last two rows, this difference is clearly attributable to the fraction of time workers spend employed. "Fraction of Quarters Employed" is the average fraction of quarters a new hire had any employment. Like $Q_{j,t}$ and $Q_{j,t}^{FQ}$, this fraction is calculated on the 8 quarter lagged window. "Fraction of Quarters Fully Employed" is the fraction of quarters the worker held a job throughout the entire quarter. The means show that newly hired workers (from nonemployment) spent between 40%

[13] Some of these are workers who held a job for several periods, then were hired at a new firm but continued working at their old job. I exclude these workers for the same reason I exclude EE hires, to avoid the selection effect of lagged wages.

and 57% of the past 2 years at work. This explains the significant differences in the levels of $Q_{j,t}$ and $Q_{j,t}^{FQ}$, since only $Q_{j,t}$ includes time spent nonemployed.

Variable	Mean	Std. Dev.
$Q_{j,t}$	8.06	1.03
$Q_{j,t}^{FQ}$	8.82	0.70
Fraction of Quarters Employed	0.57	0.21
Fraction of Quarters Fully Employed	0.40	0.18

Table 2: Characteristics of New Hires

2.4 Results

In my baseline specifications I estimate $m(\cdot)$ using $Q_{j,t}$ and $Q_{j,t}^{FQ}$ as the dependent variables. In $\mathbf{Z}_{j,t}$ I include an extensive set of controls, including location (state), industry, quarter, interactions of all three, flexible controls for firm size, interacted industry and seasonal dummies, and firm fixed effects.[14] After these controls are included, $m(\cdot)$ is identified by within firm variation that is not correlated with aggregate, industry or location effects. Figure 1 plots the estimates of $m(\cdot)$ using $Q_{j,t}$ and $Q_{j,t}^{FQ}$ as the dependent variables. The horizontal axis is firm employment growth, restricted to the [-30%,30%] interval. The vertical axis measures $Q_{j,t}$ (All Quarter Earnings) and $Q_{j,t}^{FQ}$ (Full Quarter Earnings). The estimates of $m(\cdot)$ have been normalized to average zero on the interval [-30%,0].[15]

For both specifications $m(\cdot)$ is relatively flat when growth is negative, but increases when growth is positive. Both relationships are overwhelmingly statistically significant. The t-statistic for the average slope on the [-30%,30%] interval is 9.37 for $Q_{j,t}^{FQ}$ and 40.98 for $Q_{j,t}$. Thus the positive average relationships shown in Figure 1 are significant well beyond conventional thresholds.

In terms of magnitudes, the results for $Q_{j,t}$ imply that when a firm goes from zero growth to 30% growth they hire workers with roughly 9% higher lagged earnings. Using $Q_{j,t}^{FQ}$, the same change in employment growth leads to a 2% increase in lagged full quarter earnings. While both are statistically significant, the magnitude of the effect for $Q_{j,t}$ is larger. This is due to the inclusion of nonemployment. When a firm grows more quickly, they not only hire workers with higher wages in their old jobs (the $Q_{j,t}^{FQ}$ results), but they also hire workers who have experienced less nonemployment.

[14] Firms change industries, so industry is not perfectly collinear with the firm fixed effects. The interactions of industry with other variables are decidedly not collinear with firm fixed effects.

[15] For ease of presentation Figure 1 shows the three bin moving average of $m(\cdot)$, rather than the raw estimate. This has the effect of smoothing the relationship slightly, but this effect is quantitatively and qualitatively insignificant.

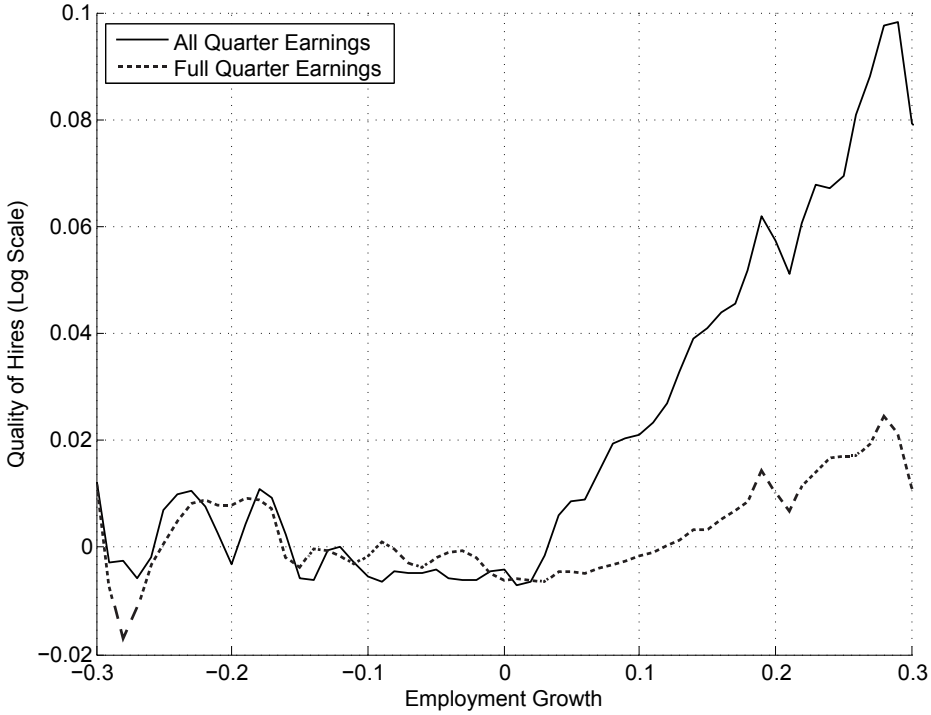

Figure 1: Baseline Specification

It might have been expected that increased firm growth would be associated with *lower* quality of hires, as growing firms exhaust their pool of potential employees and are forced to lower hiring standards. The results in Figure 1 show this is decidedly not the case. Instead, the results appear to support positive assortative matching. Firms which grow more quickly are presumably more productive, and are observed to be hiring more productive workers. In other words better workers sort themselves into more productive firms. As discussed earlier, the empirical literature has had difficulty identifying assortative matching. This paper contributes to that literature by showing evidence of assortative matching from a novel specification.

The results in Figure 1 and the associated t-tests establish that, *conditional on the control variables* ($\mathbf{Z}_{j,t}$), faster growing firms hire better workers. Now I turn to the importance of the control variables, in order to understand whether the positive growth-quality of hires correlation holds in the cross section. Figure 2 plots the estimated $m(\cdot)$'s when no control variables are included. It is apparent that the relationship is highly non-monotonic. These estimated $m(\cdot)$'s imply that low absolute values of growth are associated with better quality hires. Large positive or negative values of growth are associated with lower quality hires, with positive values having somewhat better average quality. A linear regression confirms that the average relationship on this interval is still positive and statistically significant, despite the non-monotonicity. Taken together, Figures 1 and 2 show that controlling for firm characteristics and other factors matters a great deal. Omitting controls would lead to the conclusion that the *absolute value* of growth is the most important factor influencing the quality of hires at a given firm. But this is not the case. The correct comparison in Figure 1 shows a monotone increasing relation between growth and quality of hires, once all else is held constant.

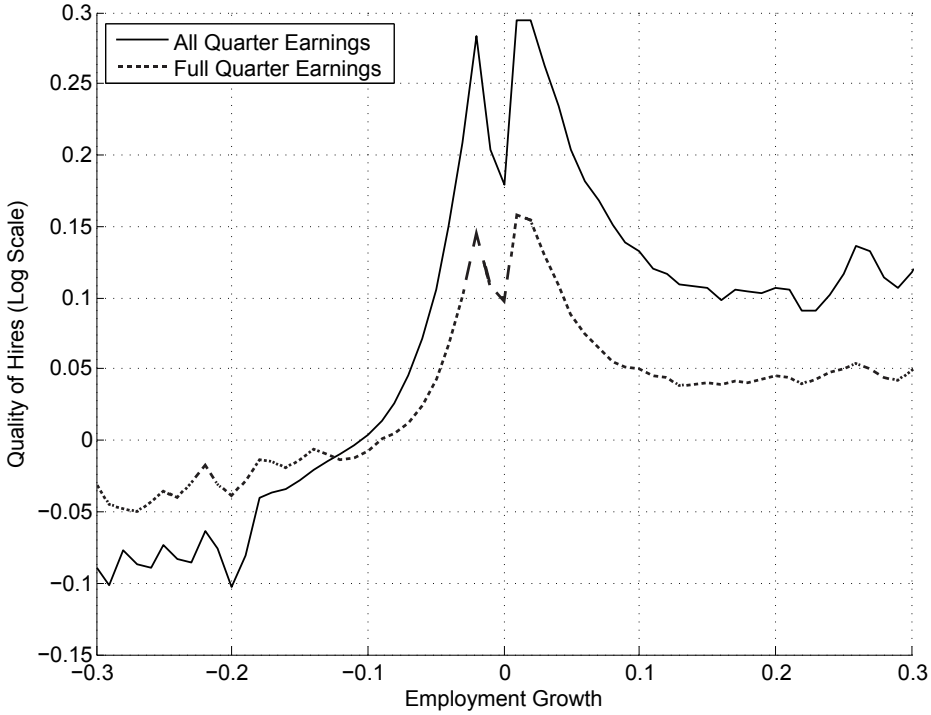

Figure 2: No Controls

Figures 1 and 2 are extreme cases. Figure 2 has no controls, while Figure 1 includes controls for many firm characteristics, as well as seasonal, industry-level and aggregate fluctuations. There are many possible intermediate specifications, some of which I discuss in Appendix Appendix B. As more controls are added, the results look less like Figure 2 and more like the baseline specification. Controls for aggregate and industry level fluctuations are relatively unimportant. In contrast, firm fixed effects and controls for firm size and interacted industry-period-location dummies are critical in explaining the difference between Figures 1 and 2. These are all *firm characteristics*. The interpretation is simple: as more firm characteristics are included, the specification is closer to the ideal of holding all else constant.

2.5 Discussion

The empirical results demonstrate the importance of conditioning on firm characteristics. Without firm controls the relation between growth and quality of hires is ambiguous and non-monotonic. This echoes the results in the assortative matching literature, where AKM-style regressions find little correlation between worker type and firm type. AKM regressions typically do not condition on firm size, industry or other firm characteristics, so all inter-industry or cross-size variation is loaded into the estimated fixed effect. In addition, AKM-style regressions cannot include firm fixed effects as controls, since the firm fixed effect is the variable of interest.[16] In this way the AKM models are similar to my specification without

[16] Mortensen et al. (2010) extend the AKM methodology to estimate worker fixed effects and firm-year fixed effects. With firm-year effects it would be possible to use time variation to study

controls, where all inter-industry or cross-size variation is included. My specification without controls can be seen as an indirect replication of the typical AKM result: when no firm characteristics are conditioned on, there is little evidence of assortative matching.

The importance of firm characteristics may reflect segmentation in the labor market. When I condition on firm characteristics I implicitly compare firms only to very similar firms, which are likely to inhabit similar labor markets. This holds for firm fixed effects as well, under the assumption that firms do not change markets over time. Thus the baseline specifications, controlling for industry and other characteristics, pick up positive sorting within each segment of the labor market. The best doctors work for the best hospitals, the best pilots work for the best airlines, and so on. When I do not control for firm characteristics I am comparing doctors to pilots and hospitals to airlines. Firms in different markets may hire different workers and have different growth patterns for technological or other reasons. The comparison of Figures 1 and 2 suggests that markets with low firm volatility (growth near zero) are markets with higher quality workers. In this paper I do not seek to explain the differences between labor markets, or the entire joint distribution of firm growth and worker productivity. Rather I focus on the new fact that within markets (i.e. conditioning on firm characteristics), there does appear to be assortative matching. Based on my results, the AKM methodology may find different results when firm effects are estimated conditional on firm characteristics.

3 Theory

In this section I describe a general model that can account for the positive firm growth-quality of hires correlation described in Section 2. The model captures the intuition that if worker and firm productivity, are complementary in production then fast growing, highly productive firms should tend to hire more productive workers. Thus the model tightly links assortative matching to firm growth and the quality of hires. The model extends the influential work of Shimer and Smith (2000) to include multi-worker firms, idiosyncratic firm productivity shocks, and firm dynamics. Once solved, the extended model turns out to be highly tractable and yields analytical comparative statics. I derive sufficient complementarity conditions that ensure assortative matching and allow the model to replicate the observed growth-quality of hires correlation. Previously Shimer and Smith (2000) and Lentz (2010) have derived sufficient conditions for assortative matching, but their models abstracted from firm dynamics.

In the model workers and firms are characterized by a scalar type, interpreted as productivity. Search is random, and when workers and firms match the wage is determined by generalized Nash bargaining. Only unemployed workers search, though the qualitative results are unchanged if search while employed is allowed.

sorting while controlling for a time-invariant firm fixed effect. This is a potentially promising approach but it faces the same problems as the original AKM approach, as described in Section 1.1.

I also allow firms to employ multiple workers by posting costly vacancies. Production is linear, so firm size is a function of the convex vacancy posting costs.

The most important feature of my model is that firms can employ multiple workers, whereas Shimer and Smith (2000) and others focus on single worker firms, or firms that cannot freely post additional vacancies. This difference turns out to matter a great deal. The reason is that when firms can post additional vacancies (and production is linear), they have no opportunity cost of hiring a given worker. A central feature of the Shimer and Smith (2000) model is that both parties face a tradeoff between matching immediately or waiting for a better partner. By contrast, in my model, firms can hire a given worker today *and continue searching* (posting vacancies) in subsequent periods. This eliminates the opportunity cost of hiring for firms that are in contact with workers, greatly simplifying the analysis. A similar point is made by Lentz (2010) in a different context.

Section 3.1 introduces the model and derives the worker and firm decision rules. Section 3.2 briefly summarizes the steady-state equilibrium. In Section 3.3 I further characterize the worker's acceptance choices, and derive sufficient conditions for assortative matching. In Section 3.4 I turn to firm type and firm growth. In the body of the paper I skip some details of the firm's problem and sometimes focus on a model without productivity shocks, for the sake of exposition. Appendix Appendix C fully analyzes the firm's problem, and Appendix Appendix D analyzes sorting under productivity shocks. All results in the body of the paper carry through in the model with productivity shocks. Omitted proofs may be found in Appendix Appendix E.

3.1 Environment

Workers and firms are risk neutral. Worker types are indexed by x. Firm types are indexed by y. Without loss of generality, I can assume that $x \in [0,1]$ and $y \in [0,1]$.[17] Let $l(x)$ be the mass of workers of type x, and let $h(y)$ be the density of firm types. I normalize the mass of firms to unity, but allow the total mass of workers to vary. When a worker and firm are matched, their per-period output is $F(x,y) > 0$. I assume h, l and F are twice continuously differentiable. I also assume $F_1 > 0$ and $F_2 > 0$, so that higher types have *absolute advantage* over low types.

Each firm potentially employs a continuum of workers. Production of each worker-firm pair is independent of the firm's workforce. Let $N(x)$ be the mass of workers of type x working at a firm in a given period. Then total output at the firm is

$$Y = \int_0^1 N(x) F(x,y)\, dx$$

Within any match, the wage $w(x,y)$ is negotiated via Nash bargaining. Firms choose the level of vacancies to post each period. Each period some matches end due to exogenous separation shocks. Other matches end endogenously, when the worker and firm find it is more profitable to separate.

[17] See Hagedorn et al. (2012) for a discussion of normalizing agent types in this type of model.

The main objects of interest are the *acceptance sets.* In equilibrium, each worker type x will be willing to match with a subset of firm types y. Let $A(x) \subset [0,1]$ be the firm types that worker type x accepts offers from. The complement $A^c(x) = [0,1] \setminus A(x)$ is the rejection set. Upon meeting, a worker and firm directly observe each other's types.

3.1.1 Productivity Shocks

Firm type, or productivity, is not constant over time. Each period there is a probability p that a given firm draws a new productivity value. The new value is drawn from the exogenous density $h(y)$. Note that when there is no firm exit, the density of firm types is given by $h(y)$ as well. Conditional on drawing a new type, the new type is independent of the firm's old type (and everything else). A low value of p implies that productivity is more persistent. This process is a tractable way to model a persistent, mean-reverting process. It has been used by Mortensen and Pissarides (1994), Lise et al. (2013) and Elsby and Michaels (2013) among others.

3.1.2 Timing

Time is discrete. Within a period, the timing of events is as follows:

1. At the beginning of period t firm type y is realized.
2. Exogenous separations occur. Separated workers cannot search within the same period. They enter the matching market in $t+1$.
3. Firms choose vacancy levels. Unemployed workers search, and meetings occur. Workers and firms who meet choose who they are willing to match with, anticipating the wage bargain and the future path of productivity.
4. Nash wage bargaining. After bargaining but before production either party can dissolve the match, resulting in endogenous separations.
5. Production occurs. Wages and unemployment benefits are paid. Consumption occurs.

3.1.3 Value Functions

In this section I briefly derive the firm and worker value functions. I describe the firm's problem in detail in Section Appendix C. Each period the firm posts vacancies v subject to a strictly convex, increasing, and differentiable cost function $c(v)$, where $c(0) = 0$ and $\lim_{v \to 0} c'(v) = 0$. Endogenous vacancy posting is important in what follows. Let $J(y)$ be the marginal net value of a vacancy to a type y firm, inclusive of posting costs. Since the firm is always free to adjust the number of vacancies, in equilibrium it must be that $J(y) = 0$. Otherwise the firm could profit by increasing or decreasing their postings. Let $J(x,y)$ be the firm's match surplus when paired with a type x worker, at the bargaining stage.

It will turn out that $J(x,y)$ is the value both for newly formed matches as well as for continuing matches. If the match were to yield negative surplus for the firm ($J(x,y)<0$), the firm could immediately dissolve the match and obtain a payoff of zero. Thus $J(x,y)\geq 0$ always. Assuming that $J(x,y)>0$ the firm does not dissolve the match, and $J(x,y)$ can be written

$$\begin{aligned} J(x,y) &= F(x,y)-w(x,y)+(1-s)(1-p)\beta J(x,y) \\ &\quad +(1-s)p\beta\int_{A(x)} h(y')J\left(x,y'\right)dy' \end{aligned} \tag{3}$$

where $w(x,y)$ is the wage, and s is the exogenous separation probability. Linearity of production implies that the value of a match is independent of the firm's other workers. The firm value is composed of three terms. The first, $F(x,y)-w(x,y)$, is within-period production less wages. The next two terms represent continuation values. In the next period a separation shock arrives with probability s and a firm productivity shock arrives with probability p. If neither shock occurs (probability $(1-s)(1-p)$) the match remains unchanged, and the firm gets the discounted value $\beta J(x,y)$. Alternatively, if the productivity shock does arrive, the firm's productivity changes to y'. If $y'\in A(x)$ the match continues to be acceptable to the worker and continues, with the firm receiving $J(x,y')$. The match is terminated if the separation shock arrives or the productivity shock makes the match unacceptable to the worker. In either case the firm is left with nothing.

Let $U(x)$ be the value of an unemployed worker, and let $W(x,y)$ be the value of that worker when matched with a type y firm. The value of unemployment satisfies

$$U(x)=\beta\int_{A(x)} f(y)W(x,y)\,dy+\beta\left(1-\int_{A(x)} f(y)\,dy\right)U(x). \tag{4}$$

Here $f(y)$ is the arrival rate of offers from type y firms. Search is undirected so $f(y)$ is independent of worker type. If a worker receives an acceptable offer, he exits unemployment and receives continuation value $W(x,y)$. Otherwise the worker continues in unemployment. The worker receives a flow payoff of zero while unemployed. This assumption is made only for convenience.[18]

Turning to $W(x,y)$, assuming that $y\in A(x)$ so the match is acceptable and $W(x,y)>0$, the value of employed worker is

$$\begin{aligned} W(x,y) &= w(x,y)+(1-s)(1-p)\beta W(x,y) \\ &\quad +(1-s)p\beta\int_{A(x)} h(y')W\left(x,y'\right)dy' \\ &\quad +(1-s)p\left(1-\int_{A(x)} h(y')dy'\right)\beta U(x) \\ &\quad +s\beta U(x). \end{aligned} \tag{5}$$

This equation is similar to the firm value J, with the exception that the worker receives the value $U(x)$ when the match is dissolved. Within the period an employed worker receives the flow wage $w(x,y)$. If the match continues, the worker

[18] If the worker does derive utility from unemployment $b>0$, all matching and sorting results carry through unchanged for the net production function $\tilde{F}(x,y)=F(x,y)-b$.

receives $W(x,y)$ or $W(x,y')$ depending on the productivity shock. Otherwise the worker goes into unemployment. If neither the separation nor the productivity shock arrive, the worker receives the continuation value $W(x,y)$. If the match is acceptable the worker remains at the firm, with value $W(x,y')$. If it is not acceptable the worker separates into unemployment. Finally, the exogenous separation shock moves the worker into unemployment with probability s.

3.1.4 Wage Bargaining

It will be convenient to define the worker surplus:

$$S(x,y) = W(x,y) - U(x).$$

This is the gain to the worker of being matched with a type y firm. Each period the worker and firm split the match surplus according to (generalized) Nash bargaining:

$$(1-\eta)\, S(x,y) = \eta J(x,y) \tag{6}$$

where $\eta \in [0,1]$ is the worker's bargaining power. In equilibrium matches form, or remain intact, if and only if the total surplus is positive

$$J(x,y) + S(x,y) > 0.$$

Some rearrangement of these equations yields expressions for the wage, the worker surplus and the firm surplus:

Fact 1. *Given a worker's acceptance set $A(x)$, the wage satisfies*

$$w(x,y) = \eta F(x,y) + (1-\eta)\beta \int_{A(x)} f\left(y'\right) S\left(x,y'\right) dy' \tag{7}$$

the marginal value of a worker is

$$\begin{aligned} J(x,y) &= \theta(1-\eta)\left(F(x,y) - \beta \int_{A(x)} f\left(y'\right) S\left(x,y'\right) dy'\right) \\ &\quad + \theta(1-s)p\beta \int_{A(x)} h(y') J\left(x,y'\right) dy'. \end{aligned} \tag{8}$$

and the worker surplus is

$$\begin{aligned} S(x,y) &= \theta\eta\left(F(x,y) - \beta \int_{A(x)} f\left(y'\right) S\left(x,y'\right) dy'\right) \\ &\quad + \theta(1-s)p\beta \int_{A(x)} h(y') S\left(x,y'\right) dy' \end{aligned} \tag{9}$$

where $\theta = \frac{1}{1-(1-s)(1-p)\beta} > 0$.

Proof. See the Appendix. □

The expressions (7), (8) and (9) are fairly standard. In equation (7), we see that the bargained wage is a weighted average of current production and $\beta \int_{A(x)} f(y') S(x,y') dy'$, which is proportional to the worker's value of unemployment $U(x)$. This is the typical result in Nash bargaining models. Equations (8) and (9) show that the agents split production less the worker's threat point, $F(x,y) - \beta \int_{A(x)} f(y') S(x,y') dy'$, in accordance with the bargaining parameter η. Finally, note that the future values of productivity do not appear in the wage expression, even though the wage bargain is forward looking. This is due to the choice of the productivity process.

3.1.5 Acceptance Sets

Equation (9) defines the worker surplus in terms of the worker acceptance set $A(x)$ as well as the surplus from other jobs, $S(x,y')$. My goal is to obtain an expression for the acceptance sets. To this end, I eliminate $S(x,y')$ from the right hand side of (9) by substitution:

Fact 2. *The worker surplus can be written as*

$$S(x,y,A) = \theta\eta \left[F(x,y) - \beta\theta \frac{\int_{A(x)} [\eta f(y') - (1-s) p h(y')] F(x,y') dy'}{1 + \theta\beta \int_{A(x)} [\eta f(y') - (1-s) p h(y')] dy'} \right] \tag{10}$$

which is a continuous function of (x, α, A).

Proof. See the Appendix. □

Here I write S as a function of A to emphasize its dependence on the worker's acceptance decisions. At this stage I make no assumptions on A (such as optimality). Equation (10) has an intuitive interpretation. The worker's surplus is proportional to the term in square brackets. The first part, $F(x,y)$, is the productivity of the match. Obviously, surplus should be increasing in production, all else held constant. The second term captures both the probability of productivity shocks and worker's opportunity cost of matching with other firms. Consider what happens when we eliminate productivity shocks by setting $p = 0$:

$$S(x,y,A) = \theta\eta \left[F(x,y) - \frac{\beta\theta\eta \int_{A(x)} f(y') F(x,y') dy'}{1 + \theta\beta\eta \int_{A(x)} f(y') dy'} \right].$$

The final term inside the brackets is a biased weighted average of net production in alternative jobs, where the weights are the meeting probabilities $f(y')$. Without productivity shocks, the ratio is the worker's opportunity cost: matching with a type y firm today precludes matching with another firm in $A(x)$ until the job ends exogenously. An increase in $F(x,y')$ increases the opportunity cost of matching with y, and naturally decreases $S(x,y,A)$.

Equation (10) allows me to begin characterizing the acceptance sets. First, note that current firm productivity only enters this expression through the initial $F(x,y)$ term. $F_2 > 0$ then implies $S_2 > 0$, so worker surplus is strictly increasing in firm type, regardless of worker type x or their choice of acceptance set. This is a strong result, summarized in the following proposition.

Proposition 1. *Under the assumed productivity process all workers rank firms the same: $S(x,y,A)$ is monotone increasing in y for all x and A. All workers prefer high y firms to low y firms.*

Increasing surplus implies that equilibrium acceptance sets will be intervals defined by a lower threshold:

$$A(x) = \{y | y > y(x)\} \tag{11}$$

where $y(x)$ is the threshold firm type for worker x. Knowing that the $A(x)$ are intervals, with all workers preferring high y to low x firms, simplifies analysis considerably. Instead of considering acceptance sets which are arbitrary subsets of $[0,1]$, I now focus on finding the thresholds $y(x)$. This result is a consequence of allowing firms to employ multiple workers and post multiple vacancies, along with linear production as in Bagger and Lentz (2012). When a vacancy matches with a worker, the firm can post another vacancy the next period at the same cost. Thus the firm faces no opportunity cost of matching. Match surplus is production less the opportunity costs of the worker and firm, so when firms have no opportunity cost (or more generally, no variation in opportunity cost) it is natural that workers prefer higher type firms. In the models of Shimer and Smith (2000), Lopes de Melo (2013), and Hagedorn et al. (2012) firms and workers are treated *symmetrically*, in that matching with a partner precludes any other productive activity, at least until the match exogenously dissolves. The result is that firms, like workers, vary in their opportunity cost of matching, and this two-sided selectivity leads to potentially complicated acceptance patterns.[19]

So far I have treated the acceptance set $A(x)$, or equivalently $y(x)$, as exogenous. Now I solve for the worker's optimal $y(x)$. Note that in equilibrium a worker will accept a job if and only if it yields positive surplus, given their acceptance set: $S(x,y,A) > 0$. Thus the optimal threshold $y(x)$ must satisfy

$$S(x, y(x), A(x)) = 0 \ if \ y(x) > 0 \tag{12}$$

$$S(x, 0, [0,1]) > 0 \ if \ y(x) = 0 \tag{13}$$

The first case is of primary interest. If the threshold $y(x)$ is on the interior, then the surplus at the marginal firm must equal zero. Otherwise the worker would do strictly better by lowering his threshold and accepting more offers. It is possible that all jobs offer positive surplus. Then we are in the second case, where $A(x) = [0,1]$, i.e. the worker accepts all jobs. Then it must be the case that the lowest ranked job still provides surplus. Substituting (10) into (12) and (13), imposing (11) and rearranging yields

$$F(x, y(x)) + \theta\beta(1-s)p\int_{y(x)}^{1} h(y)(F(x,y) - F(x, y(x)))\,dy = \beta\theta\eta\int_{y(x)}^{1} f(y)(F(x,y) - F(x, y(x)))\,dy \quad \text{if} \quad y(x) > 0 \tag{14}$$

[19] My assumption is potentially more restrictive, but allows for tractable analysis. It is also motivated by the fact that in reality firms are large. The vast majority of workers are employed at multi-worker firm, not single worker firms.

$$F(x,y(x)) + \theta\beta(1-s)p\int_{y(x)}^{1} h(y)(F(x,y) - F(x,y(x)))\,dy \geq$$
$$\beta\theta\eta\int_{y(x)}^{1} f(y)(F(x,y) - F(x,y(x)))\,dy \quad \text{if} \quad y(x) = 0 \quad (15)$$

Thus equations (14) and (15) are necessary conditions that $y(x)$ must satisfy. Intuitively, the left hand side of (14) is the value of matching with a firm of type $y(x)$ immediately, and the right hand side is the value of waiting for another acceptable offer. Matching with $y(x)$ yields utility proportional to $F(x,y(x))$ for some duration. Once matched, there is also the possibility of a productivity shock, captured by the second term on the left hand side. Roughly speaking, an increase in firm type to $y > y(x)$ increases the utility of the worker by $F(x,y) - F(x,y(x))$. On the right hand side, the benefit of waiting is the discounted expected difference between production at $y(x)$ and production at an alternative: $\beta\theta\eta\int_{y(x)}^{1} f(y)(F(x,y) - F(x,y(x)))\,dy$.

As stated above, (14) and (15) are *necessary* conditions that an optimal $y(x)$ must satisfy. It turns out that they are sufficient conditions as well. This is because the left hand side of (14) is increasing in $y(x)$, while the right hand side is decreasing. This means that at most one point satisfies the system. The following proposition formalizes the argument.

Proposition 2. *The worker's optimal threshold $y(x)$ is the unique solution to equations (14) and (15).*

Proof. See the Appendix. □

Discussion

Consider the worker's acceptance choice in equation (14) when there are no productivity shocks ($p = 0$):

$$F(x,y(x)) = \beta\theta\eta\int_{y(x)}^{1} f(y)(F(x,y) - F(x,y(x)))\,dy. \quad (16)$$

This condition is formally identical to the reservation wage expression for a model of sequential search with an exogenous wage offer distribution. These models were first studied by McCall (1970), Gronau (1971) and Mortensen (1970). The version I use can be found in Mortensen (1986) and Rogerson et al. (2005). In this model there is a single worker type and an exogenous wage offer distribution. Each period an unemployed worker searches for a job, drawing wage offers from the distribution $\bar{f}(w)$. Jobs end with probability s each period. Then the values of unemployment U and employment $W(w)$ are

$$\begin{aligned} U &= \beta\int_{w^*}^{\bar{w}} W(w)\bar{f}(w)dw + \beta\left(1 - \int_{w^*}^{\bar{w}} \bar{f}(w)dw\right)U \\ W &= w + \beta(1-s)W + \beta sU \end{aligned}$$

where w^* is the worker's choice of reservation wage and $\bar{w}$ is the upper support of the offer distribution. This leads to the following expression for the reservation wage:[20]

$$w^* = \frac{\beta}{1-\beta} \int_{w^*}^{\bar{w}} \bar{f}(w)\,[w - w^*]\,dw. \tag{17}$$

Comparing this with equation (16) we see that the two decision rules have the same form, with the reservation wage w^* replacing the "reservation match quality" $F(x, y(x))$ in equation (17). It is interesting that the two decision rules should be so similar, given that wages are endogenous in my model. The similarity is a result of Nash bargaining and allowing firms to post multiple vacancies. Nash bargaining ensures that the wage is an additive function of a worker effect and a firm effect, as we can see in equation (7). Vacancy posting along with linear production implies that firms do not have an opportunity cost of hiring, so the firm effect in the wage equation is proportional to $F(x, y)$. If there were an opportunity cost, as in Shimer and Smith (2000), the firm effect would be $F(x, y)$ less the value of the firm's vacancy. Thus, my model can be interpreted as a general equilibrium extension to the McCall (1970) model, which retains the same simple decision rules as the partial equilibrium model. The advantage of the current model is that it allows for comparisons between different worker types, which will be taken up in Section (3.3).

3.1.6 Vacancy Posting

The results in this section and the next are not strictly necessary for understanding the assortative matching results in Section 3.3. The interested reader may skip directly to that part of the paper.

The firm posts vacancies at a convex cost until the expected value of an additional vacancy is zero. Formally, this means

$$c'(v) = \int_0^1 q(x) J(x, y) dx.$$

This is derived explicitly from the firm's problem in Section Appendix C. The left hand side is the marginal cost of a vacancy, and right hand side is the expected marginal value of the vacancy. $J(x, y)$ is the value to the firm of meeting a type x worker, and $q(x)$ is the probability the vacancy attracts such a worker within the period. The right hand side simply integrates this value across worker type. Using the Nash bargaining solution this marginal condition can be written

$$c'(v) = \frac{(1-\eta)}{\eta} \int_0^1 q(x) S(x, y) dx$$

Recalling that $c(v)$ is strictly convex, I can invert $c'(v)$ and arrive at an explicit expression for v:

$$v(y) = c'^{-1}\left(\frac{(1-\eta)}{\eta} \int_0^1 q(x) S(x, y) dx\right). \tag{18}$$

[20] See Rogerson et al. (2005) and Mortensen (1986) for explicit derivations and discussion.

Recall from Section 3.1.5 that $S(x,y)$ is increasing in y. This means that a higher type firm has a higher return to matching with a given worker. Then equation (18) implies that better firms post more vacancies: $v'(y) > 0$. This is a natural prediction, and turns out to be useful in Section Appendix D for characterizing assortative matching under productivity shocks.

3.1.7 Meeting and Matching

To close the model, I formalize the relationship between vacancies, unemployment, and the meeting rates $f(y)$ and $q(x)$. I make the standard "balanced meeting" assumption for random search models. The sorting results in Section 3.3 do not depend on the details of the matching technology, as long as search is random. Let $u(x)$ be the mass of unemployed type x workers. Then total unemployment is $U = \int u(x)dx$, and I can define $\tilde{u}(x) = \frac{u(x)}{U}$ as the density of unemployed workers. Let $v(y)$ be the vacancy posting of a type y firm, and recall that $h(y)$ is the mass of type y firms in the economy. Then total vacancy posting is $V = \int v(y)h(y)dy$ and I can define $\tilde{v}(y) = \frac{v(y)h(y)}{V}$ as the share of vacancies that are posted by type y firms.

Following the literature, I assume that the total number of meetings between workers and firms in a period is given by $M(U,V)$, an increasing, concave, linearly homogenous meeting function. The meeting function takes total unemployment and total vacancies as inputs, and gives the total number of contacts between workers and firms.

I assume that each worker contacts at most one vacancy, and each vacancy contacts at most one worker. This implies

$$M(U,V) \leq \min\{U,V\}.$$

Contacts are made in a random, fashion. Thus each worker has an equal probability of making a contact, which is given by $\frac{M(U,V)}{U}$. Likewise, each vacancy has a probability $\frac{M(U,V)}{V}$ of making a contact. Conditional on a worker making a contact, the vacancy is drawn at random from the population of vacancies. Thus, the arrival rate of type y vacancies for a searching worker is given by

$$f(y) = \frac{M(U,V)}{U}\tilde{v}(y). \tag{19}$$

Symmetrically for vacancies I have

$$q(x) = \frac{M(U,V)}{V}\tilde{u}(x) \tag{20}$$

where $q(x)$ is the probability a vacancy contacts a type x worker.

3.2 Equilibrium

In a steady state equilibrium the mass of unemployed workers of each type must remain constant over time. This requires that the outflows from unemployment

equal the inflows:

$$u(x)\int_{y(x)}^{1} f(y)dy = (l(x)-u(x))\left(s+(1-s)p\int_{0}^{y(x)} h(y)dx\right) \quad \forall x.$$

For a given x, the left hand side is the stock of unemployed times the rate at which they meet acceptable jobs. The right hand side is the mass of employed workers times the combined rate of endogenous and exogenous separations. This can be rearranged to yield

$$u(x) = \frac{\left(s+(1-s)p\int_0^{y(x)} h(y)dx\right) l(x)}{\left(s+(1-s)p\int_0^{y(x)} h(y)dx\right) + \int_{y(x)}^{1} f(y)dy} \quad \forall x \tag{21}$$

which makes clear that $u(x)$ is a continuous function of $f(y)$ and $y(x)$.

In addition to balancing worker flows, a steady state general equilibrium requires that (1) worker acceptance choices are made optimally, (2) vacancy posting decisions are made optimally and (3) meeting rates are consistent with the meeting function and the stocks of unemployed and vacancies. The exogenous parameters are the scalars β, s, p, η and the functions $l(x)$, $h(y)$ and $F(x,y)$. The endogenous objects are $y(x), v(y), f(y), q(x)$ and $u(x)$, which are jointly determined by equations (14), (15), (18), (19), (20), and (21). I reproduce the equilibrium conditions below, with intermediate notation eliminated.

First is the worker's acceptance set choice, for all x:

$$\begin{aligned} F(x,y(x)) + \theta\beta(1-s)p\int_{y(x)}^{1} h(y)\left(F(x,y)-F(x,y(x))\right)dy = \\ \beta\theta\eta\int_{y(x)}^{1} f(y)\left(F(x,y)-F(x,y(x))\right)dy \quad \text{if} \quad y(x)>0 \end{aligned} \tag{22}$$

$$\begin{aligned} F(x,y(x)) + \theta\beta(1-s)p\int_{y(x)}^{1} h(y)\left(F(x,y)-F(x,y(x))\right)dy \geq \\ \beta\theta\eta\int_{y(x)}^{1} f(y)\left(F(x,y)-F(x,y(x))\right)dy \quad \text{if} \quad y(x)=0. \end{aligned} \tag{23}$$

Equations (22) and (23) define $y(x)$ as a function of other model objects. Some additional algebra shows that $y(x)$ is a *continuous* function of $f(y)$, a fact which will be useful in showing existence of an equilibrium. Next is the firm's vacancy decision:

$$v(y) = c'^{-1}\left((1-\eta)\theta\int_0^1 q(x)\left[F(x,y) - \beta\theta\frac{\int_{y(x)}^{1}\left[\eta f(y')-(1-s)ph(y')\right]F(x,y')dy'}{1+\theta\beta\int_{y(x)}^{1}\left[\eta f(y')-(1-s)ph(y')\right]dy'}\right]dx\right) \quad \forall y. \tag{24}$$

Note that equation (24) defines $v(y)$ as a continuous function of $q(x)$, $f(y)$ and $y(x)$. Finally, recall the meeting rates

$$f(y) = \frac{M\left(\int_0^1 u(x)dx, \int_0^1 v(y)h(y)dy\right)}{\int_0^1 u(x)dx} \frac{v(y)h(y)}{\int_0^1 v(y)h(y)dy} \quad \forall y \tag{25}$$

$$q(x) = \frac{M\left(\int_0^1 u(x)dx, \int_0^1 v(y)h(y)dy\right)}{\int_0^1 v(y)h(y)dy} \frac{u(x)}{\int_0^1 u(x)dx} \quad \forall x. \tag{26}$$

Equations (25) and (26) define $f(y)$ and $q(x)$ as continuous functions of $u(x)$ and $v(y)$.

Expressions (21), (22), (23), (24), (25) and (26) define a mapping of the endogenous objects $\{y(x), v(y), f(y), q(x), u(x)\}$ into new values $\{y'(x), v'(y), f'(y), q'(x), u'(x)\}$. Under some regularity conditions, the continuity of the mapping ensures that a fixed point theorem can be applied, so an equilibrium exists. In matching models the sticking point in proving equilibrium existence is usually showing that the steady state flows equation defines a continuous mapping from acceptance sets to the masses of searchers (see Shimer and Smith (2000), Nöldeke and Tröger (2009) and Smith (2011)). In my setup this is not an issue, because (21) and (24) explicitly solve for the masses of unemployed and vacancies as a function of other variables. In general the model may have multiple equilibria. Unfortunately, this is a common feature of models with two-sided heterogeneity and search.[21]

3.3 Assortative Matching

In this section I analyze the sorting of workers across firms. I focus on deriving sufficient conditions for positive and negative assortative matching. Intuitively, positive assortative matching means that better workers tend to match with better firms. In the context of my model, positive assortative matching is equivalent to $y'(x) \geq 0 \ \forall x$. That is, better workers (higher x) set their threshold firm type $y(x)$ higher. Similarly, negative assortative matching is equivalent to $y'(x) \leq 0 \ \forall x$. The case of no sorting is also of interest, and in my model this translates to $y'(x) = 0 \ \forall x$.

The empirical pattern in Section 2 relates firm growth to worker type. In this section I derive a condition, strict log-supermodularity, under which better firm types hire better worker types. In Section 3.4 I show that under strict log-supermodularity better firms are also faster growing. Together, the firm type-worker type and firm type-growth relations imply that firm growth and quality of hires are positively correlated, reproducing the empirical results.

In what follows I use the worker's threshold condition to derive an expression for $y'(x)$. The sign of $y'(x)$, and thus the nature of sorting, depend on the sign of a simple expression that compares match production at the marginal firm to expected production at firms in the acceptance set. The expression can be unambiguously signed when production $F(x, y)$ is strictly log-supermodular or strictly log-submodular, as described below. These are sufficient conditions for positive

[21] See Sattinger (1995), Shimer and Smith (2000) and Lopes de Melo (2013).

and negative assortative matching, respectively. Interestingly, these sufficient conditions rely only on the production function, and not the distribution of agent types or any other parameter. In this sense they are similar to the sorting conditions derived by Becker (1973), Shimer and Smith (2000), Lentz (2010) and Eeckhout and Kircher (2010) in the context of other models. Given that the sorting conditions are independent of agent distributions and other parameters, it might be asked whether weaker sufficient conditions can be found. I show that any weaker sufficient conditions must depend on the distribution of types. Specifically, if the production function is not strictly log-supermodular or submodular one can find distributions of agents that lead to non-assortative matching.

For simplicity I focus on the case of no productivity shocks: $p = 0$. All of the intuition and results carry over to the case in which productivity shocks occur, but at the cost of additional notation and technicalities. I prove the main results for the general model in Appendix Appendix D.

Consider the worker's threshold condition (14) when there are no productivity shocks:

$$F(x,y(x))\left(1+\beta\theta\eta\int_{y(x)}^{1} f(y)dy\right)-\beta\theta\eta\int_{y(x)}^{1} f(y)F(x,y)\,dy \;=\; 0 \tag{27}$$

This expression implicitly defines $y(x)$ in terms of x, $F(x,y)$, and the other parameter values. As a consequence (27) can also be used to derive an expression for $y'(x)$. While this expression is cumbersome, it shows that the *sign* of $y'(x)$, and thus the nature of sorting, depend only on a simple condition, shown below.

Proposition 3. *When the worker's choice of $y(x)$ is given by (27), $y'(x) > 0$ if and only if*

$$\int_{y(x)}^{1} w(y)\frac{F_1(x,y)}{F(x,y)}dy > \frac{F_1(x,y(x))}{F(x,y(x))} \tag{28}$$

and $y'(x) < 0$ if and only if

$$\int_{y(x)}^{1} w(y)\frac{F_1(x,y)}{F(x,y)}dy < \frac{F_1(x,y(x))}{F(x,y(x))}, \tag{29}$$

where $w(y) = \frac{f(y)F(x,y)}{\int_{y(x)}^{1} f(y)F(x,y)dy} > 0$, and $\int_{y(x)} w(y) = 1$.

Proof of Proposition 3. Equation (27) implicitly defines $y(x)$, so differentiating it I can obtain an expression for $y'(x)$. Applying the implicit function theorem I have

$$y'(x) = -\frac{F_1(x,y(x))\left(1+\beta\theta\eta\int_{y(x)}^{1} f(y)dy\right)-\beta\theta\eta\int_{y(x)}^{1} f(y)F_1(x,y)\,dy}{F_2(x,y(x))\left(1+\beta\theta\eta\int_{y(x)}^{1} f(y)dy\right)} \tag{30}$$

The denominator is always positive, so the sign of $y'(x)$ depends only on the numerator, implying $y'(x) > 0$ if and only if

$$\beta\theta\eta\int_{y(x)}^{1} f(y)F_1(x,y)\,dy > F_1(x,y(x))\left(1+\beta\theta\eta\int_{y(x)}^{1} f(y)dy\right).$$

Dividing this expression by (27) I arrive at

$$\frac{\int_{y(x)}^{1} f(y) F_1(x,y)\, dy}{\int_{y(x)}^{1} f(y) F(x,y)\, dy} > \frac{F_1(x, y(x))}{F(x, y(x))}$$

which can be rearranged to yield equation (28) as desired. A symmetric argument shows that equation (29) follows as well. □

In equation (28) $\frac{F_1(x,y(x))}{F(x,y(x))}$ reflects the marginal change in output at the marginal firm type when we increase the worker type. The term $\int_{y(x)}^{1} w(y) \frac{F_1(x,y)}{F(x,y)} dy$ reflects the expected change in productivity at firms in the acceptance set when we increase worker type. If output at the the marginal firm increases more than at the average accepted firm, the relative value of waiting for better offers is lower, and the worker lowers the acceptance threshold $y(x)$. Alternatively, if expected productivity at accepted firms increases more than productivity at the marginal firm, the worker prefers to wait for better offers and raises the acceptance threshold. In this case we have (locally) positive assortative matching, and $y'(x) > 0$.

Having shown that satisfying equation (28) is equivalent to assortative matching, the question is under what conditions does (28) hold? Clearly, a sufficient condition is that $\frac{F_1(x,y)}{F(x,y)}$ be increasing in y. Then any weighted average $\int_{y(x)}^{1} w(y) \frac{F_1(x,y)}{F(x,y)} dy$ must exceed the value $\frac{F_1(x,y(x))}{F(x,y(x))}$. The economic interpretation of this condition is straightforward. When $\frac{F_1(x,y)}{F(x,y)}$ is increasing in y, it means that output is *more* sensitive to worker type at high y firms. Increasing worker type increases output the most at the highest y firms, encouraging the worker to raise the acceptance threshold and generating positive assortative matching.

The condition $\frac{\partial}{\partial y} \frac{F_1(x,y)}{F(x,y)} > 0$ is equivalent to strict log-supermodularity. In particular, a differentiable, positive, increasing function is strictly log-supermodular if and only if its logarithm is supermodular:

$$\frac{\partial^2}{\partial x \partial y} \ln F(x,y) > 0$$

which is equivalent to

$$\frac{\partial}{\partial y} \frac{F_1(x,y)}{F(x,y)} > 0$$

as asserted. Returning to equation (29), it is clear that $\frac{\partial}{\partial y} \frac{F_1(x,y)}{F(x,y)} < 0$ ensures $y'(x) < 0$, so matching is negative assortative. In turn $\frac{\partial}{\partial y} \frac{F_1(x,y)}{F(x,y)} < 0$ is equivalent to strict log-*sub*modularity:

$$\frac{\partial^2}{\partial x \partial y} \ln F(x,y) < 0.$$

Finally, equations (28) and (29) imply there is no sorting, $y'(x) = 0$, if and only if $\frac{\partial}{\partial y} \frac{F_1(x,y)}{F(x,y)} = 0$. These results are summarized in the following Proposition.

Proposition 4. *Consider the economy with no productivity shocks. Matching is positive assortative, $y'(x) > 0$, if F is strictly log-supermodular. Matching is negative assortative, $y'(x) < 0$ if F is strictly log-submodular. There is no sorting if $\frac{\partial}{\partial y}\frac{F_1(x,y)}{F(x,y)} = 0$.*

Proposition 4 provides sufficient conditions for positive and negative assortative matching. Note that the conditions depend only on the production function F. They are independent of the distribution of agent types as well as the other parameters of the economy. A natural question is whether these conditions are the minimal ones that ensure assortative matching, or whether weaker conditions may suffice. The following Proposition shows that strict log-supermodularity is necessary for positive assortative matching, in the sense that if F is not globally strictly log-supermodular we can find a distribution of vacancies that ensures $y'(x) < 0$ somewhere.

Proposition 5. *If there exists y^* and x^* such that $\frac{\partial^2}{\partial x \partial y} \ln F(x^*, y^*) < 0$, then there exists a set of meeting rates $f(y)$ such that $y'(x^*) < 0$.*

Proof. See the Appendix. □

The strategy of proof is to adjust the meeting rates $f(y)$ to keep the worker's threshold constant at $y^* = y(x)$, but make sure sorting is (locally) negative. Keeping in mind that $\frac{\partial^2}{\partial x \partial y} \ln F(x^*, y^*) < 0$, there is some $y^{**} > y^*$ such that $\frac{F_1(x^*,y^*)}{F(x^*,y^*)} > \frac{F_1(x^*,y^{**})}{F(x^*,y^{**})}$. If $f(y^{**})$ is large enough, the worker will expect to meet type y^{**} firms much more often than any alternative. The increased weight on y^{**} means that the behavior of F around(x^*, y^{**}) dominates sorting decisions. Appealing to Proposition 3 then shows that $y'(x^*) < 0$, as desired.

Similar to other assortative matching results, my conditions on the production function are sufficient for *any* values of the other parameters and distributions, and are the weakest conditions on the production function that ensure assortativity for all other parameter values. If we consider a restricted set of type distributions, or other parameters, it may be possible to find alternative necessary and sufficient conditions. This is left for future work.

3.3.1 Discussion

This paper is one of many to study the theory of sorting in a search environment.[22] The two papers most similar to this one are Shimer and Smith (2000) and Eeckhout and Kircher (2010). While neither allow for productivity shocks, the models can still be compared. In the random search model of Shimer and Smith (2000), positive assortative matching requires supermodularity not only of $\ln F$ but also of $\ln F_1$, $\ln F_2$ and $\ln F_{12}$.[23] Thus my log-supermodularity requirement is weaker than theirs. This can be attributed to my allowing multiple worker firms. In

[22] See Shimer and Smith (2000), Shi (2001), Shimer (2005), Atakan (2006), Eeckhout and Kircher (2010), Lentz (2010) and Eeckhout and Kircher (2012).

[23] While Shimer and Smith (2000) do not discuss supermodularity of $\ln F$, Eeckhout and Kircher (2010) show that their conditions, along with monotonicity, imply log-supermodularity.

Shimer and Smith (2000) firms have an opportunity cost of matching, which is increasing in firm type. This tends to make high type firms less attractive to workers, since the higher firm opportunity cost comes out of the worker's wages. In my model the workers get a constant share of match output (plus a function of the worker opportunity cost). This makes high type firms more attractive, and encourages high type workers to wait for high type firms.

Eeckhout and Kircher (2010) work in a directed search framework. They derive *root-n supermodularity* as a sufficient condition for positive assortative matching. This condition requires that the n'th root of the production function be positive, where n is the elasticity of substitution of the matching function. They show that root-n supermodularity is always weaker than log-supermodularity, and thus weaker than the requirements of Shimer and Smith (2000). This is due to the differences between directed and random search. See Eeckhout and Kircher (2010) for an illuminating discussion of these differences, and the relation of their result to that of Shimer and Smith (2000). My condition, while weaker than Shimer and Smith's, is still stronger than root-n supermodularity. This is because my model is one of random search, while Eeckhout and Kircher (2010) study directed search.

3.4 Firm Growth and the Quality of Hires

In Section 3.3 I showed that strict log-supermodularity ensures assortative matching between worker types and firm types: high type firms hire better workers. To link these results to the empirical results in Section **2,** I need a connection between firm type and firm *growth.* Intuitively, if high type firms tend to grow faster, then strictly log-supermodular match production will lead to a positive firm growth-quality of hires correlation, as seen in the data. On the other hand, if there is a weak or ambiguous correlation between firm type and firm growth then the assortative matching condition in Section 3.3 will be irrelevant for understanding the firm growth-quality of hires correlation.

It is natural to think that high type firms will tend to be faster growing: the results in Sections 3.1.6 and 3.3 show that under strict log-supermodularity better firms both post more vacancies and hire more workers per vacancy. In practice, the relationship between firm growth and firm quality is complicated. Higher type firms tend to post more vacancies and hire more workers. Ceteris paribus, this increases growth rates. However, a high type firm will also tend to be larger to begin with and have larger outflows of workers, both of which reduce growth rates. Sorting out these effects is complicated by the heterogeneity of the workforce, which causes firms to occasionally hire a substantial number of workers while simultaneously separating a large portion of their incumbent workers. This can happen in the model when a there is a negative productivity shock.

To show that firm growth is positively correlated with firm type (and therefore worker type), it suffices to show that a stochastic dominance property holds. Let g be the (random) growth rate of a firm, and let $F_g(g|y)$ be the CDF of growth rates conditional on the firm being of type y. Assume for the moment that if $y' > y$, then $F_g(\cdot|y')$ dominates $F_g(\cdot|y)$ (in the sense of first order stochastic dominance). Such a dominance property implies that firm type and firm growth have a positive

correlation coefficient, which is the desired result. So the main problem is to prove dominance.

Proposition 6. *Assume that matching is positive assortative. Then $F_g(\cdot|y')$ dominates $F_g(\cdot|y)$ whenever $y' > y$.*

Proof. See the Appendix. □

The argument has three parts

- **Step 1**: Show that firm growth $g(y,m)$ is an exact function of current productivity y and a set of state variables m, where m includes the time elapsed since the last shock and the employment distribution at the time of the shock. In addition, $h(y|m) = h(y)$.
- **Step 2**: Show that $g_1(y,m)$ is positive.
- **Step 3**: Use the population distribution of m and y to show that $g_1(y,m) > 0$ is a sufficient condition for $F_g(\cdot|y')$ to dominate $F_g(\cdot|y)$ whenever $y' > y$.

The specification of the productivity process helps in Step 1. Step 2 follows in part from Proposition 7 (see below), which helps pin down the compositional dynamics of employment by worker type.

An alternative approach would be to work with the distribution of firm types conditional on growth. This might appear more natural, since the empirical results are stated in terms of worker types conditional on firm growth. However, it turns out there is not necessarily any stochastic dominance between the type distributions conditional on growth. This necessitates working with growth conditional on firm type to prove dominance and therefore a positive correlation between firm growth and the quality of hires.

3.4.1 Discussion

I have shown that the model implies a positive correlation between firm growth and quality of hires if production is strictly log-supermodular. The same condition implies positive assortative matching, so the pattern documented in Section 2 can be taken as indirect evidence for assortative matching between workers and firms. This paper is the first to characterize matching analytically in an environment with firm growth and productivity shocks. One difficulty is that in such a rich environment the connection between firm growth and productivity can become unclear. Proposition 6 proves that a positive connection between firm growth and productivity does indeed hold within the model.

As noted earlier, Lopes de Melo (2013), Eeckhout and Kircher (2011) and others have shown that AKM fixed effects are inconsistent when the data generating process is a model along the lines of Shimer and Smith (2000). Their critique does not apply to my model. Allowing firms to post multiple vacancies ensures there is no opportunity cost to hiring, so the wage is monotone increasing in firm type. Then fixed effects estimation should recover the firm types accurately. If my model

is correct there is a puzzle: why do I find assortative matching in Section 2, when AKM and others find none? As I have discussed, one possibility is that assortative matching obtains only within narrow labor markets, while sorting across the whole economy is more ambiguous. Another possibility is that my results are based on firm *growth* while the fixed effects results assume a constant firm effect over the sample period. If productivity shocks are not too persistent, so that time-averaged productivity is close to the unconditional expectation of productivity for most firms, then we would expect fixed effects regressions to reveal relatively little sorting. The reason is that while there is sorting between worker and firm types, the firm fixed effect is an average of firm productivity over the sample. This suggests an alternative identification strategy, which would jointly estimate worker effects and firm-year effects as in Mortensen et al. (2010). Allowing the firm effect to vary over time would get rid of the averaging problem.

3.5 Additional Results

Besides addressing the empirical correlation between firm growth and the quality of hires, the model can be used to derive additional results. Here I discuss three broad sets of results.

3.5.1 Vacancy Filling Rates

The model can be used to study the relationship between firm growth and how quickly vacancies fill. Recall from Proposition 1 that each worker prefers high type firms to low type firms. This implies that vacancies at high type firms fill more quickly, since a larger share of workers are willing to match with high type firms. This, along with Proposition 6, means that within the model faster growing firms also fill their vacancies more quickly.

This fact relates to the recent work of Davis et al. (2013). They find empirically that fast growing establishments tend to fill their job openings more quickly. This is at odds with the textbook search theory, which assumes the vacancy filling rate is exogenous to the firm and consequently employment growth is linear in the vacancy rate at the firm level. The conclusion of Davis et al. (2013) is that fast growing firms are using other margins to attract workers, besides simply posting more vacancies. My model and the empirical results point to a particular channel. High growth firms are more productive, and are able to hire selective high type workers. The increased wages associated with a positive productivity shock allow the firm to attract workers that would otherwise reject an offer from the firm. In this way the firm hires more of the workers it meets, reducing vacancy durations.

3.5.2 Compositional Dynamics

Conditional on positive assortative matching, we can say something about the dynamics of worker composition at the firm.

Proposition 7. *Assume that matching is positive assortative: $y'(x) > 0$. At a given firm, the growth rate of employment of type x workers*

$$g_t(x) = \frac{N_t(x) - N_{t-1}(x)}{0.5(N_t(x) + N_{t-1}(x))} \tag{31}$$

is weakly increasing in x, as long as $N_t(x) > 0$.

Proof. See the Appendix. □

This result can be understood as follows. Fix a worker type x. When y falls below $y(x)$, a firm cannot hire any type x workers, *and all incumbent type x workers leave.* If y later rises above $y(x)$ the firm begins hiring type x workers again, and the growth rate is high because the initial stock is low. On the other hand, if y has been greater than $y(x)$ for a long time there will be a larger stock of type x workers at the firm. This implies the growth rate of type x workers will be lower. Better workers have higher thresholds $y(x)$, so the firm is more likely to have dipped below the threshold more recently.

These dynamics lead to another interesting result:

Corollary 1. *The quality of new hires at a firm is always (weakly) better than the quality of incumbents, in the sense of first order stochastic dominance.*

Proof. Follows immediately from Proposition 7. □

Since better workers have faster growth rates at a firm, it follows that the quality of new hires dominates the quality of incumbents. This holds regardless of whether the firm is growing or shrinking, since a negative shock leads the best workers to leave. So although the quality of new hires may be lower than that of incumbents one period ago, the quality is higher than those who remain. Consideration of separations leads to another result.

Proposition 8. *The quality of separating workers is always (weakly) better than the quality of incumbents, in the sense of first order stochastic dominance.*

Proof of Proposition 8. Separations are a mix of exogenous separations, at rate s, and endogenous separations in response to productivity shocks. Exogenous separation rates are constant across worker types, so if there is not a negative productivity shock the distribution of separations matches the distribution of incumbents. If there is a negative productivity shock then the best incumbents separate, and the distribution of separators will dominate the distribution of incumbents. □

Given that both hires and separators are better than incumbents, we might ask whether we can order hires and separators relative to each other. This turns out not to be possible. If there is a good productivity shock, hires will be better than separators, since all separations will be exogenous. If there is a negative shock the opposite will hold.

3.5.3 Dynamics and Selectivity

Section 3.3 showed that the sufficient conditions for assortative matching are independent of the existence or frequency of productivity shocks. In this section I study worker selectivity and its relation to productivity shocks, under the assumption of positive assortative matching. I show that reducing the persistence of shocks makes workers less selective. Formally, the level of $y(x)$ is decreasing in p.

Proposition 9. *Let $f(y)$ be fixed. Under positive assortative matching, a reduction in p reduces $y(x)$ for all x.*

This is a partial equilibrium result: it depends on $f(y)$ being held constant. In reality, different values of p would induce different unemployment rates and vacancy postings, changing $f(y)$.[24] Nonetheless, the result provides some strong intuition for how productivity shocks shape matching decisions and sorting. The effect of more frequent shocks is to make firms more similar to each other, since the expected duration of current productivity is shorter. Since firms are effectively more similar to each other, each worker is willing to accept more firm offers and $y(x)$ declines. As p increases, all workers move toward accepting all offers[25].

4 Conclusion

This paper contributes to the theory and empirics of worker-firm sorting. On the empirical side, I document a previously unknown relationship between firm growth and the quality of hires: when a given firm grows faster it hires more productive workers. This relationship survives corrections for selection bias and the endogeneity of wages. Under the assumption that faster growing firms are more productive this pattern implies positive assortative matching. This evidence of assortative matching stands in contrast to the mixed results in the literature following Abowd et al. (1999). An important caveat is that when I do not condition on firm characteristics, I do not find evidence of assortative matching. This mirrors the results in the AKM literature, which do not typically control for firm characteristics. I argue that firm characteristics proxy for the labor market a firm faces. While assortative matching may be the norm within a given industry or occupation, it need not characterize cross-market variation in firm productivity, worker productivity, and wages. Thus my baseline results suggest positive sorting within labor market segments, while my alternative specification (and perhaps the AKM regression) captures the more ambiguous relationships that hold between markets.

[24] There is a special case in which this result holds exactly in general equilibrium. In the special case of a quadratic meeting function $M = UV$, the values $f(y)$ depend only on the number of vacancies posted, not on the number of unemployed. If vacancy costs are highly convex, say $c(v) = \lim_{z\to\infty} v^z$ then vacancies per firm are constant, and $f(y)$ will be invariant to changes in p.

[25] In this discrete time model, workers may still not accept all offers when $p = 1$. This is because a very low value of y, even for one period, may be enough to deter matching. However, the same model can be set in continuous time with no change in sorting results. In continuous time, the arrival rate of shocks can be driven to infinity, and workers accept all jobs in the limit.

Motivated by the empirical results, I develop a tractable model of worker-firm sorting within a single market. The model's tractability allows me to analytically characterize sorting, as well as firm compositional dynamics and the effect of productivity shocks on sorting. The model's central equilibrium condition, the worker's first order condition, corresponds exactly to a model of sequential search with an exogenous wage distribution. I show that the model reproduces the positive correlation between firm growth and the quality of hires when worker and firm types are strong complements in production (i.e., the production function satisfies strict log-supermodularity). This result extends the results of Becker (1973) on complementarity and assortative matching to an environment with productivity shocks and firm growth.

There are many possibilities for future research. One obvious direction is to examine business cycle behavior. Since the Great Recession, there has been renewed interest in the composition of the unemployed and how hiring recovers after a downturn. Mueller (2012) shows that composition of the unemployed becomes *better* in recessions. Davis et al. (2013) and Davis et al. (2012) study the role of employer recruiting intensity in the Great Recession. These studies point to a role for worker and firm heterogeneity in understanding aggregate dynamics.

References

Abowd, John, Francis Kramarz, Sebastien Perez-Duarte, and Ian Schmutte, "A Formal Test of Assortative Matching in the Labor Market," Mimeo, Univesity of Georgia, March 2012.

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, March 1999, *67* (2), 251–334.

_ , Robert H. Creecy, and Francis Kramarz, "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," Mimeo, Cornell University 2002.

Andrews, M. J., L. Gill, T. Schank, and R. Upward, "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2008, *171* (3), 673–697.

Andrews, M.J., L. Gill, T. Schank, and R. Upward, "High Wage Workers Match with High Wage Firms: Clear Evidence of the Effects of Limited Mobility Bias," *Economics Letters*, 2012, *117* (3), 824 – 827.

Atakan, Alp E., "Assortative Matching with Explicit Search Costs," *Econometrica*, 05 2006, *74* (3), 667–680.

Bagger, Jesper and Rasmus Lentz, "An Empirical Model of Wage Dispersion with Sorting," 2009 Meeting Papers 964, Society for Economic Dynamics 2012.

Bartelsman, Eric J. and Mark Doms, "Understanding Productivity: Lessons from Longitudinal Microdata," *Journal of Economic Literature*, 2000, *38* (3), pp. 569–594.

Bartolucci, Cristian and Francesco Devicienti, "Better Workers Move to Better Firms: A Simple Test to Identify Sorting," Mimeo, Collegio Carlo Alberto, 2012.

Becker, Gary S, "A Theory of Marriage: Part I," *Journal of Political Economy*, July-Aug. 1973, *81* (4), 813–46.

Belzil, Christian, "Job Creation and Job Destruction, Worker Reallocation, and Wages," *Journal of Labor Economics*, April 2000, *18* (2), 183–203.

Bjelland, Melissa, Bruce Fallick, John Haltiwanger, and Erika McEntarfer, "Employer-to-Employer Flows in the United States: Estimates Using Linked Employer-Employee Data," *Journal of Business & Economic Statistics*, 2011, *29* (4), 493–505.

Brown, Charles and James Medoff, "The Employer Size-Wage Effect," *Journal of Political Economy*, October 1989, *97* (5), 1027–59.

Brown, Jennifer and David A. Matsa, "Boarding a Sinking Ship? An Investigation of Job Applications to Distressed Firms," Mimeo, Northwestern University July 2012.

Burdett, Kenneth and Dale T. Mortensen, "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 1998, *39* (2), pp. 257–273.

Coad, Alex, Sven-Olov Daunfeldt, Dan Johansson, and Karl Wennberg, "Who do High-growth Firms Employ, and Who do they Hire?," Mimeo, The Ratio Institute July 2011.

Dahl, Michael and Stevens Klepper, "Who Do New Firm Hire?," Mimeo, Aalborg University 2007.

Davis, Steven J., John C. Haltiwanger, and Scott Schuh, *Job Creation and Destruction*, Vol. 1 of *MIT Press Books*, The MIT Press, December 1998.

_ , R. Jason Faberman, and John C. Haltiwanger, "Recruiting Intensity During and After the Great Recession: National and Industry Evidence," *American Economic Review*, May 2012, *102* (3), 584–88.

_ , _ , and _ , "The Establishment-Level Behavior of Vacancies and Hiring," *The Quarterly Journal of Economics*, 2013, *128* (2).

Eeckhout, Jan and Philipp Kircher, "Sorting and Decentralized Price Competition," *Econometrica*, 2010, *78* (2), 539–574.

_ and _ , "Identifying Sorting In Theory," *The Review of Economic Studies*, 2011, *78* (3), 872–906.

_ and _ , "Assortative Matching with Large Firms: Span of Control over More versus Better Workers," Mimeo, University of Edinburgh, 2012.

Elsby, Michael W. L. and Ryan Michaels, "Marginal Jobs, Heterogeneous Firms, and Unemployment Flows," *American Economic Journal: Macroeconomics*, September 2013, *5* (1), 1–48.

Fallick, Bruce and Charles A. Fleischman, "Employer-to-Employer Flows in the U.S. Labor Market: the Complete Picture of Gross Worker Flows," Mimeo, Board of Governors of the Federal Reserve System, 2004.

Foster, Lucia, Cheryl Grim, and John Haltiwanger, "Reallocation In The Great Recession: Cleansing Or Not?," Mimeo, Center for Economic Studies, U.S. Census Bureau, August 2013.

_ , John C. Haltiwanger, and C. J. Krizan, "Aggregate Productivity Growth. Lessons from Microeconomic Evidence," in "New Developments in Productivity Analysis" NBER Chapters, National Bureau of Economic Research, Inc, 2001, pp. 303–372.

Gordanier, John, "The Division of Labor, Firm Size, and Employee Learning," Mimeo, Amherst College, 2012.

Gronau, Reuben, "Information and Frictional Unemployment," *The American Economic Review*, 1971, *61* (3), pp. 290–301.

Hagedorn, Marcus, Tzuo Hann Law, and Iourii Manovskii, "Identifying Equilibrium Models of Labor Market Sorting," Working Paper 18661, National Bureau of Economic Research December 2012.

Hogan, Vincent, "Wage aspirations and unemployment persistence," *Journal of Monetary Economics*, November 2004, *51* (8), 1623–1643.

Iranzo, Fabiano Schivardi Susana and Elisa Tosetti, "Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data," *Journal of Labor Economics*, 2008, *26* (2), pp. 247–285.

Kaas, Leo and Philipp Kircher, "Efficient Firm Dynamics in a Frictional Labor Market," IZA Discussion Papers 5452, Institute for the Study of Labor, January 2011.

Lazear, Edward P. and James R. Spletzer, "Firing and Hiring," Mimeo, 2013.

Lentz, Rasmus, "Sorting by search intensity," *Journal of Economic Theory*, 2010, *145* (4), 1436 – 1452.

Lise, Jeremy and Jean-Marc Robin, "The macro-dynamics of sorting between workers and firms," IFS Working Papers, Institute for Fiscal Studies, August 2013.

_ **, Costas Meghir, and Jean-Marc Robin**, "Mismatch, Sorting and Wage Dynamics," Working Paper 18719, National Bureau of Economic Research, January 2013.

Lopes de Melo, Rafael, "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence," Mimeo, 2013.

McCall, J. J., "Economics of Information and Job Search," *The Quarterly Journal of Economics*, 1970, *84* (1), pp. 113–126.

Mortensen, Dale, *Wage Dispersion: Why Are Similar Workers Paid Differently?* MIT Press Books, The MIT Press, 2003.

Mortensen, Dale T, "A theory of wage and employment dynamics," *eds.) ES Phelps et al., The Microeconomic Foundations of Employment and Inflation Theory, New York: Norton*, 1970.

Mortensen, Dale T., "Job search and labor market analysis," in O. Ashenfelter and R. Layard, eds., *Handbook of Labor Economics*, Vol. 2 of *Handbook of Labor Economics*, Elsevier, 1986, chapter 15, pp. 849–919.

_ **and Christopher A. Pissarides**, "Job Creation and Job Destruction in the Theory of Unemployment," *The Review of Economic Studies*, 1994, *61* (3), pp. 397–415.

_ **, Bent Jesper Christensen, and Jesper Bagger**, "Wage and Productivity Dispersion: Labor Quality or Rent Sharing?," Mimeo, 2010.

Moscarini, Giuseppe and Fabien Postel-Vinay, "Stochastic Search Equilibrium," *Review of Economic Studies*, 2013, *80* (4), 1545–1581.

Mueller, Andreas I., "Separations, Sorting and Cyclical Unemployment," IZA Discussion Papers 6849, Institute for the Study of Labor (IZA) September 2012.

Nöldeke, Georg and Thomas Tröger, "Matching Heterogeneous Agents with a Linear Search Technology," Bonn Econ Discussion Papers, University of Bonn, Germany January 2009.

Oi, Walter Y. and Todd L. Idson, "Firm size and wages," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3 of *Handbook of Labor Economics*, Elsevier, 1999, chapter 33, pp. 2165–2214.

Pissarides, Christopher A., *Equilibrium Unemployment Theory, 2nd Edition*, Vol. 1 of *MIT Press Books*, The MIT Press, June 2000.

Rogerson, Richard, Robert Shimer, and Randall Wright, "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, September 2005, *43* (4), 959–988.

Sattinger, Michael, "Search and the Efficient Assignment of Workers to Jobs," *International Economic Review*, May 1995, *36* (2), 283–302.

Shi, Shouyong, "Frictional Assignment. I. Efficiency," *Journal of Economic Theory*, 2001, *98* (2), 232 – 260.

Shimer, Robert, "The Assignment of Workers to Jobs in an Economy with Coordination Frictions," *The Journal of Political Economy*, 2005, *113* (5), pp. 996–1025.

_ **and Lones Smith**, "Assortative Matching and Search," *Econometrica*, 2000, *68* (2), 343–369.

Smith, Lones, "Frictional Matching Models," *Annual Review of Economics*, 2011, *3* (1), 319–338.

Woodcock, Simon D., "Wage differentials in the presence of unobserved worker, firm, and match heterogeneity," *Labour Economics*, 2008, *15* (4), 771 – 793. <ce:title>European Association of Labour Economists 19th annual conference / Firms and Employees</ce:title>.

A Data

In constructing worker and job flows I follow generally follow the definitions used in official LEHD data products.[26] The first task is defining firm employment. As noted in the main text, the total number of workers with earnings at firm j in quarter t does not necessarily reflect the firm size at any point in time. A firm with high turnover will have many workers over the course of a quarter, even though it only had a few employees at each point in time. To get around this issue I use beginning of quarter employment and end of quarter employment. Beginning of quarter employment in quarter t is the number of workers who reported positive earnings at firm in both $t-1$ and t. Intuitively, this means that the worker was at the firm on the first day of quarter t. It is possible that the worker separated during $t-1$ and was rehired sometime later in quarter t, but this is not a serious concern overall. Similarly, end of quarter employment is the number of workers with positive earnings at the firm in both quarters t and $t+1$. These workers are likely to have been employed on the last day of quarter. Thus beginning of quarter employment and end of quarter employment serve as two "point in time" estimates of firm size. Letting $bop_{j,t}$ and $eop_{j,t}$ denote the two point in time measures, I define firm employment growth using the conventional Davis et al. (1998) formula

$$g_{j,t} = \frac{eop_{j,t} - bop_{j,t}}{0.5\left(eop_{j,t} + bop_{j,t}\right)}.$$

A.1 Full Quarter Earnings

Recall my definition of $q_{i,t}^{FQ}$:

$$q_{i,t}^{FQ} = \frac{\sum_{k=1}^{8} earn_{i,t-k}^{FQ}}{\sum_{k=1}^{8} \mathbf{1}\left\{earn_{i,t-k}^{FQ} > 0\right\}}.$$

The intention here is to measure the worker's average earnings per quarter *while employed.* The difficulty is that when a worker is hired or separated within a quarter I do not observe how much of the quarter they were employed for. The solution is to look for cases where I can reasonably confident that the worker was employed the entire quarter. In particular, if a worker is observed at firm j in period $t-1$, t and ,$t+1$ it is likely they were employed at j continuously throughout period t. I follow this logic in calculating $earn_{i,t}^{FQ}$:

$$earn_{i,t}^{FQ} = earn_{i,t} \times \mathbf{1}\left\{earn_{i,t-1} > 0\right\} \times \mathbf{1}\left\{earn_{i,t} > 0\right\} \times \mathbf{1}\left\{earn_{i,t+1} > 0\right\}.$$

A.2 Hires

I emphasize the importance of focusing on hires from nonemployment. Here again I must take account of the quarterly nature of the data. If a worker is observed

[26] See http://lehd.ces.census.gov/doc/QWI_101.pdf and the related documentation for more information on public-use LEHD data.

to have worked at several firms in a quarter it is difficult to determine which jobs were concurrent, which jobs ended before others began, and whether there were any nonemployment spells in the quarter.

A hire occurs when a worker has positive earnings at firm j in quarter t, but zero earnings at j in quarter $t-1$. This implies that at some point in quarter t the worker was hired at j. If the worker has positive earnings *only* at j in quarter $t-1$ we may assume that they were nonemployed prior to being hired.[27] Thus, to measure hires from nonemployment (NE) I use only workers who were hired in t and had exactly one job in t.

Measuring job to job hires (EE) also requires some assumptions. I consider a hire to be an EE hire if the worker (1) was hired in t, (2) had another job that continued from $t-1$ into t, and (3) no jobs from $t-1$ continued into $t+1$. These restrictions allow for job overlap or periods of nonemployment during t, but do not allow the worker's old jobs (from $t-1$) to continue beyond the transition quarter. This measure doubtless includes some transitions that included nonemployment between jobs. It also excludes job to job flows involving very short duration jobs. These concerns are not not too serious for me, since I only use the EE measures as a comparison for the NE measure. I have also tried several variations on these definitions and found no significant differences in the results.

B Additional Empirical Results

Here I present results from some intermediate specifications which include some of the controls from the baseline specification. First, Figure 3 plots the estimated $m(\cdot)$'s when I only include size controls. These estimates fall somewhere in between Figure 1 and Figure 2. Apparently controlling for size brings us closer to the "within firm" specification, but does not eliminate the non-monotonicity seen in Figure 2.

[27] The LEHD data are based on employer payroll records, which are organized (within the employer) by pay period. Since pay periods may vary across employers it is in principal possible that a worker could make a job-to-job transition near the first day of a quarter, but only be recorded at a single employer in each quarter. This problem is attenuated by the short duration of pay periods relative to a calendar quarter.

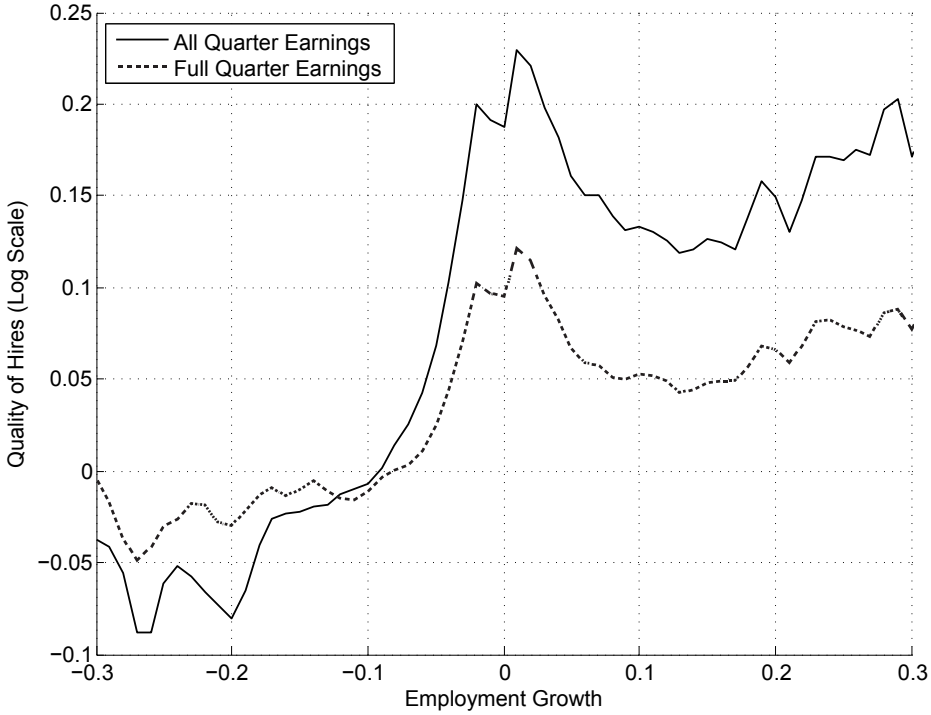

Figure 3: Size Controls Only

Next, Figure 4 adds firm fixed effects industry-season controls to the size controls. The results are much closer to the baseline specification, especially for all quarter earnings ($Q_{j,t}$). Both relations are on average positive and statistically significant. Qualitatively, the combination of firm fixed effects and size controls is enough to bring the results close to the the much richer baseline specification. A natural question is whether the firm fixed effects are truly necessary. Firm fixed effects are a convenient way to absorb all time invariant heterogeneity, but they might just proxy for data moments that are observable. In Figure 5 I experiment with a specification that is similar to the baseline specification, except that the firm fixed effects are replaced with a set of variables reflecting the firm's average worker and job flows. In particular, I include the mean and standard deviation of employment growth, the hires rate, and the separations rate over the observed lifetime of the firm. These six variables vary across firms but not over time, and capture the firm's average level and volatility of worker and job flows. Thus, they capture some characteristics that would have been picked up by a firm dummy. In Figure 5 we see that this modified regression does produce a nearly monotone relationship between the variables, though the shape of the relation is quiet different from Figure 1. This suggests that while the worker/job flow variables are useful controls, they do not substitute for the firm fixed effects.

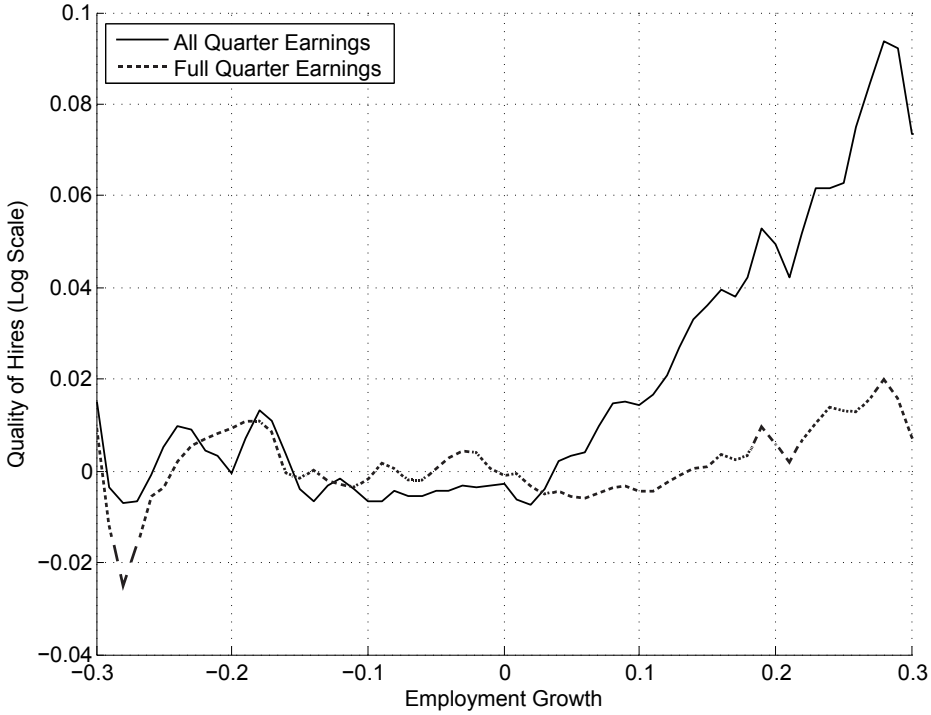

Figure 4: Size Controls and Firm Fixed Effects

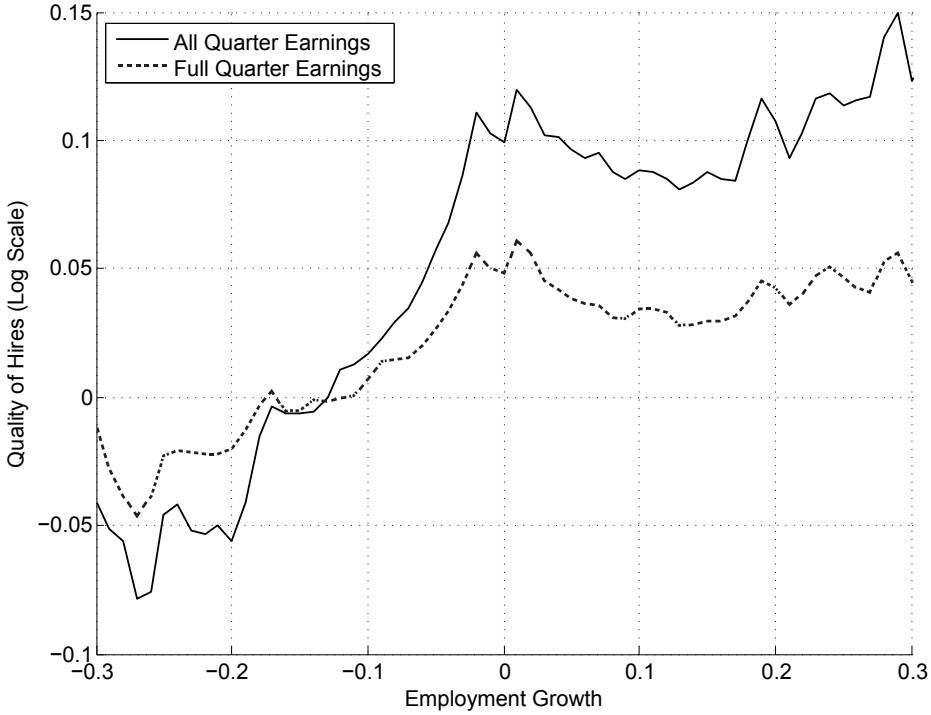

Figure 5: All Controls Except Firm Fixed Effects (Including Average Flows)

C The Firm's Problem

The firm's profit maximization problem is

$$\pi\left(N_{-1}(x),y\right) = \max_{N(x),v}\left\{\int_0^1 N(x)\left[F(x,y)-w(x,y)\right]dx - c(v) + \beta(1-p)\pi\left((1-s)N(x),y\right) + \beta p\int_0^1 h(y')\pi\left((1-s)N(x),y'\right)dy'\right\} \tag{32}$$

subject to

$$0 \leq N(x) \leq N_{-1}(x)+v\cdot q(x) \quad \forall x \tag{33}$$

$$y\in A(x) \quad \text{if} \quad N(x)>0 \ \forall x. \tag{34}$$

Turning to (33) first, $N_{-1}(x):[0,1]\to\mathbb{R}_+$ is the inherited stock of workers from the previous period, $q(x):[0,1]\to\mathbb{R}_+$ is the arrival rate of workers of each type, v is the firm's vacancy postings (a choice variable) and $N(x)$is the workforce in the current period (another choice variable). The constraint simply states that current employment of type x workers must be no greater than surviving employment from the past, $N_{-1}(x)$, plus new recruitment $v\cdot q(x)$. Note that the firm posts a single set of vacancies for all worker types.

The second constraint (34) imposes that the firm must be in the acceptance set of any worker they employ. If the worker did not find the match acceptable, they would depart into unemployment before production began.

In the profit maximization problem (32) the firm takes last period's employment and their current type as state variables. The firm chooses current employment and vacancy posting, taking the bargained wages $w(x,y)$ as given. Within the period, the firm enjoys profits net of wages $F(x,y)-w(x,y)$ from each worker, and pays the convex vacancy posting cost $c(v)$. If a productivity shock does not arrive the firm has discounted continuation value $\beta\pi\left((1-s)N(x),y\right)$, where s is the exogenous separation rate. If a shock does arrive the firm gets the expected value $\beta\int_0^1 h(y')\pi\left((1-s)N(x),y'\right)dy'$.

The Lagrangian for the firm's problem is

$$\pi\left(N_{-1}(x),y\right) = \int_0^1 N(x)\left[F(x,y)-w(x,y)\right]dx - c(v) + \beta(1-p)\pi\left((1-s)N(x),y\right) + \beta p\int_0^1 h(y')\pi\left((1-s)N(x),y'\right)dy' + \int_0^1 \mu(x)\left(N_{-1}(x)+v\cdot q(x)-N(x)\right)dx.$$

Assume that $y\in A(x)$. Then the match yields positive surplus, and the hiring constraint will bind. In this case the FOC with respect to $N(x)$ is

$$F(x,y)-w(x,y)+\beta(1-p)(1-s)\frac{d\pi}{dN_*} + \beta p(1-s)\int_0^1 h(y')\frac{d\pi}{dN_*}dy' - \mu(x) = 0.$$

Next, note that by the envelope condition

$$\frac{\partial \pi\left(N_{-1}(x), y\right)}{\partial N_{-1}(x)} = \mu_x \tag{35}$$

so when $y \in A(x)$ I can use the FOC to write

$$\begin{aligned} \frac{\partial \pi\left(N_{-1}(x), y\right)}{\partial N_{-1}(x)} &= F(x,y) - w(x,y) \\ &+ (1-s)(1-p)\beta \frac{\partial \pi\left(N_*(x), y\right)}{\partial N_*(x)} \\ &+ (1-s)p\beta \int_0^1 h(y') \frac{\partial \pi\left(N_*(x), y'\right)}{\partial N_*(x)} dy'. \end{aligned} \tag{36}$$

On the other hand, if $y \notin A(x)$ then worker separates from the firm before production occurs. This implies

$$\frac{\partial \pi\left(N_{-1}(x), y\right)}{\partial N_{-1}(x)} = 0 \text{ if } \quad y \notin A(x) \tag{37}$$

Taking equations (36) and (37) together for $y \in A(x)$ I have

$$\begin{aligned} \frac{\partial \pi\left(N_{-1}(x), y\right)}{\partial N_{-1}(x)} &= F(x,y) - w(x,y) \\ &+ (1-s)(1-p)\beta \frac{\partial \pi\left(N_*(x), y\right)}{\partial N_*(x)} \\ &+ (1-s)p\beta \int_{A(x)} h(y') \frac{\partial \pi\left(N_*(x), y'\right)}{\partial N_*(x)} dy' \end{aligned}$$

Which is identical to equation (3) when we set $\frac{\partial \pi(N_{-1}(x),y)}{\partial N_{-1}(x)} = J(x,y)$. This shows how we can formally derive equation (3) from the firm's problem. Taking the FOC with respect to v I have

$$c'(v) - \int_0^1 q(x)\mu(x)dx = 0.$$

The left hand side of this equation is the value of a vacancy. Optimality dictates that this be set to zero. Using equation 35 I can write

$$c'(v) = \int_0^1 q(x)J(x,y)dx.$$

D Assortative Matching with Productivity Shocks

Consider the worker's threshold condition (14):

$$F(x, y(x)) - \beta\eta\theta \int_{y(x)}^1 \left[\eta f(y) - (1-s)ph(y)\right]\left[F(x,y) - F(x, y(x))\right] dy = 0 \tag{38}$$

The following generalizes Proposition 3 to the case with productivity shocks.

Proposition 10. *When the worker's choice of $y(x)$ is given by (38), $y'(x) > 0$ if and only if*

$$\int_{y(x)}^{1} w^*(y) \frac{F_1(x,y)}{F(x,y)} dy > \frac{F_1(x,y(x))}{F(x,y(x))} \tag{39}$$

and $y'(x) < 0$ if and only if

$$\int_{y(x)}^{1} w^*(y) \frac{F_1(x,y)}{F(x,y)} dy < \frac{F_1(x,y(x))}{F(x,y(x))}, \tag{40}$$

where $w^*(y) = \frac{[\eta f(y)-(1-s)ph(y)]F(x,y)}{\int_{y(x)}^{1}[\eta f(y)-(1-s)ph(y)]F(x,y)dy}$ *and* $\int_{y(x)}^{1} w^*(y) = 1$.

Proof. Equation (38) implicitly defines $y(x)$, so differentiating it I can obtain an expression for $y'(x)$. Applying the implicit function theorem I have

$$y'(x) = -\frac{F_1(x,y(x)) - \beta\eta\theta \int_{y(x)}^{1} [\eta f(y) - (1-s)ph(y)] [F_1(x,y) - F_1(x,y(x))] dy}{F_2(x,y(x)) \left(1 + \beta\eta\theta \int_{y(x)}^{1} [\eta f(y) - (1-s)ph(y)] dy\right)}.$$

It is easy to show that the denominator is positive by substituting in for θ. So the sign of $y'(x)$ depends only on the numerator, implying $y'(x) > 0$ if and only if

$$\beta\theta\eta \int_{y(x)} [\eta f(y) - (1-s)ph(y)] F_1(x,y) dy \quad >$$
$$F_1(x,y(x)) \left(1 + \beta\theta\eta \int_{y(x)}^{1} [\eta f(y) - (1-s)ph(y)] dy\right)$$

Dividing this expression by (38) I arrive at

$$\frac{\int_{y(x)}^{1} [\eta f(y) - (1-s)ph(y)] F_1(x,y) dy}{\int_{y(x)}^{1} [\eta f(y) - (1-s)ph(y)] F(x,y) dy} > \frac{F_1(x,y(x))}{F(x,y(x))}$$

Setting $w^*(y) = \frac{[\eta f(y)-(1-s)ph(y)]F(x,y)}{\int_{y(x)}^{1}[\eta f(y)-(1-s)ph(y)]F(x,y)dy}$ I have equation (39)

$$\int_{y(x)}^{1} w^*(y) \frac{F_1(x,y)}{F(x,y)} dy > \frac{F_1(x,y(x))}{F(x,y(x))}$$

as desired. A symmetric argument yields equation (40). □

Note that the weights $w^*(y)$ may be negative if $f(y)$ is small and $h(y)$ is large. The sorting results summarized in Proposition 4 relied on the weights $w(y)$ being *positive*. To generalize Proposition 4 I need to show a single-crossing property holds for $w^*(y)$. This requires two intermediate results.

Result 1 First, I need to show

$$\int_{y(x)}^{1} [\eta f(y) - (1-s)ph(y)] F(x,y)\, dy > 0. \tag{41}$$

That is, although $[\eta f(y) - (1-s)ph(y)] F(x,y)$ may take negative values, the integral is always positive at an optimal threshold $y(x)$. To see this, rearrange (38) to get

$$F(x,y(x)) \left(1 + \beta\eta\theta \int_{y(x)}^{1} [\eta f(y) - (1-s)ph(y)]\, dy\right) = \beta\eta\theta \int_{y(x)}^{1} [\eta f(y) - (1-s)ph(y)] F(x,y)\, dy.$$

I have already argued in the proof of Proposition 10 that the term in parentheses is greater than zero, which then implies that the right hand side is positive, so equation (41) holds.

Result 2 Second, I need to show that there exists $\bar{y}$ such that

$$[\eta f(y) - (1-s)ph(y)] < 0 \iff y < \bar{y}. \tag{42}$$

That is, $[\eta f(y) - (1-s)ph(y)]$ has a single-crossing property, and is strictly positive above some threshold. To see this recall that $h(y)$ is not only the density of productivity shocks but is also the density of firm types (this follows from the nature of the productivity process). Then I can write $f(y) = \frac{M(U,V)}{U} \frac{h(y)v(y)}{V}$, where $v(y)$ is the number of vacancies posted by a type y firm and V is the aggregate number of vacancies. Using this I factor out $h(y)$ from the expression:

$$[\eta f(y) - (1-s)ph(y)] = h(y) \left[\eta \frac{M(U,V)}{U} \frac{v(y)}{V} - (1-s)p\right]$$

Next, recall from Section **3.1.6** that vacancy posting $v(y)$ is increasing in y. This shows that $\left[\eta \frac{M(U,V)}{U} \frac{v(y)}{V} - (1-s)p\right]$ will turn positive once and only once, proving (42). Intuitively, the difference between $f(y)$ and $h(y)$ is the number of vacancies posted per firm. More productive firms post more vacancies, so it is understandable that above a threshold the difference $\eta f(y) - (1-s)ph(y)$ will be positive.

Finally, taking equations (41) and (42) together, it is clear that the weights $w^*(y)$ also satisfy the same single crossing property:

$$w^*(y) < 0 \iff y < \bar{y}. \tag{43}$$

Given that $w^*(y)$ satisfied this condition, I can show a useful result.

Lemma 1. *Let $w^*(y)$ satisfy equation (43) and $\int_a^b w^*(y) = 1$. If $g(y)$ is increasing and positive then*

$$\int_a^b w^*(y) g(y) > g(a)$$

and if $g(y)$ is decreasing and positive then

$$\int_a^b w^*(y)g(y) < g(a)$$

Proof. Say that $g(y)$ is increasing and positive. Then I want to show

$$\int_a^b w^*(y)\left(g(y) - g(a)\right) > 0.$$

Note that $g^*(y) = g(y) - g(a)$ is increasing and non-negative. Then the desired condition can be written as:

$$\int_{\bar{y}}^b w^*(y)g^*(y) \quad > \quad \int_a^{\bar{y}} \left(-w^*(y)\right) g^*(y).$$

where $-w^*(y)$ is a positive weight for $y < \bar{y}$. Recall that $\int_a^b w^*(y)dy = 1$, so

$$\int_{\bar{y}}^b w^*(y) > \int_a^{\bar{y}} \left(-w^*(y)\right)$$

$$\int_{\bar{y}}^b w^*(y)g^*(\bar{y})dy > \int_a^{\bar{y}} \left(-w^*(y)\right) g^*(\bar{y})dy$$

Since $g^*(y)$ is increasing I have $\int_{\bar{y}}^b w^*(y)g^*(y) > \int_{\bar{y}}^b w^*(y)g^*(\bar{y})dy$ and $\int_a^{\bar{y}} \left(-w^*(y)\right) g^*(\bar{y})dy > \int_a^{\bar{y}} \left(-w^*(y)\right) g^*(y)$. Putting these all together I have

$$\int_{\bar{y}}^b w^*(y)g^*(y) > \int_a^{\bar{y}} \left(-w^*(y)\right) g^*(y)$$

as desired. □

Substituting $g(y) = \frac{F_1(x,y)}{F(x,y)}$ in Lemma 1, it is clear that $\frac{\partial}{\partial y}\frac{F_1(x,y)}{F(x,y)} > 0$ is sufficient for positive assortative matching in the generalized model. The conditions for negative sorting and no sorting follow from analogous arguments. I summarize the generalized results below.

Proposition 11. *Consider the general model with productivity shocks. Matching is positive assortative, $y'(x) > 0$, if F is strictly log-supermodular. Matching is negative assortative, $y'(x) < 0$ if F is strictly log-submodular. There is no sorting if $\frac{\partial}{\partial y}\frac{F_1(x,y)}{F(x,y)} = 0$.*

E Omitted Proofs

Proof of Fact Fact 1: First I derive an expression for $S(x,y)$. Rearranging (4) I can write

$$U(x) \quad = \quad \beta \int_{A(x)} f(y)\, S(x,y)\, dy + \beta U(x)$$

and rearranging equation (5) I have

$$\begin{aligned} W(x,y) &= w(x,y)+\beta W(x,y)-[s+(1-s)p]\beta S(x,y) \\ &\quad +(1-s)p\beta\int_{A(x)} h(y')S\left(x,y'\right)dy' \end{aligned}$$

Putting these together I have

$$\begin{aligned} S(x,y) &= w(x,y)+(1-s)(1-p)\beta S(x,y) \\ &\quad +(1-s)p\beta\int_{A(x)} h(y')S\left(x,y'\right)dy' \\ &\quad -\beta\int_{A(x)} f(y)S(x,y)\,dy. \end{aligned} \tag{44}$$

Next, note that the bargaining equation (6) implies $J(x,y)=\frac{1-\eta}{\eta}S(x,y)$. Using this in equation (3) yields

$$\begin{aligned} J(x,y) &= F(x,y)-w(x,y)+(1-s)(1-p)\frac{1-\eta}{\eta}\beta S(x,y) \\ &\quad +(1-s)p\frac{1-\eta}{\eta}\beta\int_{A(x)} h(y')S\left(x,y'\right)dy'. \end{aligned} \tag{45}$$

Combining (44), (45), the Nash bargaining equation and rearranging I have

$$w(x,y) = \eta F(x,y)+(1-\eta)\beta\int_{A(x)} f(y)S(x,y)\,dy \tag{46}$$

Using this in the worker surplus expression I have

$$\begin{aligned} S(x,y) &= \theta\eta\left(F(x,y)-\beta\int_{A(x)} f(y)S(x,y)\,dy\right) \\ &\quad +\theta(1-s)p\beta\int_{A(x)} h(y')S\left(x,y'\right)dy' \end{aligned} \tag{47}$$

where $\theta=\frac{1}{1-(1-s)(1-p)\beta}$. Similarly, the equation 45 yields

$$\begin{aligned} J(x,y) &= \theta(1-\eta)F(x,y)-\theta\eta\beta\int_{A(x)} f(y)J(x,y)\,dy \\ &\quad +\theta(1-s)p\beta\int_{A(x)} h(y')J\left(x,y'\right)dy'. \end{aligned}$$

□

Proof. Start with equation (9):

$$\begin{aligned} S(x,y) &= \theta\eta\left(F(x,y)-\beta\int_{A(x)} f(y)S(x,y)\,dy\right) \\ &\quad +\theta(1-s)p\beta\int_{A(x)} h(y')S\left(x,y'\right)dy' \end{aligned}$$

Rearranging this yields

$$\begin{aligned} S(x,y) &= \theta\eta F(x,y) \\ &\quad +\theta\beta\int_{A(x)}\left[(1-s)ph(y')-\eta f(y')\right]S\left(x,y'\right)dy' \end{aligned} \tag{48}$$

Multiplying by $[(1-s)ph(y)-\eta f(y)]$ and integrating I have

$$\begin{aligned} \int_{A(x)}[(1-s)ph(y)-\eta f(y)]\,S(x,y)\,dy &= \\ \theta\eta\int_{A(x)}[(1-s)ph(y)-\eta f(y)]\,F(x,y)\,dy \end{aligned}$$

$$\left\{\int_{A(x)}[(1-s)ph(y)-\eta f(y)]\,dy\right\}\theta\beta\int_{A(x)}\left[(1-s)ph(y')-\eta f(y')\right]S\left(x,y'\right)dy'$$

Which implies

$$\int_{A(x)}[(1-s)ph(y)-\eta f(y)]\,S(x,y)\,dy = \frac{\theta\eta\int_{A(x)}[(1-s)ph(y)-\eta f(y)]\,F(x,y)\,dy}{1-\theta\beta\int_{A(x)}[(1-s)ph(y)-\eta f(y)]\,dy}$$

And substituting this into (48)

$$\begin{aligned} S(x,y) &= \theta\eta F(x,y) \\ &\quad -\theta\eta\frac{\theta\beta\int_{A(x)}[\eta f(y)-(1-s)ph(y)]\,F(x,y)\,dy}{1+\theta\beta\int_{A(x)}[\eta f(y)-(1-s)ph(y)]\,dy} \end{aligned}$$

which is the desired result.

Continuity: S fails to be continuous only when $1-\theta\beta\int_{A(x)}[(1-s)ph(y)-\eta f(y)]\,dy=0$. For this to hold it must be that

$$1-\beta\theta(1-s)p\left[\int_A h(y)\,dy\right]\leq 0.$$

After some rearranging this becomes

$$\frac{1-\beta(1-s)+\beta(1-s)p\left(1-\left[\int_A h(y)\,dy\right]\right)}{1-\beta(1-s)(1-p)}\leq 0.$$

Note that $\left[\int_A h(y)\,dy\right]\leq 1$, so this is clearly a contradiction. So I have shown that $1-\beta\theta\left[\int_A((1-s)ph(y)-\eta f(y))\,dy\right]>0$, and thus S is continuous. □

Proof. Step 1 shows that there are never two points satisfying (14). Steps 2 and 3 show that (14) has a solution if and only if $y(x)$ is not a solution.

Step 1: There is at most one solution to (14). The derivative of the right hand side is

$$\begin{aligned} \frac{\partial}{\partial y^*}\beta\theta\eta\int_{y^*}^{1}f(y)\left(F(x,y)-F(x,y^*)\right)dy &= \\ -\beta\theta\eta F_2(x,y^*)\int_{y^*}f(y)dy &< 0 \end{aligned}$$

And the derivative of the left hand side is

$$\frac{\partial}{\partial y^*}\left[F(x,y^*)-b+\theta\beta(1-s)p\int_{y^*}^{1}h(y)(F(x,y)-F(x,y^*))\,dy\right] = F_2(x,y^*)\left[1-\theta\beta(1-s)p\int_{y^*}^{1}h(y)dy\right]$$

To see that $\left[1-\theta\beta(1-s)p\int_{y^*}^{1}h(y)dy\right]$ is positive, substitute in for θ to get

$$\left[1-\theta\beta(1-s)p\int_{y^*}^{1}h(y)\right]=\frac{1-\beta(1-s)\int_{y^*}^{1}h(y)dy}{1-\beta(1-s)(1-p)}>0.$$

Thus the left hand side of (14) is strictly decreasing and the right hand side is strictly increasing, so there is at most one solution.

Step 2: If there is no solution to (14) then (15) must hold. First, let

$$\begin{aligned} z(y^*) &= (F(x,y^*)-b)+\theta\beta(1-s)p\int_{y^*}^{1}h(y)(F(x,y)-F(x,y^*))\,dy \\ r(y^*) &= \beta\theta\eta\int_{y^*}^{1}f(y)(F(x,y)-F(x,y^*))\,dy \end{aligned}$$

From Step 1 we know that

$$\begin{aligned} z'(y^*) &> 0 \\ r'(y^*) &< 0 \end{aligned}$$

Note that

$$\begin{aligned} z(1) &> 0 \\ r(1) &= 0 \end{aligned}$$

Then, since z and r are continuous, the only way to have no solution to (14) is to have

$$z(x)>r(x)\ \ \forall x.$$

This implies $z(0)>r(0)$, which is equivalent to (**15**) so I am done.

Step 3: If (15) holds then (14) has no solution. Reversing the logic from Step 2, if $z(0)>r(0)$ then $z(x)>r(x)\ \ \forall x$, so (14) cannot hold. □

Proof of Proposition 5. The goal is to show that if F is not strictly log-supermodular then matching is not necessarily positive assortative. I start by assuming that

$$\frac{\partial^2}{\partial x\partial y}\ln F(x^*,y^*)<0 \tag{49}$$

for some $x^*$and y^*, and then show that I can pick $f(y)$ to make $y'(x^*)<0$. First, note that under these conditions we can find $y^{**}>y^*$ such that

$$\frac{F_1(x^*,y^*)}{F(x^*,y^*)}>\frac{F_1(x^*,y^{**})}{F(x^*,y^{**})}. \tag{50}$$

This will be used in the final step of the proof. Now I argue we can set $f(y)$ such that $y(x^*) = y^*$, and then show we can adjust $f(\cdot)$ to put more weight on y^{**}. It is trivial to show that there exist many meeting rate functions $f(\cdot)$ such that $y(x^*) = y^*$, so I simply assuming that we have $f(\cdot)$ such that the optimally condition (27) holds:

$$F(x,y^*)\left(1+\beta\theta\eta\int_{y^*}^{1} f(y)dy\right)-\beta\theta\eta\int_{y^*}^{1} f(y)F(x,y)\,dy = 0. \tag{51}$$

Next, define

$$\bar{f} \;=\; \int_{y(x)}^{1} f(y)dy \tag{52}$$

And define $\bar{y}$ as the solution to

$$F(x,\bar{y}) = \frac{\int_{y^*}^{1} f(y)F(x,y)\,dy}{\int_{y^*}^{1} f(y)dy} \tag{53}$$

Note that $\bar{y} > y^*$. $\bar{y}$ is the firm type that yields the expected match production within the acceptance set, while $\bar{f}$ is the probability the gets an acceptable offer. Then I can write

$$F(x,y^*)\left(1+\beta\theta\eta\bar{f}\right)-\beta\theta\eta\bar{f}F(x,\bar{y}) = 0. \tag{54}$$

This can be interpreted as the indifference condition of a counter factual problem. Say there is only one firm type $\bar{y}$, and the worker gets offers with probability $\bar{f}$. Obviously the worker accepts the offers and the decision problem is trivial. But, we can also examine what would happen if the worker were offered a one time chance to match with an alternative firm type y. Then this condition says that y^* is the lowest acceptable alternative firm type. Equivalently, we could imagine there is a measure zero population of firms with $y \neq \bar{y}$, and that the worker has to define their acceptance set on these alternatives as well.

Rearranging I have

$$\bar{f} = \frac{F(x,y^*)}{\beta\theta\eta\left(F(x,\bar{y})-F(x,y^*)\right)}.$$

Now define a new function

$$\bar{f}(y) = \frac{F(x,y^*)}{\beta\theta\eta\left(F(x,y)-F(x,y^*)\right)} \tag{55}$$

That is, $\bar{f}(y)$ is the *total* arrival rate necessary to ensure the worker's reservation type is y^* when all offers (except a set of measure zero) are of type y. $\bar{f}(y)$ is well defined for all $y > y^*$, and in particular $\bar{f}(y^{**})$ exists. In words, we can find an offer arrival rate that ensures the worker's reservation productivity is y^* when almost all firms are type y^{**}.

Equation (55) presumes a point mass of firms at type y. The results in Proposition 3 relied on continuous, differentiable densities. This is not a serious gap, since we can approximate the point mass as the limit of a sequence of smooth densities. The

critical point here is that F is twice continuously differentiable. This means that $\frac{\partial^2}{\partial x \partial y} \ln F(x,z) < 0$ implies $\frac{\partial^2}{\partial x \partial y} \ln F(x,z') < 0$ for some neighborhood U around z. Then we can replace the point mass at z with a density that accumulates on U.

Recall from Proposition 3 that $y'(x^*) > 0$ if and only if

$$\int_{y^*}^{1} w(y) \frac{F_1(x^*,y)}{F(x^*,y)} dy > \frac{F_1(x^*,y^*)}{F(x^*,y^*)} \tag{56}$$

where

$$w(y) = \frac{f(y)F(x,y)}{\int_{y^*}^{1} f(y)F(x,y)\,dy} > 0. \tag{57}$$

Clearly, if $f(y^{**}) \gg f(y)$ then $w(y^{**}) \gg w(y)$ as well. Under these circumstances

$$\int_{y^*}^{1} w(y) \frac{F_1(x^*,y)}{F(x^*,y)} dy \approx \frac{F_1(x^*,y^{**})}{F(x^*,y^{**})} < \frac{F_1(x^*,y^*)}{F(x^*,y^*)} \tag{58}$$

The inequality shows that $y'(x^*) < 0$, so we are done. □

Proof of Proposition 6. The proof has three steps.

Step 1: Growth is an exact function of current y and a set of state variables m that satisfy $h(y|m) = h(y)$. Consider a firm in period t. The firm's growth rate is defined as

$$g_t = 2\frac{E_t - E_{t-1}}{E_t + E_{t-1}}$$

Where E_t is total employment, measured in the production stage. E_t satisfies the law of motion $E_t = H_t + (1-s_t)E_{t-1}$, where H_t is total hires and s_t is the separation rate. We observe $s_t > s$ only when there is a negative productivity shock, leading to endogenous separations. Recall also that total hires $H_t = H(y_t)$ is a function only of current productivity y_t: both the number of vacancies posted and the acceptance decisions of workers only depend on y_t. Assuming $y_t = y_{t-1} = y_{t-2}$ and working recursively, we can write

$$E_t = H(y_t)(1 + (1-s)) + (1-s)^2 E_{t-2}$$

and assuming that $y = y_t$ has been constant since period $t-j$ we have

$$E_t = H(y) \sum_{i=0}^{i=j-1} (1-s)^i + (1-s)^j E_{t-j}. \tag{59}$$

Thus E_t is an exact function of (1) current productivity y, (2) the duration of the current productivity regime j, and (3) employment at the start of the regime E_{t-j}. Note that E_{t-j} is not independent of y: if y is very low, we expect E_{t-j} to be low as well, since productivity changed to y at the beginning of period $t-j$ or earlier. This means any high type worker will have separated before E_{t-j} is measured.

In what follows, I assume that a productivity shock arrived at the start of period $t-j$, and there has been no subsequent shock. Let $N : [0,1] \to \mathbb{R}_+$ be the firm's employment distribution in period $t-j-1$. That is, the firm employed $N(x)$

workers of type x in the production stage of period $t-j-1$. If the shock does not cause any separations then

$$\begin{aligned} E_{t-j} &= H(y)+(1-s)E_{t-j-1} \\ &= H(y)+(1-s)\int_0^1 N(x)dx. \end{aligned}$$

$\int_0^1 N(x)dx$ gives total employment at the end of period $t-j-1$. At the start of $t-j$, productivity changes to y and exogenous separations occur. So hiring is given by $H(y)$ and separation by $\int_0^1 N(x)dx$. In the general case, where endogenous separations may occur, the law of motion is

$$E_{t-j} = H(y)+(1-s)\int_0^{x(y)} N(x)dx. \tag{60}$$

Here $x(y)$ is the highest worker type that is willing to work for type y firm. The integral makes clear that any high type workers, $x > x(y)$, separate from the firm at the start of $t-j$. Putting (59) and (60) together I have

$$E_t = H(y)\sum_{i=0}^{i=j}(1-s)^i + (1-s)^{j+1}\int_0^{x(y)} N(x)dx.$$

Thus, E_t is an exact function of y, j and $N(x)$. Note y is independent of $N(x)$ and j: since $N(x)$ is only a function of *lagged* productivity values, it contains no information about y. In particular, the conditional density of y, $h(y|j,N(x))$ is equal to the unconditional density $h(y)$. Letting $m=(j,N(x))$, I am done.

Step 2: $g_1(y,m)$ is positive. Define

$$E^*_{t-j} = \int_0^{x(y)} N(x)dx. \tag{61}$$

where $x(y)$ and $N(x)$ are defined in Step 1. E^*_{t-j} is the mass of workers that remain at the firm after the productivity shock, but before exogenous separations occur. Using this notation I can write period t employment as be employment

$$E_t = H\sum_{i=0}^{i=j-1}(1-s)^i + (1-s)^j E^*_{t-j}$$

where H is the level of hiring for a type y firm. Differencing I have

$$E_t - E_{t-1} = (1-s)^{j-1}\left(H - sE^*_{t-j}\right)$$

so the growth rate is

$$g = 2r^{j-1}\frac{H - sE^*_{t-j}}{H\left[\sum_{i=0}^{i=j-2} r^i + \sum_{i=0}^{i=j-1} r^i\right] + E^*_{t-j}\left[r^j + r^{j-1}\right]}$$

where $r=(1-s)$. Substituting in equation (61) and $H=\int_0^{x(y)} H(x)dx$ I have

$$g = 2r^{j-1}\frac{\int_0^{x(y)} H(x)dx - s\int_0^{x(y)} N(x)dx}{\int_0^{x(y)} H(x)dx\left[\sum_{i=0}^{i=j-2} r^i + \sum_{i=0}^{i=j-1} r^i\right] + \int_0^{x(y)} N(x)dx\left[r^j + r^{j-1}\right]}$$

Differentiating this expression with respect to y, I find that $g' > 0$ if and only if

$$x'(y)\frac{H(x(y)) - sN(x(y))}{H(x(y))\left[\sum_{i=0}^{i=j-2} r^i + \sum_{i=0}^{i=j-1} r^i\right] + N(x(y))\left[r^j + r^{j-1}\right]} >$$
$$x'(y)\frac{\int_0^{x(y)} H(x)dx - s\int_0^{x(y)} N(x)dx}{\int_0^{x(y)} H(x)dx\left[\sum_{i=0}^{i=j-2} r^i + \sum_{i=0}^{i=j-1} r^i\right] + \int_0^{x(y)} N(x)dx\left[r^j + r^{j-1}\right]}$$

Under the assumption of positive assortative matching, $x'(y)$ is positive and can be canceled on both sides. Note that the remaining term on the left hand side is the growth rate of type $x(y)$ workers at the firm, while the right hand side is the growth rate of total employment at the firm. Appealing to Proposition 7, we see that the left hand side must be larger, so we are done.

Step 3: $g_1(y,m) > 0$ implies $F_g(\cdot|y')$ dominates $F_g(\cdot|y)$ whenever $y' > y$. Using the results in Step 1, I can write the distribution of growth conditional firm type as

$$F_g(g|y) = \frac{\int_{m,y} \mathbf{1}\{g(y,m) < g\} \times \mathbf{1}\{y^* = y\} dF(m,y^*)}{\int_{m,y} \mathbf{1}\{y^* = y\} dF(m,y^*)}$$

Where $F(m,y)$ is the joint CDF of m and y and $F_g(g|y)$ is the CDF of growth rates conditional on y. Using the fact that $h(y|m) = h(y)$ this simplifies to

$$F_g(g|y) = \int_m \mathbf{1}\{g(y,m) < g\} dF(m) \tag{62}$$

From equation (62) it is clear that increasing y reduces $F_g(g|y)$ if $g_1(y,m) > 0$. This means that $F_g(\cdot|y')$ dominates $F_g(g|y)$ whenever $y' > y$, so I am done. □

Proof of Proposition 7. Define $j(x)$ by

$$\begin{aligned} y_{t-j(x)} &\leq y(x) \\ y_{t-i} &> y(x) \quad \forall i > j(x). \end{aligned}$$

That is, $t - j(x)$ is the most recent period when the firm's productivity was below the worker's acceptance threshold. This implies that $N_{t-j(x)}(x) = 0$, so I can write employment in t as the accumulation of surviving hires since $t - j(x)$:

$$N_t(x) = \sum_{i=0}^{i=j(x)-1} (1-s)^i H_{t-i}(x). \tag{63}$$

Plugging this expression into the growth rate formula I have

$$g_t(x) = 2\frac{H_t(x) - s\sum_{i=1}^{i=j(x)-1}(1-s)^{i-1}H_{t-i}(x)}{H_t(x) + (2-s)\sum_{i=0}^{i=j(x)-1}(1-s)^{i-1}H_{t-i}(x)} \tag{64}$$

Clearly, a ceteris paribus increase in $j(x)$ reduces $g_t(x)$, This comes about through two channels. A higher $j(x)$ means there has been more time for the firm to accumulate workers, so the denominator of (31) is larger. Second, the larger stock of type x workers implies higher level of separations, reducing the numerator. The

final step is to note that $H_t(x) = q(x)v(y_t)$: hires of type x workers equals the arrival rate $q(x)$ times the number of vacancies posted, $v(y_t)$. Substituting in this expression, the $q(x)$ cancel leaving

$$g_t(x) = 2\frac{v(y_t) - s\sum_{i=1}^{i=j(x)-1}(1-s)^{i-1}v(y_{t-i})}{v(y_t) + (2-s)\sum_{i=0}^{i=j(x)-1}(1-s)^{i-1}v(y_{t-i})}. \tag{65}$$

This makes clear that changing x changes $g_t(x)$ only through $j(x)$, and not through any differences in hiring subsequent to $t - j(x)$. Under positive assortative matching, increasing x lowers $j(x)$, so $g_t(x)$ increases in x. □