Ranking Firms Using Revealed Preference*

Isaac Sorkin
FRB Chicago and the University of Michigan

November 2015

[Latest version]

Abstract

Firms account for a substantial share of earnings inequality. Although the standard explanation for why is that search frictions support an equilibrium with rents, this paper finds that compensating differentials are at least as important. To reach this finding, this paper develops a structural search model and estimates it on U.S. administrative data with 1.5 million firms and 100 million workers. The model analyzes the revealed preference information contained in how workers move between firms. Compensating differentials are revealed when workers systematically move to lower-paying firms, while rents are revealed when workers systematically move to higher-paying firms. With the number of parameters proportional to the number of firms (1.5 million), standard estimation approaches are infeasible. The paper develops an estimation approach that is feasible for data on this scale. The approach uses tools from numerical linear algebra to measure central tendency of worker flows, which is closely related to the ranking of firms revealed by workers' choices.

*isaac.sorkin@gmail.com. Thanks to Matthew D. Shapiro, John Bound, Daniel Ackerberg and Josh Hausman for patient advising and support. Thanks also to John Abowd, Audra Bowlus, Charles Brown, Jediphi Cabal, Varanya Chaubey, Tim Conley, Cynthia Doniger, Matthew Fiedler, Eric French, Paul Goldsmith-Pinkham, Henry Hyatt, Patrick Kline, Pawel Krolikowski, Margaret Levenstein, Kristin McCue, Erika McEntarfer, Andreas Mueller, Michael Mueller-Smith, Justin Wolfers, Mary Wootters, Eric Zwick and seminar and conference participants at University of Michigan, U.S. Census Bureau, UWO/MSU/UM Labo(u)r Day, European Search and Matching Conference, Barcelona GSE Summer Forum: Search, Matching and Sorting, SOLE/EALE, Census RDC Conference and the Chicago Fed for helpful comments and conversations. Thanks to Kristin McCue for help with the disclosure process. Thanks to David Gleich for making Matlab BGL publicly available. This research uses data from the U.S. Census Bureau's Longitudinal Employer Household Dynamics Program, which was partially supported by the following National Science Foundation Grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation. This research was supported in part by an NICHD training grant to the Population Studies Center at the University of Michigan (T32 HD007339) and the Robert V. Roosa Dissertation Fellowship. This research was also supported by a grant from the Sloan Foundation to the University of Michigan and by the Michigan Node of the NSF-Census Research Data Network (NSF-SES 1131500). Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the the Federal Reserve Bank of Chicago, the Federal Reserve System, or the U.S. Census Bureau. All results have been reviewed to ensure no confidential information is disclosed.

Firms play a central role in explaining worker earnings. Conditional on person fixed effects, firms account for over 20% of the variance of worker earnings (e.g., Abowd, Kramarz, and Margolis (1999) and Card, Heining, and Kline (2013)).[1] But there is little work asking why firms play such a central role. There are two main explanations that differ in whether high-paying firms are high-value firms: rents and compensating differentials. Rents is the leading explanation in the literature (e.g., Postel-Vinay and Robin (2002)). In the rents explanation, frictions prevent workers from bidding away the rents at high-paying firms. In this explanation, workers are lucky to end up at high-paying firms. As a result, high-paying firms are high-value firms. In contrast, in the compensating differentials explanation (e.g., Rosen (1986)), firms differ both in how much they pay and in nonpay characteristics. In this explanation, higher pay compensates for variation in unpleasant nonpay characteristics. As a result, high-paying firms are not high-value firms. Thus, measuring the relative importance of rents and compensating differentials means figuring out whether high-paying firms are high-value firms.

To distinguish between high-paying and high-value firms, this paper estimates the value of working at a firm without using information in pay. Specifically, I use information in quantities. To do so, I exploit the revealed preference idea embedded in search models that workers move to firms with higher value. Using U.S. matched employer-employee data with 1.5 million firms, I map the 1.5 million by 1.5 million matrix of worker flows across firms into the value of working at each firm.

By comparing the firm-level values to firm-level pay, I find that both rents and compensating differentials explanations are operative, but compensating differentials are more important. This finding has four (closely related) implications. First it shows that in many cases the conventional interpretation of high-paying firms as "good" firms is not warranted. Second, it contrasts with the conventional wisdom that compensating differentials are unimportant in explaining the variance of worker earnings. Third, it resolves a puzzle that benchmark search models are unable to reproduce the extent of earnings dispersion while also yielding plausible values of nonemployment. Fourth, the distribution of compensating differentials across firms means that the variance of earnings is larger than it would be if all jobs were equally pleasant.

In the first part of the paper, I write down a simple model of the labor market that contains both the rents explanation and the compensating differentials explanation and develop tools to estimate it using only quantity information. The model is a benchmark partial equilibrium utility-posting model in the spirit of Burdett and Mortensen (1998). The nonstandard ingredients in the model are first, that firms post a utility offer that consists of both earnings and a nonpecuniary bundle and second, that workers receive idiosyncratic utility draws (taste shocks) each period. On the one hand, the rents explanation is contained in the model because there is the possibility of equilibrium

[1] See also Andersson et al. (2012), Barth et al. (2014), and Song et al. (2015) for analyses of the role of employers and firms in the growth of earnings inequality in the U.S. In this paper, I use the word firm and employer interchangeably.

dispersion: different firms offer different levels of utility. On the other hand, the compensating differentials explanation is contained in the model because high earnings might be offset by a low nonpecuniary bundle. The role of the idiosyncratic utility draws is to add preference heterogeneity and allow workers to make different decisions when faced with the same choice. Specifically, these draws explain why there would be flows both from firm A to firm B and from firm B to firm A.

I estimate the model in two steps. In the first step, I isolate the transitions that reveal preferences by using information about what the worker's coworkers were doing at the time of the separation. In the spirit of the displaced worker literature (Jacobson, LaLonde, and Sullivan (1993)), if an unusually high share of coworkers were also separating, then a firm-level shock caused the workers to leave, and there is a high probability that any particular separation was exogenous. In contrast, if turnover levels look normal, then this separation was likely idiosyncratic to the worker.

In the second step, I measure the central tendency of worker flows. I record worker flows in a 1.5 million by 1.5 million matrix, where one cell is the number of workers who go from firm A to firm B. The model implies a set of linear restrictions on the entries in this matrix and a flow-relevant firm-level value. The flow-relevant firm-level value captures the central tendency of worker flows and is a known function of structural parameters.

Computing this central tendency of worker flows—and showing that it exists and has a meaningful economic interpretation—is the main technical contribution of this paper. The central tendency of worker flows is captured by the top eigenvector of a normalized matrix of worker flows. Showing when this eigenvector exists and is unique requires a new analytical result. Computing the eigenvector relies on techniques from numerical linear algebra that are scalable to massive datasets such as a 1.5 million by 1.5 million matrix.

In addition to the value of a firm, two factors affect the central tendency of mobility. First, a large firm will naturally have more workers moving away from it than a small firm. I account for this because I observe firm size. Second, a firm that makes a lot of offers will naturally have more workers moving towards it. I account for this because I estimate the offer distribution using information in nonemployment-to-employer flows. By jointly estimating the offer distribution and the value of nonemployment, I allow nonemployed workers to reject offers.

In the second part of the paper, I estimate the earnings that firms post and show how to combine these with the estimates of firm values to decompose the variation in firm-level earnings into compensating differentials and rents. The model implies that—as in Abowd and Schmutte (2014)—I can estimate the earnings that firms post using a selection-corrected version of the statistical decomposition pioneered by Abowd, Kramarz, and Margolis (1999) (and also used by Card, Heining, and Kline (2013)), which controls for person effects.

Because I use revealed preference at the firm level, I can identify compensating differentials. At the individual level, about a third of employer-to-employer moves result in earnings cuts. As the literature recognizes, this finding does not identify compensating differentials because the pay cut might reflect idiosyncratic factors that are not priced in the labor market. At the firm level, however, systematic patterns of workers moving from higher-paying to lower-paying firms indicates factors that are valued by all workers and are priced in the labor market.

Formally, I prove an identification result to show how to measure the relative roles of compensating differentials and rents. Combining utilities and earnings gives a lower bound on the variance of nonpecuniary characteristics, which is the extent of compensating differentials. The complement is the role of rents in explaining the variance of earnings. The identification result is consistent with an argument in the search literature that compensating differentials are hard to find because desirable nonearnings factors might be positively correlated with earnings (Hwang, Mortensen, and Reed (1998) and Lang and Majumdar (2004)). Specifically, I distinguish between *Rosen* and *Mortensen* amenities. *Rosen* amenities are priced in the labor market and generate compensating differentials. In contrast, *Mortensen* amenities are positively correlated with earnings and generate equilibrium dispersion; for example, some high-paying firms might also offer great benefits. The result shows that in my framework the variance of *Mortensen* amenities is not identified.

I estimate the model on the U.S. Census Bureau's Longitudinal Employer Household Dynamics (LEHD) dataset. Compensating differentials explain twice as much of the variance of firm-level earnings as rents. The estimated ranking of sectors is intuitively plausible, as is the implied distribution of nonpay characteristics. For example, education has good nonpay characteristics, while many blue-collar sectors have bad nonpay characteristics. The central finding that compensating differentials are at least as important as rents holds within subgroups defined by age and gender.

The finding that compensating differentials are important contrasts with conventional wisdom.[2] Research that looks at specific amenities rarely finds robust evidence that amenities are priced in the labor market. By using revealed preference at the firm-level, however, I overcome a few challenges in this literature. First, I distinguish between *Rosen* and *Mortensen* amenities, so allow for some amenities to be unpriced and contribute to dispersion in firm-level value. Second, revealed preference takes into account the entire bundle of amenities. As a result, it does not take a stand on what workers value, or how firms use amenities to offset other amenities. Third, at the firm-level, revealed preference uncovers amenities that are priced. In contrast, at the individual-level revealed preference can uncover idiosyncratic factors that are unlikely to be priced.

The finding of a large role for compensating differentials helps resolve the puzzle emphasized by Hornstein, Krusell, and Violante (2011) that benchmark search models cannot generate the extent of observed residual earnings inequality. I find that workers act as if a large share of the variance of firm-level earnings does not reflect variation in value. I also find an empirically reasonable value for nonemployment, and thus pass a crucial test they propose.

Finally, if the estimated nonpay characteristics were removed and earnings changed to compensate workers, then earnings inequality would decline. The effect depends on the correlation

[2]For example, Hornstein, Krusell, and Violante (2011, pg. 2883) write that compensating differentials "does not show too much promise" in explaining earnings dispersion.

between earnings potential and nonpay characteristics. In this counterfactual, earnings inequality as measured by the variance of earnings declines. Many low-paying jobs come with good amenities so this reduction comes mainly from the lower tail of the income distribution shifting up.

This paper unites two distinct literatures that seek to understand the structure of earnings. On the one hand, a reduced-form, descriptive literature has shown that industries and firms pay different amounts to identical workers (see, e.g., Krueger and Summers (1988), Abowd, Kramarz, and Margolis (1999), Card, Heining, and Kline (2013), Barth et al. (2014), and Song et al. (2015)). On the other hand, this reduced-form literature has inspired a structural literature that uses frictional search models to understand and quantify the forces that explain how high- and low-paying firms can coexist in equilibrium (see, e.g., Burdett and Mortensen (1998), Postel-Vinay and Robin (2002), Cahuc, Postel-Vinay, and Robin (2006) and Bagger and Lentz (2015)). This paper uses ideas from the structural literature to interpret the reduced-form literature. Specifically, by estimating a search model using quantities, I allow it to be result rather than an assumption that high-paying firms are high-value firms.

This paper uses how workers move across firms in a new way relative to existing literature. Bagger and Lentz (2015) also emphasize patterns in worker reallocation across firms, but they do not allow for nonpecuniary characteristics and do not exploit the complete structure of employer-to-employer moves. Similarly, Moscarini and Postel-Vinay (2014), Haltiwanger, Hyatt, and McEntarfer (2015) and Kahn and McEntarfer (2014) explore worker flows and ask whether these are consistent with a job ladder defined by a particular observable characteristic (e.g., size or wages). I invert the approach in these papers and instead construct the job ladder implied by worker flows.

The idea that earnings cuts identify amenities is shared with a few papers (e.g., Becker (2011), Hall and Mueller (2013), Sullivan and To (2014), and Taber and Vejlin (2013)). Only the last paper operates at the firm-level and relates variation in amenities to compensating differentials. In section 6, I discuss these papers in more detail.[3]

The estimation approach applies conditional choice probability estimation (Hotz and Miller (1993)) to matched employer-employee data, which allows gross worker flows between firms to exceed net flows. Other papers exploit similar modeling insights to study situations where gross flows exceed net flows; see e.g. Kline (2008) and Artuc, Chaudhuri, and McLaren (2010).

1 Ranking firms using revealed preference

1.1 A model with utility-posting firms

This section develops a partial equilibrium search model in the spirit of Burdett and Mortensen (1998) where firms post utility offers. This model is sometimes described through the metaphor of

[3] There is also literature (e.g., Gronberg and Reed (1994), Dey and Flinn (2005), Bonhomme and Jolivet (2009) and Aizawa and Fang (2015)) which estimates the value of specific observable amenities in a search environment. See section 6 for further discussion of this approach.

a job ladder, where there is a common ranking of firms and workers try to climb the ladder through employer-to-employer mobility.

To allow the model to contain both the rents and compensating differentials explanations and to rationalize workers making different choices, the structure of job value is nonstandard in two ways. First, it is in units of utility. By posting a level of utility, firms can create value for workers through both earnings and nonpecuniary characteristics. To connect to the second part of the paper, the value that firm i posts is given by:

$$\underbrace{V_i^e}_{\text{value}} = \underbrace{\omega}_{\text{conversion factor}} (\underbrace{\Psi_i}_{\text{earnings}} + \underbrace{a_i}_{\text{amenities}}).$$

The key innovation in this paper relative to the structural search literature is to directly estimate the V^e. The potential trade-off between earnings and nonpecuniary characteristics reflected in this equation allows the model to contain both the compensating differentials explanation and the rents explanation.

The second way in which the structure of job value is nonstandard is that in each period a worker receives a new idiosyncratic utility draw, which is the preference heterogeneity in the model. This preference heterogeneity explains why two workers would make different choices, and so we would observe workers moving from employer A to employer B and B to A. It also explains why over time a given worker's feelings about her employer might change and gives the model a theory of endogenous separations. While this preference heterogeneity is novel to the search literature, it is restrictive and is probably the most controversial assumption in the model. According to this assumption, all workers consider the same factors when making choices between firms, whereas in other models different workers might consider different things. Put differently, the i.i.d. assumption rules out persistent preference heterogeneity. A justification for this restriction is that firms appear to be governed by an equal treatment norm, where a desirable firm for a janitor is also a desirable firm for an executive. For example, Goldschmidt and Schmieder (2015, Figure 6) show that in Germany, high-paying firms for food, cleaning, security and logistics workers are also high-paying for workers outside these occupations. Similarly, I have estimated my model by men and women and young (18-34 years old) and old (35-61 years old) workers separately, and firms that are high-paying or high-value for one group are high-paying or high-value for the other group. (See Appendix G).

An additional assumption in the model is what Hall and Mueller (2013) term the proportionality-to-productivity hypothesis. Persistent heterogeneity only enters through a worker-specific constant that shifts the flow payoff to all employers as well as nonemployment. Specifically, this assumption means that the search parameters are the same for all workers, and so I can use the structure of the search model to infer rejected offers.

Because workers sometimes move between employers or to nonemployment involuntarily, the model contains both exogenous and endogenous moves. Exogenous moves are related to a firm-level shock and would typically be labelled involuntary. I identify these moves by building on the displaced worker literature, which aims to capture mass layoff events. In contrast, endogenous moves result from maximizing decisions in the model. From the worker perspective, some of these moves would be perceived as involuntary because they are in response to a negative idiosyncratic draw at the current firm. For example, in Mortensen and Pissarides (1994) all separations are endogenous and unpleasant for workers.

The following Bellman equation summarizes this verbal discussion of the model. A worker at employer i has the following value function:[4]

$$
\begin{aligned}
\underbrace{V^e(v_i)}_{\text{value of being at i}} = \underbrace{v_i}_{\text{flow payoff}} + \underbrace{\beta}_{\text{discounter}} \mathbb{E}\Bigg\{ & \underbrace{\delta_i \int_{\iota_1} \{V^n + \iota_1\} dI}_{\text{exogenous job destruction}} + \underbrace{\rho_i(1-\delta_i) \int_{V^e(v')} \int_{\iota_2} \{V^e(v') + \iota_2\} dI d\tilde{F}}_{\text{exogenous employer-to-employer (reallocation)}} \\
& + \underbrace{(1-\rho_i)(1-\delta_i)}_{\text{no exogenous shocks}} \times \Bigg[\underbrace{\underbrace{\lambda_1}_{\text{offer}} \int_{V^e(v')} \int_{\iota_3} \int_{\iota_4} \max\{\underbrace{V^e(v'') + \iota_3}_{\text{accept}}, \underbrace{V^e(v_i) + \iota_4}_{\text{reject}}\} dI dI dF}_{\text{endogenous employer-to-employer}} \\
& + \underbrace{\underbrace{(1-\lambda_1)}_{\text{no offer}} \int_{\iota_5} \int_{\iota_6} \max\{\underbrace{V^n + \iota_5}_{\text{accept}}, \underbrace{V^e(v_i) + \iota_6}_{\text{reject}}\} dI dI}_{\text{endogenous job destruction}} \Bigg] \Bigg\}. \qquad (1)
\end{aligned}
$$

Reading from left to right, a worker employed at i has value $V^e(v_i)$. This value consists of the deterministic flow payoff, v_i, and the continuation value, which she discounts by β. The flow payoff is the same for all workers at employer i and is the basis on which the model ranks and values employers. It represents the utility-relevant combination of pay, benefits, and non-wage amenities such as working conditions, status, location, or work-life balance at employer i. In addition, in every state workers also receive an idiosyncratic utility draw ι, which is drawn from a type I extreme value distribution.

The continuation value weights the expected value of four mutually exclusive possibilities. Two possibilities generate employer-to-employer transitions. A worker can be hit by a reallocation shock and forced to take a random draw from the offer distribution, or she can receive an offer and make a maximizing decision of whether to accept or reject it. And two possibilities generate employer-to-nonemployment transitions. A worker can be hit by a job destruction shock and forced to move to nonemployment, or she can make a maximizing choice to quit to nonemployment.

To estimate the offer distribution, I use information on where workers who are hired from

[4] The fact that the idiosyncratic shock shows up on the forward-looking values, rather than in the flow payoff, may look odd but is standard in the conditional choice probability literature. See Arcidiacono and Ellickson (2011, pg. 368).

nonemployment end up. A worker who is nonemployed has the Bellman equation:

$$\underbrace{V^n}_{\text{value of being nonemployed}} = \underbrace{b}_{\text{flow payoff}} + \underbrace{\beta}_{\text{discounter}} \mathbb{E}\{\underbrace{\underbrace{\lambda_0}_{\text{offer}} \int_{V^e(v')} \int_{\iota_7} \int_{\iota_8} \max\{\underbrace{V^e(v') + \iota_7}_{\text{accept}}, \underbrace{V^n + \iota_8}_{\text{reject}}\} dI dI dF}_{\text{endogenous nonemployment-to-employment}}$$

$$+ \underbrace{(1-\lambda_0) \int_{\iota_9} \{V^n + \iota_9\} dI}_{\text{no offer}}\}. \tag{2}$$

Reading from left to right, a nonemployed worker receives a total value of nonemployment of b, which includes both unemployment benefits as well as the value of nonmarket time and household production. Then each period two things might happen. She might receive an offer from an employer, in which case she decides whether or not to accept it. Or nothing might happen in which case she receives a new idiosyncratic draw associated with nonemployment.

1.2 Estimating the utility levels that firms post

This section shows how to estimate the utility levels that firms post. To allow for an arbitrary relationship between utilities and earnings, I do not use the earnings data in my estimation of the utilities.

The estimation procedure is semiparametric and is in the spirit of Postel-Vinay and Robin (2002). I do not impose parametric assumptions on the distribution of firm values or on the offer distribution. The estimation proceeds in two steps.

This section describes the two aspects of estimation that leverage novel features of the data, while Appendix B provides the complete details. The first novel aspect is in the calibration step. I use firm-level information in the spirit of the displaced worker literature to estimate which worker transitions do not reflect worker choices. The second novel aspect is that I develop a flow-relevant firm-level value—and prove when it exists and how to solve for it—that summarizes the central tendency of worker flows across employers (and to nonemployment) in terms of parameters of the model. With this value determined, the model reduces to an overidentified (by one equation) system of equations. I can then unravel this central tendency into the underlying structural parameters.

1.2.1 Calibration step: identifying displaced workers probabilistically

To calibrate the firm-specific shocks experienced by workers, I develop a continuous measure of displaced workers, or workers who left their employer because it was contracting. This amounts to identifying exogenous transitions as excess correlations in the transitions. Flaaen, Shapiro, and Sorkin (2015) use similar reasoning to detect whether workers report the proximate or ultimate reason for separations in surveys.

When an employer contracts by a lot, the displaced worker literature presumes the employer contraction caused the worker separations. To see this graphically, consider Figure 1a, which, using data and definitions described in Section 2, shows employer-to-employer and employer-to-nonemployment separation probabilities as a function of quarterly employer growth. The figure makes two points. First, even when an employer grows, workers separate. Second, as an employer starts to shrink, the probability of a separation rises. Starting with Jacobson, LaLonde, and Sullivan (1993), the displaced worker literature draws a vertical line at -30% and presumes all separations to the left were caused by the employer contracting and all separations to the right were not.

The continuous notion of displacement assigns an exogenous probability as the excess probability of separation at a contracting employer relative to that at an expanding employer. To see this graphically, Figure 1a draws a horizontal dashed line at the average of the employer-to-nonemployment separation rates at the expanding employers. The level of separations captured by this line are what would be expected in the absence of a firm-level shock.[5] Any separations above this line are excess and I attribute to a firm-level shock. The exogenous weight is then the excess probability relative to the total probability of a separation given the employer growth rate.[6]

I use the displacement probabilities to assign exogenous weights to each transition in the data. Figure 1b depicts the weights as a function of employer growth rates. To compute the overall probability of an exogenous job destruction shock (δ in the model), I sum over the exogenous probabilities of all employer-to-nonemployment transitions in the data and divide by the number of person-years in the data; while to compute the probability of an exogenous reallocation shock (ρ in the model), I sum over the exogenous probability of all the employer-to-employer transitions in the data and divide by the number of person-years in the data.[7]

The inference from administrative data that workers are being laidoff at contracting employers and are more likely to perceive separations as voluntary at expanding employers is supported by both employer- and worker-side survey data. Davis, Faberman, and Haltiwanger (2012, Figure 6) compute employers' reported reasons for separations as a function of firm growth rates. They find

[5]Figure A1 in Appendix H shows that the model reproduces the relationship between firm growth and the endogenous employer-to-employer and employer-to-nonemployment transitions in the data. The figure plots the probability of endogenous employer-to-employer and employer-to-nonemployment transitions as a function of firm growth rate in the data, as well as in the model. The figure shows that the model reproduces the upward slope on both the left and right of both the employer-to-employer and employer-to-nonemployment probabilities.

[6]Formally, let $\Pr(\text{EN}|g)$ be the probability of an employer-to-nonemployment transition at a firm that grows at rate g. Then let g_c be a particular rate of employer contraction. The endogenous probability at a firm contracting by g_c is $\Pr(\text{EN \& Endogenous}|g_c) = \frac{\Pr(\text{EN}|g>0)}{\Pr(\text{EN}|g_c)}$, while the exogenous probability is $\Pr(\text{EN \& Exogenous}|g_c) = 1 - \frac{\Pr(\text{EN}|g>0)}{\Pr(EN|g_c)}$.

[7]This approach to measuring exogenous separations using information on employer-side performance differs from typical approaches. The most common approach uses the rate of flows to nonemployment to estimate exogenous shocks (e.g., Postel-Vinay and Robin (2002, pg. 2319)), or uses the exogenous shocks to explain unexplained mobility (e.g., flows to worse firms as in Moscarini and Postel-Vinay (2014, pg. 15)). An alternative approach to removing layoffs more closely in the spirit of this paper is followed by Taber and Vejlin (2013, pg. 22, note 15), who drop all employer-to-employer transitions from employers that contract by more than 70%; similarly, Fox (2010, pg. 364) eliminates firm-years in which the firm closed. In both cases, they treat all other mobility as reflecting workers' choices.

patterns of employers' reports of quits and layoffs that are similar to the employer-to-employer line and employer-to-nonemployment lines in Figure 1a. Flaaen, Shapiro, and Sorkin (2015) compute workers' reported reasons for separations as a function of firm growth rates and find similar patterns. Specifically, the probability of separating and reporting all reasons for separations rise as firms contract, but "distress"-related reasons rise most rapidly.

1.2.2 Summarizing the central tendency of endogenous worker flows

The model implies a flow-relevant firm-level value that summarizes the complete structure of how workers move across firms and can be written in terms of underlying parameters of the model. To estimate this firm-level value, I derive a set of linear restrictions from the model on the values that firms post. To show when this firm-level value exists, I prove a graph theoretic result.

The goal of the firm-level value is to find values that rationalize the structure of flows between employers. I record endogenous flows between employers in a mobility matrix, denoted by M. The (i, j) entry in M is the number of nondisplaced workers flowing to employer i from employer j.[8] In the model, workers receive one offer at a time and, therefore, only ever make binary choices. Because I adopt the standard continuum assumption in discrete choice models, such flows from employer j to employer i are given by

$$\underbrace{M_{ij}}_{\text{j to i flows}} = \underbrace{g_j W}_{\text{\# of workers at j}} \underbrace{(1-\delta_j)(1-\rho_j)}_{\text{no shocks}} \underbrace{\lambda_1 f_i \Pr(i \succ j)}_{\text{get offer from i and accept}}, \tag{3}$$

where W is the number of employed workers. To interpret this equation, note that there are $g_j W$ workers at employer j and $(1-\delta_j)(1-\rho_j)$ share of them do not undergo exogenous separations. These workers get an offer from i with probability $\lambda_1 f_i$ and accept the offer with probability $\Pr(i \succ j)$.

The model implies a simple expression for the flow-relevant value of an employer and nonemployment. To derive this expression, consider relative flows between pairs of employers, which are given by:

$$\frac{M_{ij}}{M_{ji}} = \frac{f_i g_j (1-\delta_j)(1-\rho_j)\Pr(i \succ j)}{f_j g_i (1-\delta_i)(1-\rho_i)\Pr(j \succ i)}. \tag{4}$$

It is helpful to use the type I extreme value distribution to simplify $\Pr(i \succ j)$. Specifically, the type I extreme value distribution implies the differences in the error terms are distributed logistically so

[8] To continue with the formality from footnote 6, suppose that there is one worker who moves from j to i. Then the *observed* matrix is $1 = M^o_{ij}$. Further, suppose that firm j was contracting when the worker made the move. Then the *endogenous* matrix is $M_{ij} = \frac{\Pr(\text{EN}|g>0)}{\Pr(\text{EN}|g_c)} \times M^o_{ij}$.

that

$$\Pr(i \succ j) = \underbrace{\frac{exp(V^e(v_i))}{exp(V^e(v_i)) + exp(V^e(v_j))}}_{\text{logistic (difference of type I errors)}} \Rightarrow \frac{\Pr(i \succ j)}{\Pr(j \succ i)} = \frac{exp(V^e(v_i))}{exp(V^e(v_j))}. \tag{5}$$

Combining (4) and (5) gives:

$$\underbrace{\frac{M_{ij}}{M_{ji}}}_{\text{relative flows}} = \underbrace{\frac{f_i}{f_j}}_{\text{relative offers}} \times \underbrace{\frac{g_j}{g_i}}_{\text{relative size}} \times \underbrace{\frac{(1-\delta_j)(1-\rho_j)}{(1-\delta_i)(1-\rho_i)}}_{\text{relative shocks}} \times \underbrace{\frac{exp(V^e(v_i))}{exp(V^e(v_j))}}_{\text{relative values}}. \tag{6}$$

Relative flows (accepted offers) are directly related to relative values, but multiplied by relative offers and effective size. These additional terms account for the rejected offers. A firm that is large relative to the number of offers it makes must have had few rejected offers.[9] Now introduce notation which defines the flow-relevant firm-level value that summarizes the determinants of relative flows:

$$\underbrace{exp(\tilde{V}_i)}_{\text{flow-relevant value}} \equiv \underbrace{\frac{f_i exp(V^e(v_i))}{g_i(1-\delta_i)(1-\rho_i)}}_{\text{offer} \times \text{value / effective size}}. \tag{7}$$

$exp(\tilde{V}_i)$ is the flow-relevant value of an employer. It combines differences in the underlying value of an employer, as well as differences in (effective) size and the offer rate. For flows between employers and the nonemployed state, an analogous derivation implies:

$$\frac{M_{ni}}{M_{in}} = \frac{(1-\lambda_1)g_i W(1-\delta_i)(1-\rho_i)\Pr(n \succ i)}{\lambda_0 f_i U \Pr(i \succ n)} = \frac{exp(\tilde{V}_n)}{exp(\tilde{V}_i)}, \tag{8}$$

where

$$\underbrace{exp(\tilde{V}_n)}_{\text{flow-relevant value}} = \underbrace{\frac{(1-\lambda_1)W exp(V^n)}{\lambda_0 U}}_{\text{offers} \times \text{values / effect size}}. \tag{9}$$

Reading across, note the "offers" of nonemployment to employed workers occur when a worker does not get an outside offer, and the effective size of the nonemployed pool for valuation purposes is the number of workers who have offers.

[9] In the steady state where all firms are a constant size, this ratio is in fact sufficient to rank firms. If firms are growing or shrinking this ratio also reflects the recent growth trajectory. The approach developed here is not mechanically related to the rate of firm growth.

Combining (6) and (7) gives relative flows between employers in terms of $exp(\tilde{V}_i)$:

$$\underbrace{\frac{M_{ij}}{M_{ji}}}_{\text{relative flows}} = \underbrace{\frac{exp(\tilde{V}_i)}{exp(\tilde{V}_j)}}_{\text{relative flow-relevant values}}. \tag{10}$$

While it would be possible to take equation (10) directly to the data, doing so would run into two problems. First, for many pairs of firms there are not flows in both directions so this approach would not yield well-defined values of employers. Second, there is no guarantee that this approach would yield consistent valuations of employers. For example, there might be Condorcet cycle-like cases where combining the comparisons of employer A with employer B and employer B with employer C would give different relative valuations of employer A and employer C than the direct comparison of employer A and employer C.

To estimate the firm-level value, I now show that the model implies a set of linear restrictions. These restrictions address the two problems mentioned directly above. First, the restrictions imply a unique set of firm-level values that best explain all the flows. Second, the condition for uniqueness, which I discuss formally later in this section, is a restriction on the pattern of zeros in the flows that is much weaker than that each pair of firms have flows in both directions. Let $\mathcal{E}$ be the set of employers and n be the nonemployment state. Cross-multiplying (10) gives

$$M_{ij} exp(\tilde{V}_j) = M_{ji} exp(\tilde{V}_i), \ \forall j \in \mathcal{E} + n,$$

where the "for all" holds because the derivation of (10) goes through for all employers (as well as nonemployment). Summing across all employers on both sides gives

$$\underbrace{\sum_{j \in \mathcal{E}+n} M_{ij}}_{\text{\# of hires by i}} \underbrace{exp(\tilde{V}_j)}_{\text{flow-relevant value}} = \underbrace{\sum_{j \in \mathcal{E}+n} M_{ji}}_{\text{\# of exits from i}} \underbrace{exp(\tilde{V}_i)}_{\text{flow-relevant value}}. \tag{11}$$

Dividing by the summand on the right-hand side gives

$$\frac{\overbrace{\sum_{j \in \mathcal{E}+n} M_{ij} exp(\tilde{V}_j)}^{\text{value weighted hires}}}{\underbrace{\sum_{j \in \mathcal{E}+n} M_{ji}}_{\text{exits}}} = \underbrace{exp(\tilde{V}_i)}_{\text{flow-relevant value}}. \tag{12}$$

Equation (12) implies one linear restriction per firm (and one for nonemployment). The equation generates a recursive definition of employer quality: the quality of an employer depends on the quality of the employers it hires from, which in turn depends on the quality of the employers it hires from. In other words, the equation says that a good firm hires a lot from other good firms and has few workers leave.

To solve for the flow-relevant values, create the matrix version of equation (12). Specifically, define a square matrix S that is all zeros off-diagonal with the i^{th} diagonal entry being $S_{ii} = \sum_{j \in \mathcal{E}+n} M_{ji}$. Then letting $exp(\tilde{V})$ be the $|\mathcal{E}+n| \times 1$ vector that contains the firm-level $exp(\tilde{V}_i)$ and $exp(\tilde{V}_n)$ yields the following:

$$\underbrace{S^{-1}M}_{\text{normalized flows}} \quad \underbrace{exp(\tilde{V})}_{\text{flow-relevant values}} \quad = \quad \underbrace{exp(\tilde{V})}_{\text{flow-relevant values}} \quad . \tag{13}$$

This equation allows me to solve for $exp(\tilde{V})$. Intuitively, $exp(\tilde{V})$ is the fixed point of the function $S^{-1}M : \mathbb{R}^{|\mathcal{E}+n|} \rightarrow \mathbb{R}^{|\mathcal{E}+n|}$. In many settings in economics, fixed points can be found by starting with an initial guess and repeatedly applying the function to the resulting output until it converges. Despite the very high-dimensionality of the function, the same idea applies here.

To show when the $exp(\tilde{V})$ vector exists requires tools from graph theory. In the context of a linear system, the fixed point is an eigenvector corresponding to an eigenvalue of 1. For the iteration idea to work, the technical condition is that $S^{-1}M$ has an eigenvalue of 1 and this is the largest eigenvalue of the matrix.[10] Moreover, in order for the values to be interpretable, the $exp(\tilde{V})$ vector needs to be all positive so that $\tilde{V}$ is defined (the log of a negative number is not defined).

Result 1. *Let $S^{-1}M$ be matrices representing the set of flows across a set of employers and be defined as above. If the adjacency matrix associated with M represents a set of strongly connected employers, then there exists a unique-up-to-multiplicative-factor vector of the same sign $exp(\tilde{V})$ that solves the following set of equations:*

$$S^{-1}Mexp(\tilde{V}) = exp(\tilde{V}).$$

Proof. See Appendix A (also for graph theory definitions). □

This result shows that I can only estimate the value of employers in the *strongly connected* set. Strongly connected is a restriction on the pattern of zeros in the M matrix. To be in the strongly connected set, an employer has to both hire a worker from *and* have a worker hired by an employer in the strongly connected set. This result is intuitive. The information used to estimate values is relative flows. If an employer either never hires, or never has anyone leave, then we cannot figure out its relative value. To see this, consider equation (12). If a firm never hires, then its value is mechanically zero. Alternatively, if a firm has no workers leave, then the denominator is zero and

[10] In many other contexts the (top) eigenvector of matrices have been shown to have interesting economic content. Some examples include the following: the Leontief inverse of the input-output tables, and, in a network context, eigenvector centrality and Bonacich (1987) centrality. For some applications and development of these network ideas, see Ballester, Calvo-Armengol, and Zenou (2006), Acemoglu et al. (2012), and Elliott and Golub (2014).

the value of the firm is infinite. This result is related to the identification result in Abowd, Creecy, and Kramarz (2002) who show that the employer fixed effect in Abowd, Kramarz, and Margolis (1999) can only be estimated in the *connected* set of employers. To be in the connected set, an employer has to either hire a worker from *or* have a worker hired by an employer in the connected set.

Remark on Result 1: Because the search model implies that the S matrix is different than in standard applications, the novelty in result 1 is showing that the top eigenvalue is 1 (the Perron-Frobenius theorem is used to show that the top eigenvector is unique). S divides the i^{th} *row* of M by the i^{th} column sum of M. In other applications (e.g., Pinski and Narin (1976), Page et al. (1998) (Google's PageRank), and Palacios-Huerta and Volij (2004)), the normalizing matrix instead divides the i^{th} *column* of M by the i^{th} column sum. This normalization makes the resulting matrix a transition matrix and standard results imply that the top eigenvalue is 1. With the alternative normalization implied by the discrete choice model, standard results do not apply.

The diagonal entries in M are not defined using the model. The following result shows that because of the normalization, the top eigenvector of $S^{-1}M$ is invariant to the value of the diagonal entries in M.

Result 2. *Suppose that $exp(\tilde{V})$ is a solution to $exp(\tilde{V}) = S^{-1}Mexp(\tilde{V})$ for a particular set of $\{M_{i,i}\}_{i\in\mathcal{E}}$. Pick arbitrary alternative values of the diagonal: $\{M'_{i,i}\}_{i\in\mathcal{E}} \neq \{M_{i,i}\}_{i\in\mathcal{E}}$. Let S' and M' be the natural variants on S and M. If $exp(\tilde{V})$ solves the equation $exp(\tilde{V}) = S^{-1}Mexp(\tilde{V})$, then it also solves the equation $exp(\tilde{V}) = S'^{-1}M'exp(\tilde{V})$.*

Proof. See Appendix A. □

1.3 Summary

Appendix B shows how to take the employer-level value defined in equations (7)and (9) and infer the underlying values of employers and nonemployment. Here I summarize these steps. To facilitate this discussion, rearrange equation (7) to the following:

$$\underbrace{exp(V^e(v_i))}_{\text{value}} = \underbrace{\frac{exp(\tilde{V}_i)g_i(1-\delta_i)(1-\rho_i)}{f_i}}_{\text{flow-relevant value}\times\text{effective size/offers}} .$$

The object of interest is the firm value on the left-hand side. From the calibration step described in Section 1.2.1, I get the M matrix of endogenous moves and $\{\delta_i, \rho_i\}$.[11] The approach developed

[11] In practice, I use the population averages of δ_i and ρ_i instead of the firm-specific values. The reason is that with the firm-specific values, the model failed the overidentification test and was not able to find a value of nonemployment that generated the extent of employer-to-employer transitions. With the population averaged values, the model was able to do so. The reason the value of nonemployment is relevant for the extent of employer-to-employer transitions is that it affects the estimate of the arrival rate of offers while employed.

in Section 1.2.2 takes the M matrix and gives $exp(\tilde{V}_i)$. In addition, I observe g_i. The remaining term on the right-hand side is f_i, or the share of offers a firm makes. Intuitively, the structure of the employer-to-nonemployment transitions tells me the value of nonemployment, while the nonemployment-to-employer transitions tells me the offer distribution facing workers. By estimating the value of nonemployment and the offer distribution jointly, I allow nonemployed workers to reject offers.

2 Matched employer-employee data

This section describes the U.S. Census Bureau's Longitudinal Employer Household Dynamics data, which is a quarterly data that is constructed from unemployment insurance records.[12] The LEHD is matched employer-employee data and so allows me to follow workers across firms.

2.1 Data description

Three features of the LEHD should be kept in mind when interpreting the results. First, the notion of an employer in this dataset is a state-level unemployment insurance account, though the dataset follows workers across states.[13] Second, only employers that are covered by the unemployment insurance system appear in the dataset.[14] Overall, in 1994 the unemployment insurance system covered about 96% of employment and 92.5% of wages and salaries (BLS (1997, pg. 42)). Third, the unemployment insurance system measures earnings, but not hours.[15] Thus, variation in hours as well as in benefits will be included in my measure of compensating differentials.

Being able to track employers over time is central to measuring employer-to-employer flows and administrative errors in the employer identifiers would lead to an overstatement of flows. Following Benedetto et al. (2007), I assume that large groups of workers moving from employer A to employer

[12] See Abowd et al. (2009) for details.

[13] This can understate firm size for two reasons. First, for employers that operate in multiple states, this understates true employer size. Second, it is also possible for a given employer to have multiple unemployment insurance accounts within a state, which would also lead to an understatement of true employer size, though this is quantitatively unimportant (personal communication from Henry Hyatt (dated June 12, 2014): "the employment weighted fraction of firmids with multiple SEINs [state employer identification number] in a given state is about 1.5%, and...this fraction is actually lower in some of the larger states.") That said, working conditions are probably more similar within establishments than within employers, so having a "smaller" notion of an employer is desirable from the perspective of measuring compensating differentials.

[14] This restriction results in the exclusion of certain sectors of the economy. In particular, small nonprofit (organizations employing fewer than four workers), domestic, self-employed, some agricultural and federal government (but not state and local government) workers are excluded. For more complete discussions, see Kornfeld and Bloom (1999, pg. 173), BLS (1997, pg. 43) and `http://workforcesecurity.doleta.gov/unemploy/pdf/uilawcompar/2012/coverage.pdf`.

[15] The notion of earnings captured by UI records is as follows: "gross wages and salaries, bonuses, stock options, tips and other gratuities, and the value of meals and lodging" (BLS (1997, pg. 44)). This omits the following components of compensation: "employer contributions to Old-age, Survivors, and Disability Insurance (OASDI); health insurance; unemployment insurance; workers' compensation; and private pension and welfare funds" (BLS (1997, pg. 44)).

B in consecutive periods—especially if employer B did not previously exist—likely reflects errors in the administrative data rather than a genuine set of flows. As such, I correct the employer identifiers using worker flows. I use the Successor-Predecessor File and assume that if 70% or more of employer A's workers moved to employer B, then either 1) employer B is a relabelling of employer A or 2) employer B acquired employer A. Therefore, I do not count such "moves" as employer-to-employer transitions.

I pool data from 27 states from the fourth quarter of 2000 through the first quarter of 2008.[16] Pooling data means that I keep track of flows between as well as within these states. Like the dataset used by Topel and Ward (1992), the LEHD contains limited worker covariates. In particular, it includes age, race, and sex.

2.2 Dataset construction

I treat the model as an annual model, and I make several choices to go from the raw data to model-relevant objects.

To define a worker's employer, I reduce my dataset to one observation per person per year. The observation is the worker's *annual dominant* employer: the employer from which the worker made the most money in the calendar year. In addition, to facilitate coding transitions, I require that to count as an annual dominant employer the worker had two quarters of employment at the employer and that the second quarter occurred in the calender year.[17] I also restrict attention to workers aged 18-61 (inclusive) and, following Card, Heining, and Kline (2013), require that the annualized earnings exceed $3,250 (in 2011 dollars). Earnings are annualized by adjusting earnings for the number of quarters a worker was at a particular employer. With an annual dataset it is not possible to infer whether a change in dominant job was an employer-to-employer transition or an employer-to-nonemployment-to-employer transition.

To understand more about the transition, I use the quarterly detail of the LEHD to code transitions as employer-to-employer or employer-to-nonemployment-to-employer and to construct the displacement probabilities. I construct a quarterly dataset building on ideas in Bjelland et al. (2011) and Hyatt et al. (2014). First, I code a worker's transition between annual dominant employers as employer-to-employer or employer-to-nonemployment-to-employer. The transition is coded as employer-to-nonemployment-to-employer if between the annual dominant employers there is a quarter when the worker is nonemployed or has very low earnings. Second, I use the quarterly dataset to measure whether (and by how much) the employer was contracting in the quarter the worker separated, and to construct the relevant exogenous weight (see section 1.2.1). See Appendix

[16] I use the following states: CA, FL, GA, HI, ID, IL, IN, KS, MD, ME, MN, MO, MT, NC, ND, NJ, NM, NV, PA, OR, RI, SC, SD, TN, VA, WA, and WI. See Figure A2 in Appendix H for a map.

[17] Reduction to one observation per person per year is common. See Abowd, Kramarz, and Margolis (1999) (France), Abowd, Lengermann, and McKinney (2003) (US), Card, Heining, and Kline (2013) (Germany), and Card, Cardoso, and Kline (2014) (Portugal). Even outside of estimating statistical wage decompositions, Bagger et al. (2014) also reduce to one such observation per year.

C for details on my dataset construction (including how earnings are annualized) and Appendix D for a discussion of my computation.

2.3 Sample sizes

The sample restrictions that I impose to estimate the model eliminate many of the smallest firms where it would be hard to plausibly estimate a firm effect. A standard step in the estimation of search models is to impose a minimum size threshold; for example, Postel-Vinay and Robin (2002, pg. 2311) impose a minimum size threshold of 5. Table 1, column (1) shows the full annual sample. The subsequent columns show what happens to sample sizes and descriptive statistics as I restrict attention to the largest set of firms I can estimate the firm effects in (column (2)), the set of firms I use to compute the top eigenvector (column (3)), and the set that I estimate the model in (column (4)). Moving from column (1) to column (4), I lose very few person-years and people, but I do lose a large number of employers: I keep over 90% of person-years but only about 25% of employers. Under the assumption that the firms that I eliminate while moving from column (1) to column (4) exist for 7 years, the firms had 1.5 people per year on average. The second portion of Table 1 shows that the mean and variance of earnings are quite stable as I lose person-years and employers.

3 Features of the choice data

This section displays the patterns in the flow data that go into estimating the model and shows that the model is able to reproduce them.

3.1 Isolating endogenous moves: constructing the M matrix

Identifying compensating differentials relies on endogenous moves across firms. In the model, these moves are recorded in the M matrix. This section reports on the quantitative importance of the two steps I take to focus on endogenous moves. First, I separate transitions between annual dominant employers into employer-to-employer and employer-to-nonemployment-to-employer transitions. Second, even within employer-to-employer transitions, I use my firm-level calibration of shocks to downweight the moves where it is less likely that the worker had a choice to stay because the employer was contracting. Finally, even when I label a transition as endogenous, it does not necessarily mean that it feels good to the worker because it might be in response to a negative idiosyncratic (worker-match-specific) shock.

Fewer than half of the transitions between annual dominant employers are employer-to-employer. For workers in column (1) of Table 1, I code each transition between annual dominant jobs as employer-to-employer or employer-to-nonemployment-to-employer. The top rows of Table 2 shows that the annual separation probability is 25% and only 40% of these are employer-to-employer.[18]

[18] To compute a probability, the denominator is the number of person-years minus one per worker. The level of

Among employer-to-employer transitions, about three-quarters are endogenous. I assign the endogenous weight to each employer-to-employer separation. The table shows that 26% of employer-to-employer transitions are exogenous. At contracting employers, the share of exogenous employer-to-employer transitions is 48%.[19]

Only about a third of employer-to-nonemployment moves are associated with contracting employers. Aggregating across the exogenous weights on each employer-to-nonemployment transition, I find that 34% of the employer-to-nonemployment transitions are exogenous; that is, there are many separations to nonemployment in the absence of what looks like a firm-level shock. In the model, this is because workers, especially at the worst firms, sometimes quit to nonemployment following an idiosyncratic shock. At contracting firms, 56% of employer-to-nonemployment transitions are exogenous.[20]

Summing up across firms, I get novel estimates of the rate of two sources of exogenous mobility: the job destruction rate and the reallocation rate. Combining the exogenous weights and the separation probabilities gives an annual exogenous employer-to-nonemployment, or job destruction, rate (δ) of 0.05.[21] Combining the exogenous weights and the separation probabilities, gives an annual exogenous employer-to-employer, or reallocation, rate (ρ) of 0.03.[22] The remaining rows of Table 2 contain the other model parameters.[23]

transitions is slightly lower than—but the share of employer-to-employer transitions is similar to—what previous literature using similar definitions has found. For example, Bjelland et al. (2011, pg. 498) find that the quarterly employer-to-employer rate is about 4% (on an annualized basis about 16%), and employer-to-employer flows make up 27% of all separations, while Hyatt and McEntarfer (2012, Figure 2) find quarterly dominant job separation rates of about 16% (on an annualized basis about 50%) and close to half of such separations are employer-to-employer. I find lower separation rates for two reasons. I look at an annual frequency and so miss multiple separations within a year, and I also select for a slightly more stable populaton by imposing an earnings test.

[19] This exogenous share is between the estimates of Sullivan and To (2014, pg. 482, Table 1) who find that about 15% of employer-to-employer transitions are reallocations, and those of Moscarini and Postel-Vinay (2014, pg. 30), who find that about half of employer-to-employer transitions (49.7%) are exogenous in the sense used here. However, I do not use worker outcomes to categorize employer-to-employer transitions as being exogenous or endogenous.

[20] This ratio is the opposite of what Moscarini and Postel-Vinay (2014, pg. 21, Figure 11) find when separating employer-to-nonemployment transitions into layoffs and quits; they find that about $\frac{2}{3}$ of employer-to-nonemployment transitions are layoffs and $\frac{1}{3}$ are quits. This emphasizes that an endogenous separation is simply due to a shock idiosyncratic to the worker and that the separation might be labelled by the firm as a layoff.

[21] This annual job destruction rate is quite similar to Bagger and Lentz (2015, pg. 34), who estimate the annual job destruction rate for the low type worker in their model (which dominates their data) at 0.063. However, it is much lower than annualized versions of the monthly layoff rates estimated by Hornstein, Krusell, and Violante (2011, pg. 2879) (0.03). In general, monthly/quarterly views of the labor market do not aggregate in an i.i.d. way to annual values.

[22] The reallocation rate is quite a bit lower than the estimate of Bagger and Lentz (2015) of 0.106, but again, I do not use outcomes to code the reallocation rate.

[23] I estimate the arrival rate of offers on the job, λ_1, by matching the level of endogenous employer-to-employer transitions. The annual probability of receiving an offer is 0.20. (This is lower than the annualized offer rates reported by Moscarini and Postel-Vinay (2014, pg. 30) (about 30%) and Hornstein, Krusell, and Violante (2011, pg. 2889 and Figure 3) (at least 60%). This is presumably because I have annual data.)

3.2 The slippery ladder

Figure 2 shows that the model generates what the literature has termed a "slippery ladder." The figure sorts firms into 20 bins on the basis of firm-value. Within each bin, I compute the average probability of each kind of exogenous separation shock by summing across the exogenous weights, which are constructed using the variation depicted in Figure 1. The figure shows that job destruction and reallocation shocks are more likely at the worst firms. This means that the bottom rungs of ladder are "slippery" in that workers are more likely to experience exogenous shocks. This result is not mechanical, since the model does not use the exogenous transitions in computing firm values. This feature of the data has also been emphasized—or conjectured—by Jarosch (2014), Krolikowski (2014), Pinheiro and Visschers (2015) and Bagger and Lentz (2015).

3.3 Assessing the fit of the search model

The model does a reasonable job of fitting the choice ("revealed preference") information in the data along three dimensions. First, the model fits the pattern of workers at better firms being less likely to quit to take another job. Second, the model fits the pattern of workers at better firms being less likely to quit to nonemployment. Third, the model fits the detailed structure of the patterns of movements between employers on employer-to-employer transitions.

I first consider whether the model can match both the slope of the probability of making an employer-to-employer transition and the slope of the probability of making employer-to-nonemployment transition by firm value. This tests whether there is a single employer value that can rationalize the heterogeneity in both kinds of transitions. Heuristically, the slope of the probability of the employer-to-employer transitions tells us how different the firms are: in the model, the slope reflects the different probabilities across firms of workers accepting an offer from a random firm. Given these values, the model then picks the value of nonemployment to match the overall probability of endogenous employer-to-nonemployment transitions, but not the slope. It is not mechanical that a single firm value would match both slopes.

First, the model matches the probability of endogenous employer-to-employer transitions across employers. Figure 3a shows the probability of an endogenous employer-to-employer separation in the data and in the model. To construct the figure, I take all firms and sort them on the basis of the estimated firm values (V^e) into 20 bins with the same number of person-years. For each firm, I then compute the model-implied probability of an endogenous employer-to-employer separation. Within a bin, I average over the firm-specific probabilities implied by the model and in the data. The model implies that the probability of an endogenous employer-to-employer transition varies by a factor of about 5 from the bottom 5% of firms to the top 5% of firms. This variation is also present in the data.

Second, the model predicts that workers at better firms are less likely to make endogenous employer-to-nonemployment transitions, a pattern that is also in the data. Figure 3b shows the probability of an endogenous employer-to-nonemployment separation in the data and the model. The construction of the figure is analogous to the employer-to-employer figure. The figure shows that the model implies that the probability of an endogenous employer-to-nonemployment transition (one from an expanding firm) varies by a factor of about 10 from the bottom 5% of firms to the top 5% of firms, which nearly exactly matches the pattern in the data.

Third, the search model matches the detailed patterns of how workers make employer-to-employer transitions. To study how well the model matches the detailed structure of employer-to-employer transitions, I compare rankings based on "global" and "local" information. The model uses "global information" and says that A wins if $\tilde{V}_A > \tilde{V}_C$.[24] The local information is contained in binary comparisons. A binary comparison of employer A and employer C occurs when I observe accepted offers from A to C and C to A. In the binary comparison, I say that employer A wins if more workers join A from C than vice versa.

Any disagreement between the global and the local rankings should be explained by small samples. In small samples, these two rankings can differ. For an example of this phenomenon, consider Figure 4. In this figure, employer C wins on the binary comparison over employer A $(2 > 1)$. But when the accepted offers related to employer B are taken into account, the estimated value has A winning, or $\tilde{V}_A > \tilde{V}_C$. In large samples, this cannot happen.

The extent of disagreement between the global and local rankings is consistent with the model being the data-generating process. When I weight the comparisons by the number of accepted offers represented in each comparison, the model and the binary comparisons agree on 70.39% of comparisons. Is 70% big or small? This number allows me to reject the null of the model being equivalent to all firms having the same value (the neoclassical model of the labor market).[25] I find that the 90% confidence interval under the random null is [49.67%, 50.28%]. Under the null that the model is the data-generating process, the 90% confidence interval is [75.38%, 75.56%].[26] This means that the data are statistically inconsistent with the model being the data-generating process, but the economic magnitude of the rejection is not large. Thus, I conclude that the top eigenvector of the mobility matrix does a reasonable job of summarizing the structure of the employer-to-employer transitions.

[24] That is, I focus on the information in the accepted offers to parallel the binary comparisons. I use the values defined in equation (7).

[25] If in the data I observe 5 workers flowing from A to B and 10 workers flowing from B to A, then I take $10 + 5 = 15$ draws from a binomial distribution, where the probability of choosing A is 0.5. I ask what share of weighted comparisons the model and the binary comparisons agree on. I repeat this procedure 50 times to generate a null distribution under the hypothesis of all firms are equally appealing.

[26] I repeat the procedure described in footnote 25 except that the probability of choosing A is given by $\frac{exp(\tilde{V}_A)}{exp(\tilde{V}_A)+exp(\tilde{V}_B)}$, where $exp(\tilde{V}_A)$ is what I estimate in the model (and similarly for B) and the probability of choosing B is the remaining probability.

3.4 Dispersion in the labor market

There is dispersion in the labor market when measured either in terms of the value that employers provide to their workers or in terms of earnings (I discuss how I estimate the firm-level earnings in section 4.) Strikingly, however, there is more dispersion when it is measured in terms of earnings. This section also shows that patterns of left-shifts in the offer distribution predicted by search theory are found in both value and earnings space.

Figure 5a provides a new source of evidence that there is equilibrium dispersion. The figure plots the dispersion in the common values of firms (V^e) and the idiosyncratic draw (ι). In a benchmark frictional model (i.e., Burdett and Mortensen (1998)), workers agree on the ranking of firms and only voluntarily move towards better firms. All mobility is explained by common firm values and the model would find no dispersion in ι. Alternatively, in a benchmark neoclassical model in which firms play no special role (i.e., Topel and Ward (1992)), all mobility is explained by the idiosyncratic draw and the model would find no dispersion in V^e. The data are inconsistent with both extreme views. The dispersion in the common firm values and the dispersion in the idiosyncratic draws are similar. Based only on the information in quantities, there is clear evidence that there are good and bad firms. But these firm-level differences do not fully explain worker choices.

Figure 5b shows that there is less dispersion in values than in earnings. The figure plots the dispersion in the firm-level earnings (Ψ) and the residual in the earnings equation (r). A comparison of the two panels makes it clear that the common component is relatively more important in terms of earnings than in terms of values. This feature indicates that the systematic patterns in mobility are weaker than the systematic patterns in earnings changes.

Figure 6 shows that the offer distribution is left-shifted relative to the distribution of where workers are employed, whether this is measured in terms of values (top panel) or earnings (lower panel). The left shift in the earnings distribution is not mechanical, since the earnings information is not used in the estimation of the search model. In addition, the figure's panels compare the distributions of where workers are hired following nonemployment spells and reallocation shocks. In specifying the model, I do not allow the reallocated workers to reject offers (though this is not particularly central to the estimation). It is clear, however, that the reallocated workers on average end up at better firms than the workers hired from nonemployment. Thus, related to Bowlus and Vilhuber (2002), who argue that potentially displaced workers treat nonemployment as an outside option, there is evidence that workers who successfully make employer-to-employer transitions from contracting firms are more selective than workers who move from nonemployment.

4 Earnings

So far this paper has shown how to estimate the value of each firm on the basis of the choices that workers make. In this section I turn to estimating the earnings posted by each employer. Section

5 shows how to combine the values and earnings to decompose the variation in firm-level earnings into a rents component and a compensating differentials component.

4.1 Estimating earnings that firms post

To measure the earnings offered by firms, I use the following equation for log earnings (known as the Abowd, Kramarz, and Margolis (1999) decomposition):

$$\underbrace{y_{wt}}_{\text{log earnings}} = \underbrace{\alpha_w}_{\text{person effect}} + \underbrace{\Psi_{\boldsymbol{J}(w,t)}}_{\text{firm effect}} + \underbrace{x'_{wt}\beta}_{\text{covariates}} + \underbrace{r_{wt}}_{\text{residual/error term}}, \tag{14}$$

where y_{wt} is log earnings of person w at time t, α_w is a person fixed effect, $\Psi_{\boldsymbol{J}(w,t)}$ is the firm fixed effect at the employer j where worker w is employed at time t (denoted by $\boldsymbol{J}(w,t)$), and r is an error term. Canonically, x is a set of covariates including higher-order polynomial terms in age.[27]

Once I use the search model to selection-correct the equation, the firm effects in equation (14), Ψ, are the model-consistent notion of earnings, where the Ψ is the same firm-level earnings as discussed in section 1. The firm effects are identified by workers who switch firms. As such, the firm effects remove a time-invariant worker effect and, therefore, capture the earnings at a firm shared by all workers.[28] The one conceptual difference from the search model is that the search model contains a theory of the error term, whereas consistent estimation of equation (14) using movers requires that workers do not move on the basis of the error term. To remedy this inconsistency, I selection-correct equation (14) by inserting the expected value of the idiosyncratic utility draw calculated from the search model. Table A10 in Appendix H shows that this does not affect the firm effects. See Appendix E for details.

Knowledge of the firm effects allows me to quantify the role of firms in earnings using the following decomposition of the variance of earnings:

$$\underbrace{\mathrm{Var}(y_{wt})}_{\text{variance of earnings}} = \underbrace{\mathrm{Cov}(\alpha_w, y_{wt})}_{\text{person effect}} + \underbrace{\mathrm{Cov}(\Psi_{\boldsymbol{J}(w,t)}, y_{wt})}_{\text{firm effect}} + \underbrace{\mathrm{Cov}(x'_{wt}\beta, y_{wt})}_{\text{covariates}} + \underbrace{\mathrm{Cov}(r_{wt}, y_{wt})}_{\text{residual}}. \tag{15}$$

The share of the variance in earnings due to firm effects is given by:

$$\frac{\mathrm{Cov}(\Psi_{\boldsymbol{J}(w,t)}, y_{wt})}{\mathrm{Var}(y_{wt})}. \tag{16}$$

Firms play an important role in earnings determination. The third portion of Table 1 performs

[27] Because I only use seven years of data, the linear terms in the age-wage profile are highly correlated with the person fixed effects and thus, following Card, Heining, and Kline (2013), are omitted.

[28] To estimate the firm effects, researchers typically use both employer-to-employer and employer-to-nonemployment-to-employer movers to estimate the firm effects. Table A10 in Appendix H shows that restricting to just employer-to-employer transitions barely changes the firm effects (the correlation with the benchmark is 0.96). As such, I use both types of transitions in my main results.

the Abowd, Kramarz, and Margolis (1999) decomposition (equation (15)). About 22-23% of the variance of earnings is explained by employer-level heterogeneity.[29]

4.2 Earnings cuts are an important feature of the data

This section shows that seemingly voluntary cuts in earnings are widespread, are captured by the firm effects, and are not offset by future earnings increases. Besides emphasizing the prevalence of earnings cuts, this section serves two broader purposes in the paper. First, once selection-corrected, the model of the labor market underlying the Abowd, Kramarz, and Margolis (1999) earnings decomposition is the same as the model of the labor market reflected in the benchmark search model used in this paper. Since utility is unobserved while earnings are observed, there are more ways to test this model of the labor market using the earnings data than using the choice data. Second, by documenting that changes in firm effects are related to individual-level earnings changes and that such earnings cuts are widespread, this section begins to build the empirical case that something besides the pursuit of higher pay explains some employer-to-employer moves. In section 5, I directly relate the average direction of mobility in the labor market to the average pattern in earnings changes and develop a formal apparatus to interpret systematic patterns of moves to lower-paying firms as indicative of compensating differentials.

Individual-level earnings cuts are widespread in the data. Table 3 shows that 42% of transitions between annual dominant employers take earnings cuts. Even after restricting to endogenous transitions employer-to-employer transitions, 37% of these transitions take earnings cuts.[30] Hence, many employer-to-employer transitions cannot be explained by pursuit of higher-pay. By revealed preference, there must be some good nonearnings characteristics that justify these earnings cuts. The nonearnings characteristics, however, might be idiosyncratic to the firm-worker match and would not generate compensating differentials because such factors are not necessarily priced in the labor market.

To show that the earnings cuts are related to firm-level characteristics that we expect to be priced in the labor market, I now show that individual-level earnings cuts are related to firm-level pay in both probability and magnitude. The firm effects capture the probability of an earnings cut. Table 3 shows that 52% of the endogenous transitions to lower-paying firms have earnings cuts, while only 26% of the such transitions to higher-paying firms have earnings cuts. Figure 7 shows this fact graphically. Figure 7a plots the change in firm effects against the probability

[29] This share is broadly in line with the literature. For example, Card, Heining, and Kline (2013, Table IV) perform a related decomposition and find that the establishment share is about 20%. Comparing columns (2) and (3) of Table 1 also shows that changing the sample does not affect the decomposition.

[30] The share of earnings cuts is quantitatively consistent with evidence from survey datasets where researchers are able to calculate changes in hourly earnings. Jolivet, Postel-Vinay, and Robin (2006, Table 1) find that in the Panel Study of Income Dynamics 23.1% of job-to-job transitions come with an earnings cut. For the Survey of Income and Program Participation, Tjaden and Wellschmied (2014, Table 2) find 34%. And for the National Longitudinal Survey of Youth 1997, Sullivan and To (2014, Table 1) find 36%.

of taking an earnings cut on an employer-to-employer transition, while Figure 7b shows this for employer-to-nonemployment-to-employer transitions. The probability of an earnings cut monotonically decreases as workers move to higher-paying firms: from the largest downward moves to the largest upward moves, the probability of an earnings cut falls from about 80% to 10%.

The firm effects capture the magnitude of the earnings cuts. As emphasized by Chetty, Friedman, and Rockoff (2014), a measure of bias in firm effects (or, in their case, teacher value-added) is to consider the β_1 coefficient in the following regression:

$$\text{individual-level earnings change} = \beta_0 + \beta_1 \text{firm effect change}.$$

If the firm effects are unbiased, then we expect $\beta_1 = 1$.[31] The x-axis of Figure 8 plots the vigintiles of changes in firm effects at transitions between annual dominant employers against the average individual-level change in earnings on these transitions. The solid line plots the best-fitting line from a regression run on the individual-level data. The thin-dashed line shows the line that would be expected if the firm effects were unbiased. The two lines are not that far apart. The lower panel shows that the fit is better in the employer-to-nonemployment-to-employer transitions than in the employer-to-employer transitions. (For employer-to-employer transitions the slope coefficient is 0.82; for employer-to-nonemployment-to-employer transitions it is 1.03.)

It is not mechanical that changes in firm effects capture the magnitude of the earnings cuts. Specifically, this result is not predicted by models where mobility is on the basis of comparative advantage (e.g., Eeckhout and Kircher (2011), Lopes de Melo (2013) and Hagedorn, Law, and Manovskii (2014)). A natural concern is that the firm effects summarize the information in the individual earnings changes, which would suggest that the fit in Figure 8 is mechanical. But two features of this fit do not occur in data simulated from a model where mobility is on the basis of comparative advantage. First, in these models there are no earnings cuts, and the individual-level earnings changes always lie above the x-axis. Second, in these models there is not the approximate symmetry in earnings changes from moving to a better or a worse firm (Card, Heining, and Kline (2013, pg. 990) and Card, Cardoso, and Kline (2014, pg. 10) emphasize this symmetry property). Figure A3 plots the analogous figure to Figure 8 with data simulated from the example production function in Eeckhout and Kircher (2011). The estimate of β_1 is about 0.4, and unlike in the data, the earnings changes display a v-shape in the firm effects changes.[32]

[31] Chetty, Friedman, and Rockoff (2014) perform this exercise using a leave-one-out estimator to eliminate the mechanical correlation from the fact that the individual test scores are used to estimate the teacher value-added. In my case, leave-one-out is computationally infeasible, since the process of building the matrices and estimating the firm effects takes at least an hour. Nevertheless, some evidence that this mechanical effect is unlikely to be important comes from the fact that the firm effects are very stable when estimating using subgroups. For example, the correlation between the firm effects estimated separately on men and women is 0.92 and on people aged 18-34 and 35-61 the correlation is 0.86.

[32] The v-shape comes from earnings increases accruing to workers whose comparative advantage is working at the worst firms. In Eeckhout and Kircher (2011), there is an absolute ranking of firms and workers in terms of productivity. Because the opportunity cost of jobs varies as a function of worker and firm types, however, each

The earnings cuts captured by moving to lower-paying firms are not offset by future earnings increases. One explanation for earnings cuts is that workers accept cuts in exchange for the possibility of steeper earnings profiles. Following Abowd, Kramarz, and Margolis (1999), I estimate firm-specific earnings slopes using the wage growth of the stayers. Figure 9 shows that when workers move to lower-paying firms they do not move to firms offering steeper slopes in earnings (the coefficient is 0.00). Similarly, the firm effects in the intercept are weakly positively correlated with the slope when estimated in separate regression (the correlation is 0.02) and only weakly negatively correlated when estimated in the same regression (the correlation is -0.03). These results are quantitatively different than the leading theory that explains earnings cuts as a function of an option value of a future increase. I simulate Papp (2013)'s calibration of Cahuc, Postel-Vinay, and Robin (2006), which matches the share of earnings cuts in the data. In the simulated data, the correlation between the firm effects in the intercept and the slope is -0.90. This result implies that earnings cuts are not explained by the possibility of future earnings increases at the same firm.

5 Why do some firms pay so much and some so little?

So far I have shown how to estimate a value of each employer as revealed by worker choices, as well as the earnings at each firm. This section shows that combining these two measures allows a decomposition of the variance of firm earnings into a part explained by rents and a part explained by compensating differentials. Compensating differentials are more important than rents in explaining why some firms pay so much and some so little.

5.1 Measuring compensating differentials and rents

This section shows that the relationship between the values and earnings that firms post is sufficient to decompose the variance of earnings into a component explained by compensating differentials and a component explained by rents.

A firm i posts a combination of earnings and nonpecuniary characteristics that leads to a firm-wide value. The forward-looking value of being employed at firm i, V_i^e, takes the following additively separable form:

$$\underbrace{V_i^e}_{\text{forward-looking values}} = \underbrace{\omega}_{\text{value per log dollar}} \times (\underbrace{\Psi_i}_{\text{earnings}} + \underbrace{a_i}_{\text{amenities/nonpecuniary}}), \tag{17}$$

worker has a unique ranking of firms in terms of pay (i.e., the highest-paying firm for the low-type worker is a low-type firm, even though the low-type worker would be most productive at the high-type firm). The left-hand side of the v comes from low-type workers moving to low-type firms. These are firms that are the highest paying for the low-type workers, and thus they get a pay increase, but are low-paying for most workers. In a model of circular heterogeneity without absolute advantage like Marimon and Zilibotti (1999), there is no variance in the estimated firm effects because all firms are equally high- and low-paying for the same number of workers and in the same way.

where ω is value per log dollar, Ψ is the firm-specific earnings, and a_i is the nonpecuniary or amenity bundle at the firm. There are two related features of how a is defined worth noting. First, the nonpecuniary bundle is in the same units as earnings. These units make it meaningful to talk about trading off earnings and the nonpecuniary bundle. Second, because earnings are a flow, a is defined in flow-equivalent units. Being in flow units means that a captures differences in the riskiness of firms; for example, in the model there are firm-level differences in the rates of job destruction.

Compensating differentials generate variation in earnings when a worker makes a one-for-one trade-off between earnings and nonpecuniary characteristics. To see this formally, suppose we observe a at the firm-level and run a regression $\Psi = \beta a$. If a reflects only nonpecuniary characteristics that generate variation in earnings through compensating differentials, then in equilibrium we should estimate $\hat{\beta} = -1$, that is, workers are willing to pay one for one unit of the nonpecuniary bundle.

In contrast, rents generate variation in earnings—resulting in equilibrium dispersion—when higher-earnings firms are higher-utility firms (this can be an equilibrium outcome because of frictions). This part of earnings reflects rents because in the absence of frictions competitive pressure would push all firms to offer the same level of utility (but not necessarily the same earnings).

To understand what can be identified using the patterns of utilities and earnings that firms post, it is helpful to consider the decomposition of the nonpecuniary part of firm value into two components: $a = a_{\text{Rosen}} + a_{\text{Mortensen}}$.[33] The *Rosen* amenities are priced in the labor market and generate compensating differentials. In contrast, the *Mortensen* amenities are not priced in the labor market and contribute to equilibrium dispersion. For example, many benefits appear to be positively correlated with earnings. The following result is the basis of how I measure compensating differentials and rents.

Result 3. *Suppose that the utility function is given by equation (17) and that $\{V_i^e\}_{i \in \mathcal{E}}$ and $\{\Psi_i\}_{i \in \mathcal{E}}$ are known. Then*

$$Var(a_{Rosen}) = (1 - R^2) Var(\Psi),$$
$$Var(a_{Mortensen}) \in [0, \infty)$$

and combining them, we have

$$Var(a) \in [Var(\Psi)(1 - R^2), +\infty),$$

where $R^2 = Corr(V^e, \Psi)^2$. The willingness to pay for a Rosen *amenity is one. The amenities are*

[33] These components are labelled in honor of Rosen (1986) and Mortensen (2003). I could equally name the second term after Lang and Majumdar (2004) who, like Hwang, Mortensen, and Reed (1998), construct examples in search environments in which it is an equilibrium for Ψ and a to be perfectly positively correlated. Similarly, I could name the second term after Pierce (2001), who presents evidence that benefits are positively correlated with earnings.

related to earnings as follows:

$$Corr(\Psi, a_{Rosen}) = -\sqrt{1-R^2}, \text{ and}$$
$$Corr(\Psi, a_{Mortensen}) = \sqrt{R^2}.$$

Bounds on the variance of utility in log dollar units are:

$$Var(\Psi + a) \in [Var(\Psi)R^2, \infty).$$

Proof. See Appendix A. ☐

I use the result to decompose the variation in firm-level earnings into compensating differentials and rents. Specifically, the result makes the following accounting identity interpretable:

$$\underbrace{Var(\Psi)}_{\text{variance of firm effects}} = \underbrace{R^2 Var(\Psi)}_{\text{rents}} + \underbrace{(1-R^2)Var(\Psi)}_{\text{compensating differentials}},$$

where $R^2 = Corr(V^e, \Psi)^2$. The first term reflects rents because it is the part of the variance of firm effects that is accounted for by variation in utility in the labor market. The second term reflects compensating differentials because it is the variance of *Rosen* amenities.

Because the variance of *Mortensen* amenities is not identified, there are three interesting quantities that I cannot point identify. The first is the overall variance of nonpecuniary characteristics. The second is the overall correlation between earnings and nonpecuniary characteristics. And third, I can only provide a (informative) lower bound on the variance of utility offered by firms.

5.2 Sector-level evidence

This section discusses the utilities and earnings that firms post aggregated to the sector level.[34] At this level of aggregation, we can assess the intuitive plausibility of the results of the model.

Figure 10a shows that there are similarities between these rankings, which provides evidence of rents. It shows the scatterplot of the sector-level values and earnings, as well as the best-fitting line. The evidence for rents is that some high-paying sectors are high-utility (or "good") sectors and some low-paying sectors are low-utility (or "bad") sectors. For example, both approaches agree that hotels/restaurants is a bad sector in which to work. And both approaches agree that utilities is a good sector in which to work.

That said, there are also differences between these rankings, which provides evidence of compensating differentials. The most striking sector is education. Specifically, among sectors offering a similar level of utility (equally far out on the x-axis), it is the worst paid, which indicates the

[34]A sector is a slightly more aggregated than a 2-digit NAICS. Because of disclosure limitations, I cannot report results about individual firms.

presence of good nonearnings characteristics. Similarly, public administration is low-paying given how much utility it offers. In contrast, traditionally blue-collar and male sectors—mining, manufacturing, and construction—tend to be quite high-paying for how much utility they offer, which implies the presence of bad nonpecuniary characteristics.[35]

Aggregating the similarities and differences between the rankings, the R^2 between values and earnings—on an employment-weighted basis—is 0.45. Thus, 45% of the intersectoral wage structure is rents and 55% is compensating differentials.

5.3 First-pass answer

In Section 5.1, I showed that a quantitative answer to the question why some firms systematically pay some workers so much and some so little is provided by computing the R^2 between firm-level earnings and utility. The first row of Table 4 reports that 25% of the variance of firm-level earnings is related to utility so that compensating differentials account for 75% of the variance of firm-level earnings.

Figure 10b shows that the relationship between the firm-level values and earnings is approximately linear. It shows a binned scatterplot of the firm-level values and earnings as well as the line of best fit on the firm-by-firm data, where I have grouped firms into 20 equal employment-weighted bins (vigintiles). The approximate linearity of the relationship implies that the correlation (or R^2) captures the relationship.[36]

Figure 10b also illustrates the identification result. Rents can be seen in the upward slope in the line of best fit, which shows the variation in firm-level earnings that is reflected in variation in values. Compensating differentials—or *Rosen* amenities—are variation in firm-level earnings holding utility constant. In the figure, this variation is depicted in the red dashed lines that show plus and minus one standard deviation bands of firm-level earnings holding utility constant. The firms above the line have relatively bad *Rosen* amenities and the firms below the line have relatively good *Rosen* amenities. In contrast, the fact that the y-axis is in log dollar units, while the x-axis is in units of the standard deviation of the idiosyncratic utility draw, explains why the *Mortensen* amenities are not identified. *Mortensen* amenities stretch out the x-axis in log dollar units. But the unit conversion from log dollars to the standard deviation of the idiosyncratic draw is not known, so this framework does not identify the importance of *Mortensen* amenities.

[35] In Sorkin (2015), I explore the implications of patterns like this for gender earnings gap.

[36] Some readers might wonder how this correlation relates to tabulations showing that workers are more likely to transition to higher-paying firms than to lower-paying firms as measured by Abowd, Kramarz, and Margolis (1999) firm effects; see, for example, Card, Heining, and Kline (2012, Appendix Table 3) or Schmutte (2015, Table 3). A given transition matrix across firm effects is consistent with almost any correlation. The estimation of $\tilde{V}$ constructs the ranking that results in the "best-fitting" transition matrix (according to a specific loss function). This best-fitting transition matrix might be nearly identical to the transition matrix using the firm effects in earnings (in which case the correlation would be 1), or radically different (in which case the correlation could approach zero; the lower-bound depends on how well the firm effects in earnings predict mobility patterns). It is also important to note that the transition matrix (and $\tilde{V}$) only contains information in accepted offers.

5.4 Addressing measurement error

Measurement error in either earnings or utility leads me to overstate the role of compensating differentials. Because there is potentially measurement error in both the left-hand-side and the right-hand-side variables, standard approaches are not applicable. Grouping firms by exogenous characteristics addresses measurement error. This leads me to revise down the role of compensating differentials presented in the previous section.

Because I have administrative data, the main source of measurement error is the fact that the firm-level values and earnings are estimated. In particular, for consistent estimates of the R^2, I need a law of large numbers to obtain *within* each firm so that the estimates of firm-level earnings and values have converged. At smaller firms, the values and earnings are less precisely estimated.

Figure 11a shows that at smaller firms the correlation between values and earnings is lower and thus suggests the importance of measurement error. I sort firms on the basis of firm size and then group firms into 20 bins, with each bin representing the same number of person-years (so that the bins on the left-hand side of the graph have many more firms than the bins on the right-hand side). Within each bin, I compute the correlation between the values and earnings. Consistent with the importance of measurement error, the R^2 rises rapidly as I consider larger firms and then is flat among firms with between about 50 and $1,800$ workers per year (log size of 4 to 7.5). Inconsistent with the law of large numbers logic, for the biggest firms in the dataset the correlation then starts rising again. The interpretation of this finding is simply that compensating differentials are less prevalent among larger firms than smaller firms.

To address the possibility of measurement error in both the earnings and utility, I group firms by characteristics that predict both earnings and nonpecuniary characteristics, and then ask how related are the group-level averages. A limitation of this approach is that the within-group relationship is open to interpretation: an imperfect relationship might reflect measurement error or the role of nonpecuniary characteristics. As such, I offer bounds motivated by the observation in Figure 11a that the looser relationship at the smaller firms is due to measurement error. See Appendix F for a formal discussion.

To find grouping characteristics, I appeal to three literatures. First, the spatial equilibrium literature (e.g., Roback (1982)) argues that location-level differences in earnings reflect nonpecuniary characteristics in the form of higher house prices or other amenities such as weather. As such, I ask how related are county-level means of earnings and utilities. Second, the inter-industry-wage differential literature (e.g., Krueger and Summers (1988)) argues that there is important industry-level variation in earnings that reflects rents in the labor market. As such, I ask how related are earnings and utilities at the industry level.[37] Finally, I appeal to the firm-size wage differential literature (e.g., Brown and Medoff (1989)) and group firms by size. I use the same 20 bins of firms

[37]The reason to remove county-level means first is that industry composition is a feature of a place (i.e., Beaudry, Green, and Sand (2012)).

sorted on the basis of firm size discussed previously.

Table A9 in Appendix H shows that these groupings together explain about 60% of the variance of firm-level earnings.[38] Unconditionally, about 12% of the variance of earnings is at the county-level, about 50% is at the 4-digit industry level and 3% is based on firm-size categories. These groupings explain a similar share of the variance of the firm-level values.[39]

Table 4 presents the main results of this paper, that about a third (31%) of the variance of firm-level earnings is accounted for by rents, and the remaining two-thirds is accounted for by compensating differentials. To get there, I first extract the signal in each of the grouping variables and then aggregate. Specifically, to extract the signal, I compute the group-level mean of both the earnings and utilities, compute the R^2 between them, and then subtract off these group-level means to have only within-group variation. For the county grouping, about 12% of the variance in earnings is between-county. Consistent with the Roback (1982) model, the between-county variation in earnings is only loosely related to utility: the R^2 is 0.04. Thus, the location grouping says that $0.12 \times 0.04 \approx 0.01$ of the overall variance of firm-level earnings is due to rents and $0.12 \times 0.96 \approx 0.11$ is due to compensating differentials.

The remaining two grouping variables are industry and size. Industries account for about half the variance of earnings. Industry-level variation in earnings is more tightly linked to utility than county-level variation: about a third of the industry-level variation is related to utility. The final grouping variable is firm size. After removing the location- and industry-level means, firm size accounts for 1% share of the variance of earnings, and this variation is very weakly related to firm-level values.

After having removed the signal in the grouping variables, I am left with the within-industry, net-of-location and net-of-firm-size variation in earnings and utility, which includes measurement error, so I present bounds. This component accounts for about 40% of the variance of firm-level earnings. Figure 11b shows how the R^2 varies by firm size bin, after having removed the common component. As expected, because I have extracted much of the signal in the values and earnings, the R^2 is now lower in all size bins and it takes longer for the R^2 to flatten out. This means that we expect measurement error to be more important in this residual component of earnings and utility. An upper bound on the contribution of compensating differentials comes from assuming that there is no measurement error in this component, so the R^2 reflects the true relationship. To get a lower bound, I take the maximum value of the R^2 in the figure (in the largest firm size category). To get my preferred estimate, I take the average of the "flat" portion of the figure where the asymptote suggests that measurement error is not driving estimates.

Even at the lower bound, compensating differentials account for a majority of the firm-level

[38]Table A8 in Appendix H shows that there are many employer-to-employer transitions across these boundaries.

[39]Table A9 in Appendix H also displays these statistics for single-units. There is a concern that geography is ambiguous for multi-units. Because multi-unit status is correlated with industry and size, the single-unit numbers differ for reasons beyond measurement issues.

variation in earnings. These three options result in a range for the contribution of compensating differentials to the variance of firm-level earnings of 56% to 72%, with a preferred estimate of 69%.

Appendix G re-estimates the model on subgroups defined by age and gender. The central finding of the paper is robust within each subgroup—though compensating differentials are much more important in explaining earnings dispersion among older workers than younger workers. This appendix also shows that firm values estimated industry-by-industry or location-by-location are quite similar to the values estimated also using the between-industry and between-location variation.

6 Discussion and implications

6.1 Compensating differentials

This paper finds that compensating differentials explain about 15% of the variance of individual-level earnings in the U.S. economy.[40] While one should interpret this point estimate cautiously given the numerous strong assumptions taken to reach it, the take-away is that compensating differentials are important in explaining the variance of firm-level earnings. The feature of the data that points to this finding is that there are systematic patterns of workers moving to lower-paying firms. Moreover, the intuitive plausibility of this interpretation is supported by the sector-level analysis.

The importance of compensating differentials contrasts with the conventional wisdom summarized by Hornstein, Krusell, and Violante (2011, pg. 2883) that they do "not show too much promise" in explaining earnings dispersion. While the literature estimating willingness to pay for particular amenities has had some successes studying very narrow slices of the labor market—for example, there is evidence that compensating differentials explain the large earnings gaps between private sector and nonprofit lawyers (Goddeeris (1988) and Golden (2014))—it has not found robust evidence that compensating differentials are important in explaining earnings determination in the labor market as a whole.[41]

There are three assumptions typical in the one-amenity-at-a-time compensating differentials literature that might help explain why it has struggled to find evidence for compensating differentials. The first assumption is that, in the language of Section 5, the amenity is a *Rosen* amenity and priced in the labor market rather than a *Mortensen* amenity, which is positively related to earnings and is itself a source of dispersion in job value. Based on the evidence in Pierce (2001), however, some well-studied amenities such as employer benefits (i.e. health insurance) are more

[40] This number comes from the following calculation. According to Table 1, firms account for 22% of the variance of earnings. According to Table 4, 69% of the variance of firm-level earnings is compensating differentials. And $0.22 \times 0.69 = 0.15$.

[41] Bonhomme and Jolivet (2009, pg. 763, note 2) summarize the literature as follows: "the general picture of the literature based on hedonic regressions is rather inconclusive."

like *Mortensen* amenities than *Rosen* amenities.[42] The second assumption is that firms compensate workers for particular amenities through pay, rather than through other amenities. The fact that this literature has studied more than one amenity suggests that there is scope for this assumption not to hold. The third assumption is that the researcher knows what workers value in a firm, and can accurately measure it. A benefit of the revealed preference approach is that it considers the entire bundle of amenities as perceived and valued by workers.

Since this paper identifies compensating differentials through patterns of moving to lower-paying firms, one may wonder why previous literature documenting the large number of voluntary earnings cuts had not been interpreted as evidence of the importance of compensating differentials. The reason is that this paper operates at the firm level rather than the individual level. The intuition of the identification result in Section 5 is that compensating differentials are identified when there is behavior that is not explained by earnings. At the firm level, this definition is reasonable because we expect persistent characteristics of a firm that are valued by all workers to be priced in the market. At the individual level, however, this definition is not reasonable. Some of the behavior that is not explained by earnings might reflect individual-level idiosyncrasies—such as a worker sharing a hobby with a boss—that is unlikely to be taken into account when wages are set. Indeed, this paper allows for individual-level idiosyncrasies through the idiosyncratic utility draw, ι, and choices on the basis of ι are not attributed to compensating differentials.

To illustrate the fact that the literature has been justified in not taking individual-level earnings cuts as evidence of compensating differentials, I compute the relationship between pay and behavior implied by the estimates in Hall and Mueller (2013) and Sullivan and To (2014). Both papers estimate search models similar to the one in this paper using individual-level survey data. And both papers emphasize the intuition that voluntary earnings cuts—or rejecting higher-paying offers—implies that by revealed preference amenities are important.[43] I take the preferred estimates from these papers and compute the implied relationship between behavior and pay.[44] For the Hall and Mueller (2013) estimates, pay can only explain 8% of the variation in utility at the individual level. Sullivan and To (2014) allow for three types. For these types, pay can only explain 12% (type I), 21% (type 2), and 28% (type 3) of the variation in utility at the individual level.[45] Hence, relative to my estimate that variation in pay can explain about 31% of variation in behavior (value) at the firm level, using the individual-level earnings cut intuition would find a larger role for compensating

[42] Dey and Flinn (2005) model the positive correlation between health insurance and wages as the result of a treatment effect of health insurance on productivity.

[43] Becker (2011) and Nunn (2013) also use individual-level survey data, but estimate richer models.

[44] I take the preferred estimates from Hall and Mueller (2013) and Sullivan and To (2014) and simulate the steady state distribution of matches. Within each simulated match, I compute utility and earnings. I then compute the R^2 between utility and earnings (see Appendix I for details).

[45] The respective population proportions are 0.14, 0.50, and 0.36. One important feature of Sullivan and To (2014) is that they use a sample of unmarried men, who never attended college and are 26 or younger. Based on Table A12, this a group for which it appears equilibrium dispersion is likely to be more important in explaining the variance of earnings (i.e., my preferred estimate for this group is that compensating differentials' share is 0.46).

differentials, and this likely reflects idiosyncratic factors that we do not expect to be priced in the labor market.

Taber and Vejlin (2013) is the only other paper I am aware of that estimates the role of compensating differentials using revealed preference at the firm level. Their paper answers a different question related to the role of nonpecuniary characteristics than this paper. The counterfactual calculation in Section 6.4 answers the question: what would the variance of earnings be if we priced the part of nonpecuniary characteristics that is priced in earnings? The counterfactual calculation in Taber and Vejlin (2013) related to the role of nonpecuniary characteristics answers this question: what would the variance of earnings be if people valued only money?[46] The answers to these two questions have no mechanical relationship. In addition, this paper develops a methodology that allows for firm-level estimates of earnings and amenities, while Taber and Vejlin (2013) rely on more aggregated features of the data.

6.2 Rents

This paper provides two new sources of evidence on the importance of rents in the labor market, and frictional models more generally. The first source of evidence is systematic patterns of worker mobility across firms. Systematic patterns of worker mobility are a core prediction of the Burdett and Mortensen (1998) model. In this model, search frictions support an equilibrium with dispersion where some firms offer low levels of utility (labelled pay in the model) and some offer high levels. Workers climb the implied job ladder through employer-to-employer transitions. Despite the centrality of this job ladder to search models, this paper is novel in showing how to uncover the job ladder from worker behavior, rather than assuming that it is indexed by pay or some other observable firm-level characteristic. And by comparing the dispersion in the idiosyncratic draws to the common firm-values, this paper also quantifies the importance of differences between firms in explaining mobility. Specifically, the common component has about equal variance to the idiosyncratic component.

The second source of evidence is that the higher-value firms are also, on average, the higher-paying firms. This source of evidence is related to a long tradition in labor economics that compared industry-level variation in quit rates to industry-level variation in pay. Papers in this tradition have argued that the positive relationship provided evidence that (at least some of) the inter-industry wage structure reflected rents (see, for example, Ulman (1965, Table III) and Krueger and Summers (1988, Table IX)). The measure of firm-level pay I use, which uses the Abowd, Kramarz, and Margolis (1999) decomposition, can be viewed as the modern version of the inter-industry wage

[46] Taber and Vejlin (2013) compare the variance of earnings in two steady states. In the first steady state, workers move on the basis of the estimated total job value, which includes nonpecuniary characteristics. In the second steady state, they "turn off" the nonpecuniary characteristics and workers move on the basis of the estimated earnings in each job—that is, they do not price out the nonpecuniary characteristics for the same reason that I cannot identify the variance of all nonpecuniary characteristics.

differential literature by documenting that there are systematic differences in firm-level pay even after removing person fixed effects. Similarly, the estimated search model in this paper can be viewed as the modern version of a quit rate in documenting systematic patterns in choices.

6.3 Ability of search models to match earnings dispersion

This paper finds that search frictions cannot explain all firm-level earnings dispersion. This message is consistent with Hornstein, Krusell, and Violante (2011). They argue that benchmark search models cannot rationalize the extent of earnings dispersion—measured as the residual in a Mincerian regression—in the labor market. Their observation is that unemployed workers find jobs quickly—which suggests that workers do not face a large amount of dispersion in job value in the offer distribution since otherwise they would wait for a better offer.

By focusing on the behavior of employed workers, rather than unemployed workers, this paper provides a complementary source of evidence to Hornstein, Krusell, and Violante (2011) that rents do not explain all earnings dispersion. The key evidence is systematic patterns of employed workers making employer-to-employer transitions to lower-paying firms. This finding indicates that a large portion of measured earnings dispersion is not treated by workers as reflecting dispersion in value. I interpret this finding as evidence for compensating differentials. In addition, this paper focuses on the firm-level part of the variance of earnings. This part of earnings variation controls for person fixed effects and thus is more likely to reflect frictions than the Mincerian residual.

Hornstein, Krusell, and Violante (2011) argue that search models that explain earnings dispersion typically imply implausibly low values of unemployment. Because utility is only measured up to an additive constant, I cannot compare the value of nonemployment to the average value of a job. A statistic about the value of nonemployment that I can compare to other estimates is the share of offers accepted among the unemployed. Hall and Mueller (2013, pg. 12) report that in their sample of job seekers in New Jersey collecting unemployment insurance, unemployed workers accept 71.5% of offers. I estimate that the nonemployed accept 72.5% of offers, which provides evidence that my estimate of the value of nonemployment is plausible.

6.4 Inequality

This section considers the consequences for earnings inequality of pricing the a_{Rosen} portion of nonpecuniary characteristics (this leaves the $a_{\text{Mortensen}}$ portion unpriced). Eliminating compensating differentials would reduce earnings inequality.

Theoretically, the effect on inequality of pricing out the a_{Rosen} component is ambiguous and depends on the correlation between nonpecuniary characteristics and overall earnings. To see how eliminating compensating differentials could decrease inequality, consider an educator and a miner. They both receive the same utility from their jobs, but the miner receives more monetary compensation than the educator because the miner's job is relatively dangerous, while the educator's job is relatively meaningful. As a result, there is substantial earnings inequality: the miner is highly paid, and the Educator is poorly paid. Equalizing the pleasantness of their work leads them to both be equally well-paid, so earnings inequality—or the variance of earnings—decreases. In contrast, to see how eliminating compensating differentials could increase inequality, consider the professor (of economics) and the miner. Even given their large differences in nonearnings compensation, the professor is still better paid than the miner. Equalizing the pleasantness of their work leads the difference in their earnings to be more dramatic and increases earnings inequality.

To measure the effect on inequality of pricing out a_{Rosen}, I use the identification result in section 5.1. Capturing the dispersion in utility, V^e takes into account the good and bad amenities reflected in a_{rosen}. Pricing out a_{Rosen} means that the variance of utility in log dollar units is given by the lower bound in result 3. Hence, I replace the firm-level earnings, Ψ, with the firm-level pure rent component, which is proportional to V^e.[47] Then I recompute the variance of individual-level earnings. To illustrate what drives the counterfactual, I do this in stages, where I sequentially price out location, industry, etc.

At my central estimate of the division of firm-level earnings into rents and compensating differentials (see Table 4), Table 5 shows that removing amenities and compensating workers would reduce inequality. About half of the effect occurs at the industry level.

This counterfactual has surprising impacts on the structure of earnings. Figure 12 plots the actual distribution of earnings and the counterfactual distribution at my central estimate. Inequality is reduced in the counterfactual relative to the data primarily by shifting in the lower tail of the distribution. This is not what we would expect from the patterns of sorting workers to firms. Figure 12b shows a naive counterfactual. To compute this counterfactual, I multiply the firm effects by $1-R^2$ and recompute the variance of earnings. Relative to the data, the naive counterfactual shifts in both the lower and the upper tail of the income distribution. As documented in the last row of Panel A of Table 5, this has a much larger effect on the variance of earnings than what I estimate. The reason for this surprise is that, as shown in Panel B of Table 5, while the earnings potential of workers is positively correlated with rents, it is negatively correlated with a_{Rosen} amenities.

7 Conclusion

This paper exploits the intuition that workers move towards firms with higher value to estimate the revealed value of essentially every firm in the U.S. economy. It then combines this value with earnings data to measure the relative role of rents and compensating differentials in explaining why some firms pay so much and some so little. The paper finds evidence for rents to the extent that workers move towards higher-paying firms, while it finds evidence for compensating differentials to

[47]Specifically, I normalize the variance of V^e so that it is in log dollar units. Let $Var(\Psi)$ be the variance of the firm effects in earnings, $Var(V^e)$ be the variance of values, and $R^2 = Corr(\Psi, V^e)^2$. Then the constant is $\sqrt{\frac{Var(\Psi)}{Var(V^e)}R^2}$.

the extent that workers move towards lower-paying firms. Overall, the paper finds novel evidence for both compensating differentials and rents, but compensating differentials are more important than rents in explaining the variance in firm-level pay. The rankings of sectors are plausible, as is the implied distribution of compensating differentials.

This paper demonstrates that there is much to be learned from studying the "quantity" dimension—how people do and do not flow across firms—of matched employer-employee data. For example, in Sorkin (2015) I explore the gender earnings gap: namely, over 20% of the gender earnings gap is due to men being in higher-paying firms (and industries) than women. Similarly, the findings in Appendix G suggests that there are interesting differences by age. The firm-level value that summarizes the central tendency of worker flows also suggests a new way of studying the direction of reallocation in the labor market, a question which has recently received attention because of the work of Moscarini and Postel-Vinay (2013). More broadly, the core economic and computational insight of this paper could be fleshed out in different product-demand-type contexts. Specifically, this paper emphasizes that there is lots of identifying information in switching behavior and has developed a technique that is computationally feasible when there are a large number of options. This could be applied in a variety of other settings—for example, to study the value of locations and products (see also Bils (2009)).

References

Abowd, John M., Robert H. Creecy, and Francis Kramarz. 2002. "Computing Person and Firm Fixed Effects Using Linked Longitudinal Employer-Employee Data." Working Paper TP-2002-06, US Census Bureau: Longitudinal Employer Household Dynamics.

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica* 67 (2):251–333.

Abowd, John M., Paul Lengermann, and Kevin L. McKinney. 2003. "The Measurement of Human Capital in the U.S. Economy." Working paper.

Abowd, John M. and Ian M. Schmutte. 2014. "Endogenous Mobility." Working paper.

Abowd, John M., Bryce E. Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L. McKinney, Marc Roemer, and Simon Woodcock. 2009. "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators." In *Producer Dynamics: New Evidence from Micro Data*, edited by Timothy Dunne, J. Bradford Jensen, and Mark J. Roberts. University of Chicago Press, 149–230.

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. 2012. "The Network Origins of Aggregate Fluctuations." *Econometrica* 80 (5):1977–2016.

Aizawa, Naoki and Hanming Fang. 2015. "Equilibrium Labor Market Search and Health Insurance Reform." Working paper.

Andersson, Fredrik, Elizabeth E. Davis, Matthew L. Freedman, Julia I. Lane, Brian P. McCall, and Kristin Sandusky. 2012. "Decomposing the Sources of Earnings Inequality: Assessing the Role of Reallocation." *Industrial Relations* 51 (4):779–810.

Arcidiacono, Peter and Paul B. Ellickson. 2011. "Practical Methods for the Estimation of Dynamic Discrete Choice Models." *Annual Review of Economics* 3:363–394.

Artuc, Erhan, Shubham Chaudhuri, and John McLaren. 2010. "Trade Shocks and Labor Adjustment: A Structural Empirical Approach." *American Economic Review* 100 (3):1008–1045.

Bagger, Jesper, Francois Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin. 2014. "Tenure, Experience, Human Capital and Wages: A Tractable Equilibrium Search Model of Wage Dynamics." *American Economic Review* 104 (6):1551–1596.

Bagger, Jesper and Rasmus Lentz. 2015. "An Empirical Model of Wage Dispersion with Sorting." Working paper.

Ballester, Coralio, Antoni Calvo-Armengol, and Yves Zenou. 2006. "Who's Who in Networks. Wanted: The Key Player." *Econometrica* 74 (5):1403–1417.

Barth, Erling, Alex Bryson, James C. Davis, and Richard Freeman. 2014. "It's Where You Work: Increases in Earnings Dispersion Across Establishments and Individuals in the U.S." Working Paper 20447, NBER.

Beaudry, Paul, David E. Green, and Benjamin Sand. 2012. "Does Industrial Composition Matter for Wages? A Test of Search and Bargaining Theory." *Econometrica* 80 (3):1063–1104.

Becker, Dan. 2011. "Non-Wage Job Characteristics and the Case of the Missing Margin." Working paper.

Benedetto, Gary, John Haltiwanger, Julia Lane, and Kevin Mckinney. 2007. "Using Worker Flows to Measure Firm Dynamics." *Journal of Business and Economic Statistics* 25 (3):299–313.

Bils, Mark. 2009. "Do Higher Prices For New Goods Reflect Quality Growth or Inflation?" *Quarterly Journal of Economics* 124 (2):637–675.

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

BLS. 1997. *U.S. Bureau of Labor Statistics Handbook of Methods.*

Bonacich, Phillip. 1987. "Power and Centrality: A Family of Measures." *American Journal of Sociology* 92 (5):1170–1182.

Bonhomme, Stephane and Gregory Jolivet. 2009. "The Pervasive Absence of Compensating Differentials." *Journal of Applied Econometrics* 24:763–795.

Bowlus, Audra and Lars Vilhuber. 2002. "Displaced workers, early leavers, and re-employment wages." Work in progress.

Brown, Charles and James Medoff. 1989. "The Employer Size-Wage Effect." *Journal of Political Economy* 97 (5):1027–1059.

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

Burgess, Simon, Julia Lane, and David Stevens. 2000. "Job Flows, Worker Flows, and Churning." *Journal of Labor Economics* 18 (3):473–502.

Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin. 2006. "Wage Bargaining with On-the-Job Search: Theory and Evidence." *Econometrica* 74 (2):323–364.

Card, David, Ana Rute Cardoso, and Patrick Kline. 2014. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." Working paper.

Card, David, Joerg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics* 128 (3):967–1015.

Card, David, Jorg Heining, and Patrick Kline. 2012. "Workplace Heterogeneity and the rise of West German wage inequality." Working Paper 18522, NBER.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104 (9):2593–2632.

Davis, Steven J., R. Jason Faberman, and John Haltiwanger. 2012. "Labor Market Flows in the Cross Section and Over Time." *Journal of Monetary Economics* 59 (2012):1–18.

Dey, Matthew S. and Christopher J. Flinn. 2005. "An Equilibrium Model of Health Insurance Provision and Wage Determination." *Econometrica* 73 (2):571–627.

Eeckhout, Jan and Philipp Kircher. 2011. "Identifying Sorting—In Theory." *Review of Economic Studies* 78 (3):872–906.

Elliott, Matthew and Benjamin Golub. 2014. "A Network Approach To Public Goods." Working paper.

Flaaen, Aaron, Matthew D. Shapiro, and Isaac Sorkin. 2015. "Reconsidering the Consequences of Worker Displacements: Survey versus Administrative Measurements." Work in progress.

Fox, Jeremy T. 2010. "Estimating the Employer Switching Costs and Wage Responses of Forward-Looking Engineers." *Journal of Labor Economics* 28 (2):357–412.

Goddeeris, John H. 1988. "Compensating Differentials and Self-Selection: An Application to Lawyers." *Journal of Political Economy* 96 (2):411–428.

Golden, Joseph M. 2014. "The Importance of Preferences in Sectoral Sorting: Direct Evidence from Lawyers in the Private and Nonprofit Sectors." Working paper.

Goldschmidt, Deborah and Johannes F. Schmieder. 2015. "The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure." Working paper.

Gronberg, Timothy J. and W. Robert Reed. 1994. "Estimating Workers' Marginal Willingness to Pay for Job Attributes Using Duration Data." *Journal of Human Resources* 29 (3):911–931.

Hagedorn, Marcus, Tzuo Hann Law, and Iourri Manovskii. 2014. "Identifying Equilibrium Models of Labor Market Sorting." Working paper.

Hall, Robert E. and Andreas Mueller. 2013. "Wage Dispersion and Search Behavior." Working paper.

Haltiwanger, John, Henry R. Hyatt, and Erika M. McEntarfer. 2015. "Cyclical Reallocation of Workers Across Employers by Firm Size and Firm Wage." Working Paper 21235, NBER.

Hornstein, Andreas, Per Krusell, and Giovanni L. Violante. 2011. "Frictional Wage Dispersion in Search Models: A Quantitative Assessment." *American Economic Review* 101:2873–2898.

Hotz, V. Joseph and Robert A. Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *Review of Economic Studies* 60 (3):497–529.

Hwang, Hae-Shin, Dale T. Mortensen, and W. Robert Reed. 1998. "Hedonic Wages and Labor Market Search." *Journal of Labor Economics* 16 (4):815–847.

Hyatt, Henry R. and Erika McEntarfer. 2012. "Job-to-Job Flows and the Business Cycle." Working paper.

Hyatt, Henry R., Erika McEntarfer, Kevin McKinney, Stephen Tibbets, and Doug Walton. 2014. "Job-to-Job Flows (JTJ) Flows: New Labor Market Statistics From Linked Employer-Employee Data ." Working Paper CES 14-34.

Jacobson, Louis, Robert LaLonde, and Daniel Sullivan. 1993. "Earnings Losses of Displaced Workers." *American Economic Review* 83 (4):685–709.

Jarosch, Gregor. 2014. "Searching for Job Security and the Consequences of Job Loss ." Working paper.

Jolivet, Gregory, Fabien Postel-Vinay, and Jean-Marc Robin. 2006. "The empirical content of the job search model: Labor mobility and wage distributions in Europe and the US." *European Economic Review* 50:877–907.

Kahn, Lisa B. and Erika M. McEntarfer. 2014. "The Cyclicality of Employment and Firm Quality." Working paper.

Kline, Patrick. 2008. "Understanding sectoral labor market dynamics: An equilibrium analysis of the oil and gas field services industry." Working paper.

Kornfeld, Robert and Howard S. Bloom. 1999. "Measuring Program Impacts on Earnings and Employment: Do Unemployment Insurance Wage Reports from Employers Agree with Surveys of Individuals." *Journal of Labor Economics* 17 (1):168–197.

Krolikowski, Pawel. 2014. "Job Ladders and Earnings of Displaced Workers." Working paper.

Krueger, Alan B. and Lawrence H. Summers. 1988. "Efficiency Wages and the Inter-Industry Wage Structure." *Econometrica* 56 (2):259–293.

Lang, Kevin and Sumon Majumdar. 2004. "The Pricing of Job Characteristics When Markets Do Not Clear: Theory and Policy Implications." *International Economic Review* 45 (4):1111–1128.

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

Marimon, Ramon and Fabrizio Zilibotti. 1999. "Unemployment vs. Mismatch of Talents: Reconsidering Unemployment Benefits." *Economic Journal* 455:266–291.

Minc, Henryk. 1988. *Nonnegative matrics.* Wiley.

Mortensen, Dale T. 2003. *Wage Dispersion: Why are similar workers paid differently?* MIT Press.

Mortensen, Dale T. and Christopher A. Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies* 61 (3):397–415.

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

———. 2014. "Did the Job Ladder Fail After the Great Recession?" Working paper.

Nunn, Ryan. 2013. "Match Quality with Unpriced Amenities." Working paper.

Page, Lawrence, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1998. "The PageRank Citation Ranking: Bringing Order to the Web." Working paper, Stanford. URL `http://ilpubs.stanford.edu:8090/422/`.

Palacios-Huerta, Ignacio and Oscar Volij. 2004. "The Measurement of Intellectual Influence." *Econometrica* 72 (3):963–977.

Papp, Tamas K. 2013. "Frictional Wage Dispersion with Bertrand Competition: An Assessment." *Review of Economic Dynamics* 16:540–552.

Pierce, Brooks. 2001. "Compensation Inequality." *Quarterly Journal of Economics* 116 (4):1493–1525.

Pinheiro, Roberto and Ludo Visschers. 2015. "Unemployment risk and wage differentials." *Journal of Economic Theory* 157:397–424.

Pinski, Gabriel and Francis Narin. 1976. "Citation Influence for Journal Aggregates of Scientific Publications: Theory, with Application to the Literature of Physics." *Information Processing and Management* 12:297–312.

Postel-Vinay, Fabien and Jean-Marc Robin. 2002. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica* 70 (6):2295–2350.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90 (6):1257–1278.

Rosen, Sherwin. 1986. "The Theory of Equalizing Differences." In *Handbook of Labor Economics*, vol. 1, edited by Orley Ashenfelter and Richard Layard. Elsevier, 641–692.

Schmutte, Ian M. 2015. "Job Referral Networks and the Determination of Earnings in Local Labor Markets." *Journal of Labor Economics* 33 (1):1–32.

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter. 2015. "Firming Up Inequality." Working paper.

Sorkin, Isaac. 2015. "Why do men and women work at different firms?" Work in progress.

Sullivan, Paul and Ted To. 2014. "Search and Non-Wage Job Characteristics." *Journal of Human Resources* 49 (2):472–507.

Taber, Christopher and Rune Vejlin. 2013. "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market." Working paper.

Tjaden, Volker and Felix Wellschmied. 2014. "Quantifying the Contribution of Search to Wage Inequality." *American Economic Journal: Macroeconomics* 6 (1):134–161.

Topel, Robert H. and Michael P. Ward. 1992. "Job Mobility and the Careers of Young Men." *Quarterly Journal of Economics* 107 (2):439–479.

Trefethen, Lloyd N. and David Bau. 1997. *Numerical Linear Algebra*. SIAM Press.

Ulman, Lloyd. 1965. "Labor Mobility and the Industrial Wage Structure in the Postwar United States." *Quarterly Journal of Economics* 79 (1):73–97.

Table 1: Summary statistics and the variance of earnings

	All (1)	Connected by EE and ENE (2)	S. Connected by EE (3)	S. Connected by EE (restrictions) (4)
Sample size				
People-years	$504,945,000$	$500,584,000$	$470,387,000$	$456,746,000$
People	$105,921,000$	$104,778,000$	$100,547,000$	98,821,000
Employers	$6,155,000$	$5,258,000$	$1,971,000$	$1,528,000$
Summary statistics				
Mean log earnings	10.43	10.43	10.45	10.45
Variance of log earnings	0.70	0.70	0.69	0.68
Share of variance of earnings explained by each parameter set				
Employers	N/A	0.23	0.22	N/A
People	N/A	0.54	0.55	N/A
Covariates	N/A	0.11	0.11	N/A
Overall fit of AKM decomposition				
R^2	N/A	0.88	0.88	N/A
Adj. R^2	N/A	0.85	0.85	N/A

Notes: Sample counts are rounded to the nearest thousand. The data is at an annual frequency. There is one observation per person per year. The observation is the job from which a person made the most money, but only if she made at least $3,250 (in $2011). Earnings are annualized. The table includes person-years in which on December 31 of the year the person was aged 18-61 (inclusive). The extra restrictions in the final column are that an employer have non-missing industry information, hire a worker on an exogenous employer-to-employer transition, and hire a worker from nonemployment. AKM is the Abowd, Kramarz, and Margolis (1999) decomposition; see equation (14). EE is employer-to-employer and ENE is employer-to-nonemployment-to-employer.

Table 2: Transition probabilities and model parameters

Symbol	Description	Value
	Overall employer-to-employer Transition Probability	0.10
	Overall employer-to-nonemployment Transition Probability	0.15
	employer-to-employer Share of Transitions	0.40
	Pr(displacement | employer-to-employer)	0.26
	Pr(displacement | employer-to-nonemployment)	0.34
	Pr(displacement | employer-to-employer & contracting)	0.48
	Pr(displacement | employer-to-nonemployment & contracting)	0.56
δ	Exogenous employer-to-nonemployment probability	0.05
ρ	Exogenous employer-to-employer probability	0.03
λ_1	Probability of offer on-the-job	0.20

Notes: All probabilities and parameters are annual. The sample for the transition probabilities is column (1) of Table 1. A worker only counts as separating if she appears again in the dataset. The sample for estimating λ_1 and below is column (4) of Table 1. The ρ is related to the calculated probability of making an exogenous employer-to-employer transition by $(1-\delta)\rho$. λ_1 is estimated from the model.

Table 3: Earnings cuts are common and correlated at the firm level

$\Pr(y \downarrow)$	All	ENE	EE	EE (weighted)
Unconditional	0.42	0.47	0.38	0.37
When moving to a				
...higher paying firm	0.29	0.33	0.26	0.26
...lower paying firm	0.58	0.62	0.52	0.52

Notes: This table summarizes transitions for a worker who had a different dominant employer in consecutive years. The pay of a firm is defined by its firm effect. EE is employer-to-employer and ENE is employer-to-nonemployment-to-employer. The weight is the endogenous weight on the transition.

Table 4: Why do some firms pay so much and some so little?

	Share of the variance of firm-level earnings	Share of (1) explained by firm-level utility	Share of the variance of firm-level earnings explained by:	
			Rents	Compensating differentials
	(1)	(2)	(3)	(4)
All	1.00	0.25	0.25	0.75
County	0.12	0.04	0.01	0.11
4-digit industry	$(1 - .12) \times 0.54 = 0.47$	0.45	0.21	0.26
Size	$(1 - .12 - .47) \times 0.04 = 0.01$	0.02	0.00	0.01
Resid. (v1: upper)	0.39	0.15	0.06	0.33
Resid. (v2: preferred)	(see previous row)	0.24	0.09	0.30
Resid. (v3: lower)	(see previous row)	0.55	0.22	0.18
Total (v1)			0.28	0.72
Total (v2)			0.31	0.69
Total (v3)			0.44	0.56

Notes: This table uses firm effects estimated in the set of firms strongly connected by employer-to-employer transitions. Column (1) reports an order-dependent decomposition of the firm effects into a county, industry, size, and residual component. The first row reports the share of the variance of the firm effects that is explained by a set of county fixed effects. The second row removes county-means from the firm fixed effects and reports the share of the variance of firm fixed effects net-of-county-means that is explained by a set of 4-digit industry fixed effects. The third row removes the industry fixed effects and reports the share of the variance accounted for by firm size categories. The next three rows report the decomposition of the remaining share of the variance. Column (2) reports the R^2 on the firm-level wage against the firm-level utility at each level of aggregation. Version 1 is the actual relationship on the residual; version 2 assumes that the relationship is the same as on the flat part of Figure 11b; and version 3 assumes that it is the maximum value in Figure 11b. Column (3) = Column (1) $\times$ Column (2) and is the share of the variance of firm-level earnings that is related to utility. Column (4) = Column (1) $\times$ (1- Column (2)) and is the share of the variance of firm-level earnings that is unrelated to utility. The total rows sum up the four components. The total (v1) and the all correlations differ because of cross-terms. Calculations may not exactly reproduce because of rounding.

Table 5: Implications for inequality

Panel A. Counterfactuals

Scenario	Variance of earnings (1)	Change relative to data (2)
Data	0.68	N/A
Price location	0.66	-3%
...and industry	0.62	-9%
...and size	0.61	-10%
...and scaled residual	0.59	-14%
Remove firm effects	0.48	-30%
"Naive"	0.53	-23%

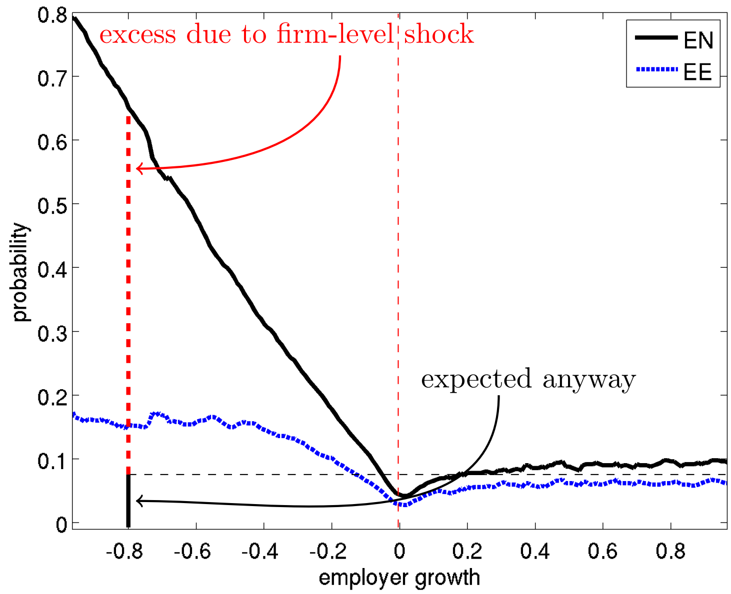
Panel B. Correlations

	Earnings minus firm effects	Rents	Amenities
Earnings minus firm effects	1	0.29	-0.07
Rents		1	0.08
Amenities			1

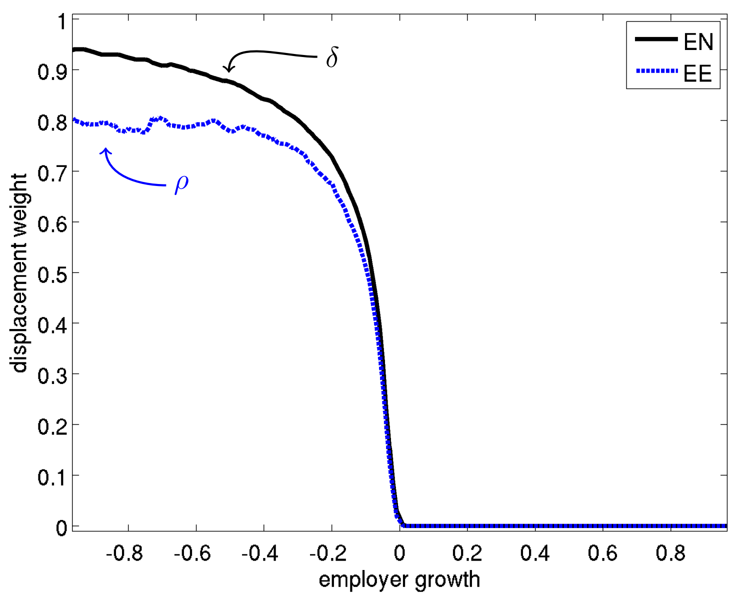
Notes: This table shows how pricing nonearnings characteristics affects the variance of earnings. Panel A shows the variance of earnings in the data and in four counterfactuals. The first counterfactual equalizes the nonearnings aspects of locations. The subsequent rows show the effect of sequentially equalizing the nonearnings aspects by industry, firm size, and the residual (the remaining variance in firm-level earnings net of location, industry, and size). Column (1) reports the variance of log annualized earnings. Column (2) reports the percent change relative to the data. Panel B shows the relationship between earnings minus firm effects (i.e., $y_{w,t} - \Psi_{\boldsymbol{J}(w,t)}$), rents (the scaled version of the V^e), and amenities (the gap between the scaled version of V^e and Ψ). The sample is the person-years in column (4) of Table 1.

Figure 1: Constructing exogenous rates

(a) Quarterly probabilities

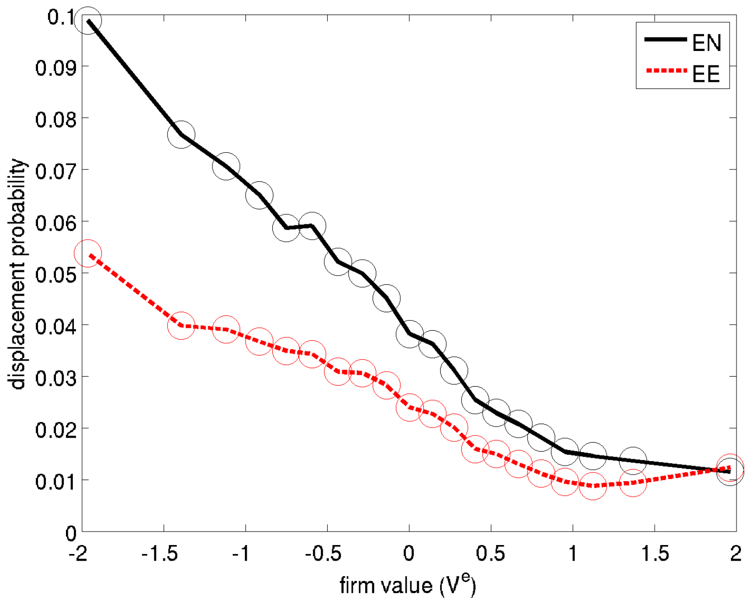


(b) Exogenous weights



Notes: These figures shows the data used to construct the exogenous weights for employer-to-employer (EE) and employer-to-nonemployer (EN) transitions. The probabilities and growth rates are quarterly. The probabilities are computed in one percentage point wide bins of employer growth rates. The figure plots a five-bin moving average. The exogenous weight is $\frac{\text{excess}}{\text{excess}+\text{expected}}$. At an expanding employer, the exogenous weight is zero by construction.

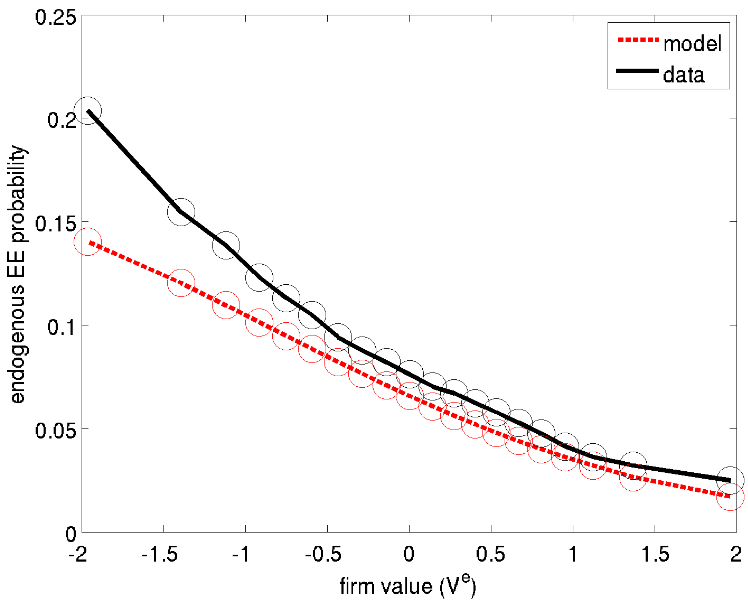
Figure 2: The "slippery ladder": exogenous shocks by firm value



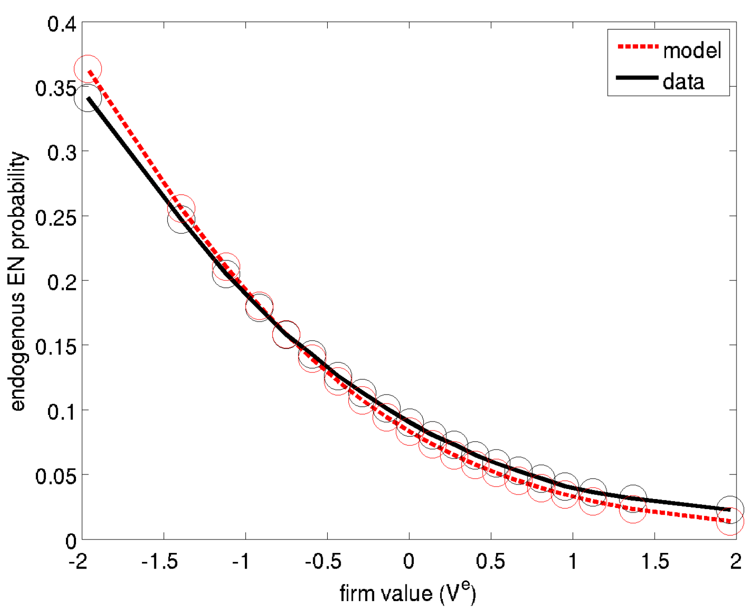
Notes: This figure sorts firms into 20 bins on the basis of firm-value. Within each bin, I compute the average probability of each kind of exogenous separation shock by summing across the exogenous weights, which are constructed using the variation depicted in Figure 1. EE is employer-to-employer and EN is employer-to-nonemployment.

Figure 3: Endogenous employer-to-employer and employer-to-nonemployment probabilities by firm value

(a) Employer-to-employer

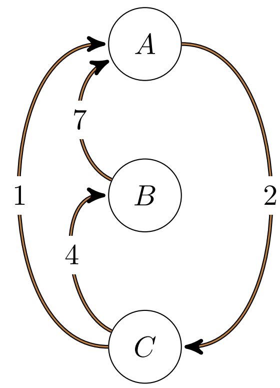


(b) Employer-to-nonemployment



Notes: This figure sorts firms into 20 bins on the basis of firm-value. For each firm, I compute the model-implied probability of an endogenous employer-to-employer and employer-to-nonemployment transition. I then take the person-year-weighted average of the model and data within each bin.

Figure 4: The model returns a different answer than a binary comparison



$$M = \begin{bmatrix} 0 & 7 & 1 \\ 0 & 0 & 4 \\ 2 & 0 & 0 \end{bmatrix}$$

$$S = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

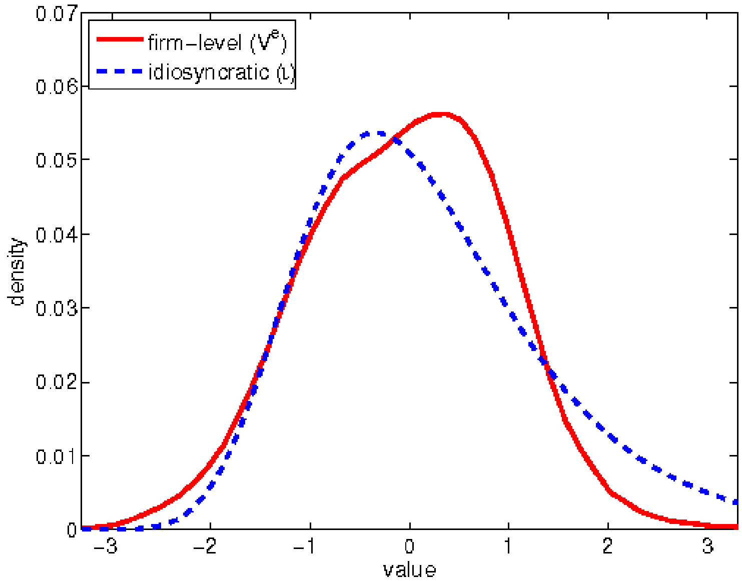
$$S^{-1}M = \begin{bmatrix} 0 & \frac{7}{2} & \frac{1}{2} \\ 0 & 0 & \frac{4}{7} \\ \frac{2}{5} & 0 & 0 \end{bmatrix}$$

$$exp(\tilde{V}) = \begin{bmatrix} 0.614 \\ 0.140 \\ 0.246 \end{bmatrix}$$

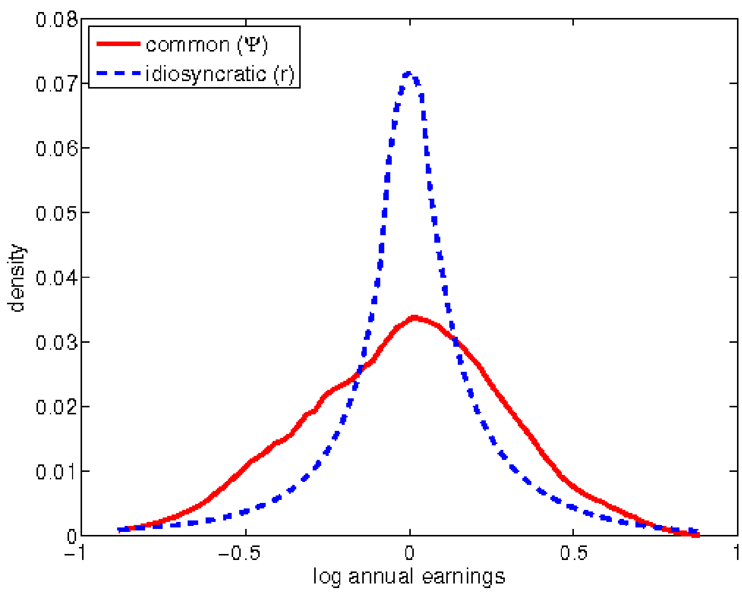
Notes: Given M, the matrix S is defined as follows: entry $S_{ii} = \sum_j M_{ji}$ (the i^{th} column sum), while the off-diagonal entries are zero. Hence $S^{-1}M$ divides the i^{th} row of M by the i^{th} column sum of M (this means that $S^{-1}M$ is not a transition matrix). $exp(\tilde{V})$ is the solution to the following equation: $exp(\tilde{V}) = S^{-1}Mexp(\tilde{V})$.

Figure 5: Dispersion in the labor market

(a) Values



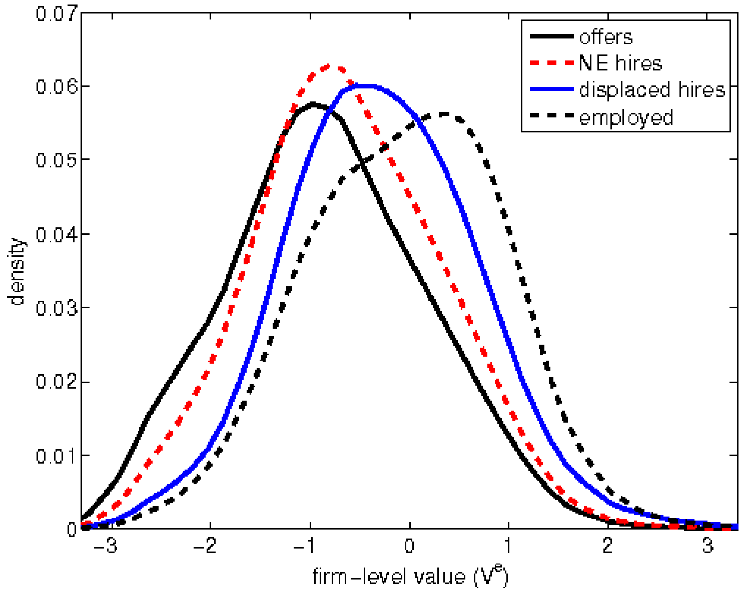
(b) Earnings



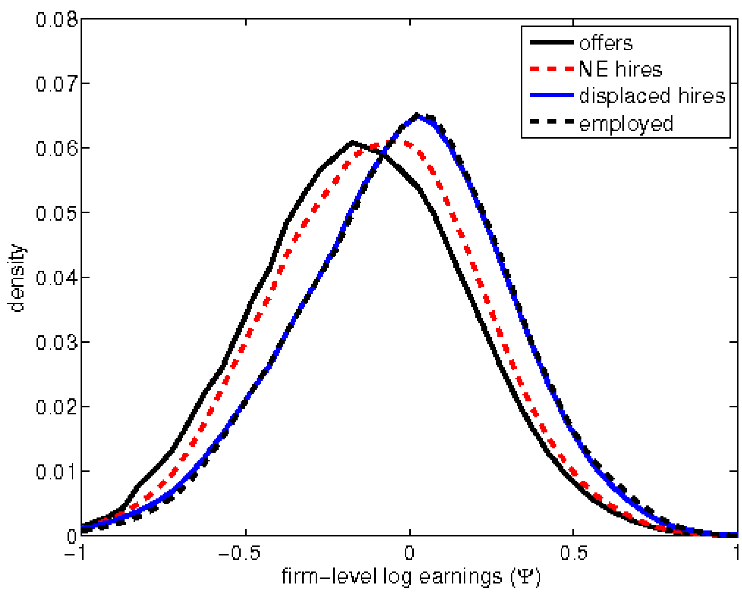
Notes: Both distributions are normalized so that the median value is zero.

Figure 6: Distributions

(a) Values



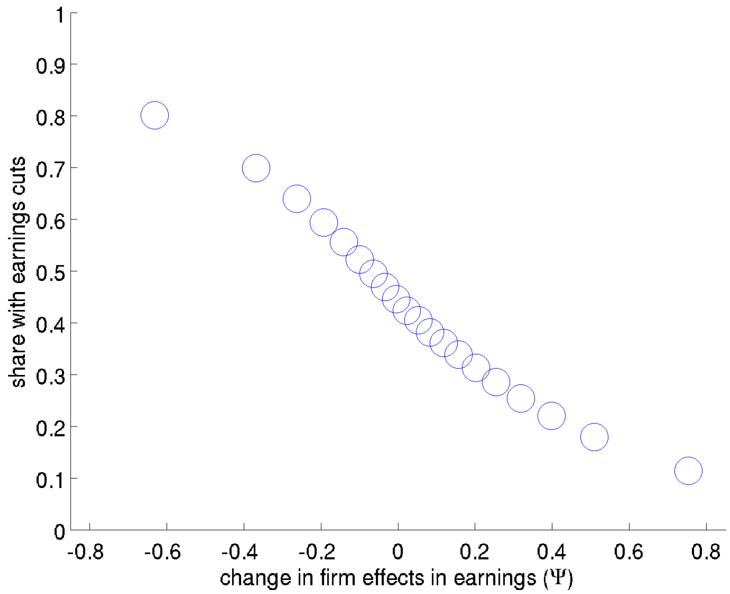
(b) Earnings



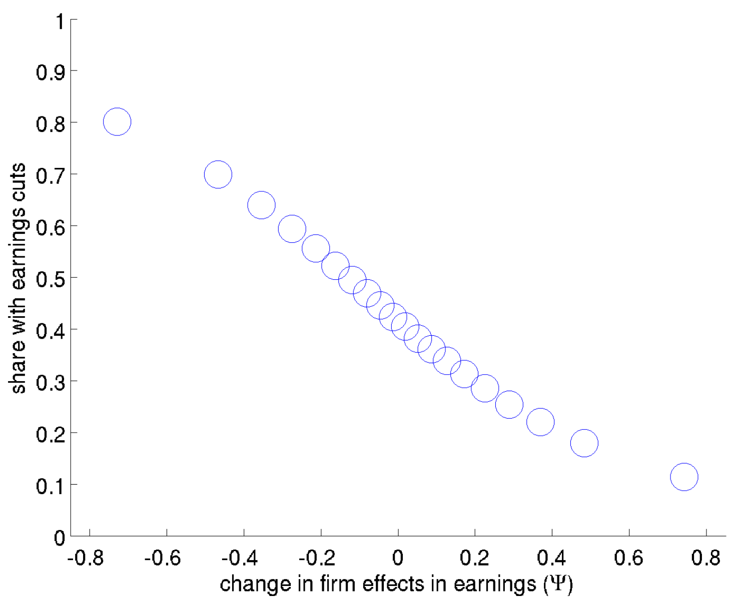
Notes: This figure plots the dispersion in the firm-level values (top panel) and earnings (bottom panel) in four distributions: the offer distribution, the distribution of where hires from nonemployment accept offers, the distribution of where displaced workers making employer-to-employer transitions accept offers, and finally the distribution among the employed workers.

Figure 7: Change in firm pay related to probability of an earnings cut

(a) Employer-to-employer



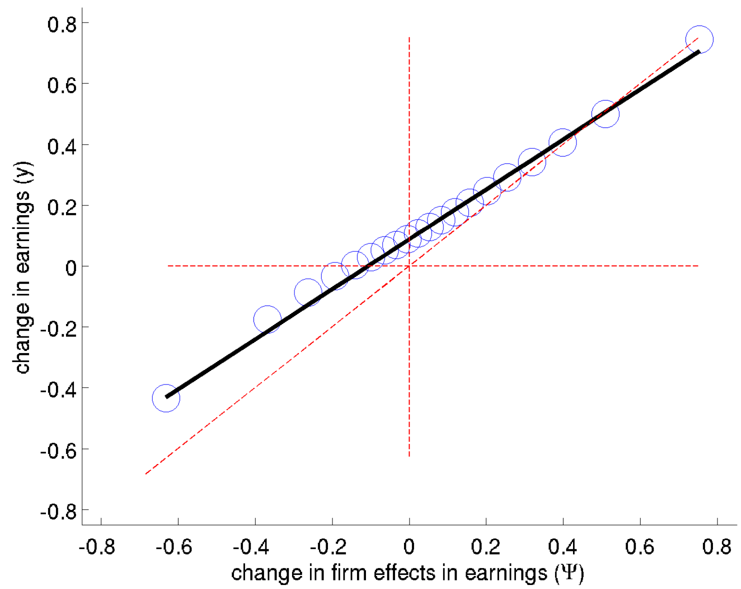
(b) Employer-to-nonemployment-to-employer



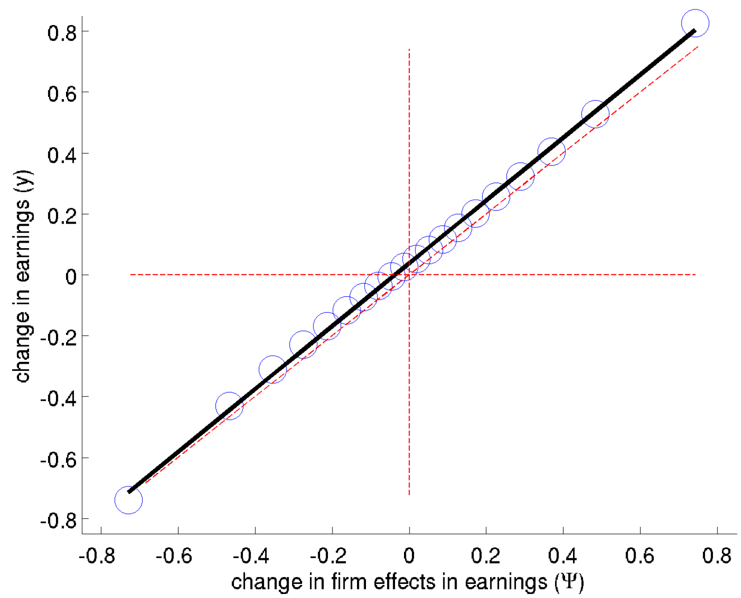
Notes: This figure considers the sample of workers who switch annual dominant jobs between consecutive years. The earnings considered are the earnings in the last year at the previous job and the earnings in the first year at the new job.

Figure 8: Change in firm pay related to magnitude of earnings change

(a) Employer-to-employer



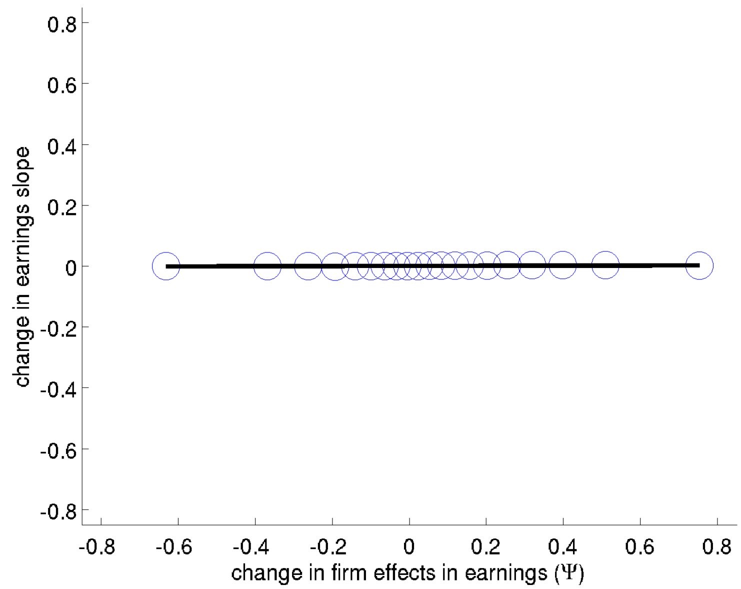
(b) Employer-to-nonemployment-to-employer



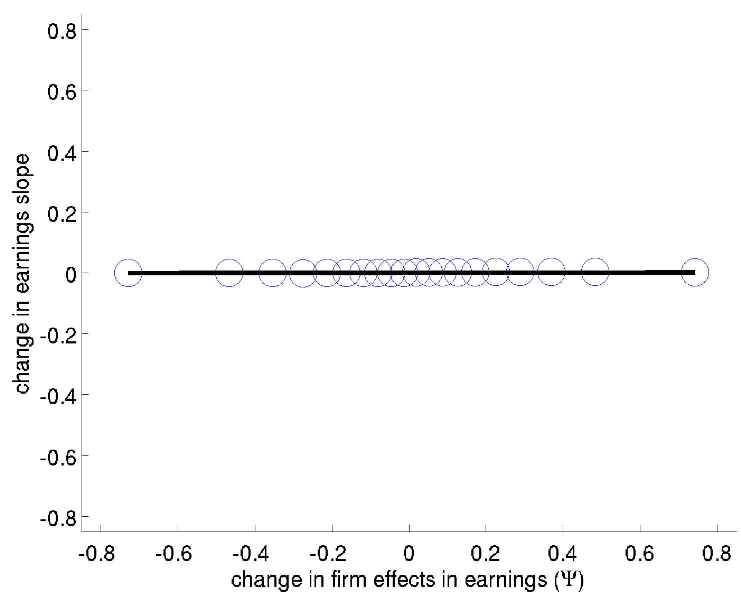
Notes: This figure considers the sample of workers who switch annual dominant jobs between consecutive years. I have sorted the job changers into 20 bins on the basis of the change in the firm effects. The circles plot the bin means. The solid line plots the best-fitting line estimated based on the micro-data. The dashed red line plots the 45 degree line. The coefficient in the upper panel is 0.82, and the coefficient in the bottom panel is 1.03 (the standard errors are essentially zero).

Figure 9: Change in firm pay is not related to change in slope of earnings

(a) Employer-to-employer



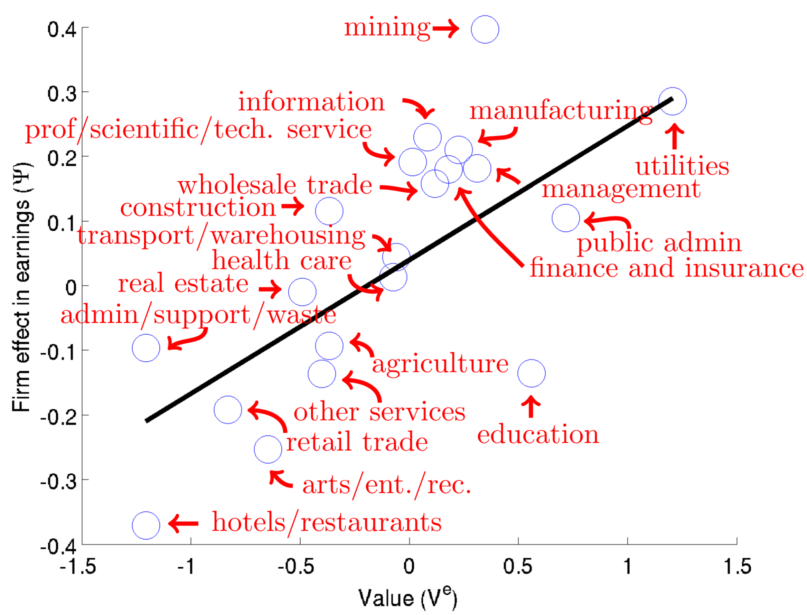
(b) Employer-to-nonemployment-to-employer



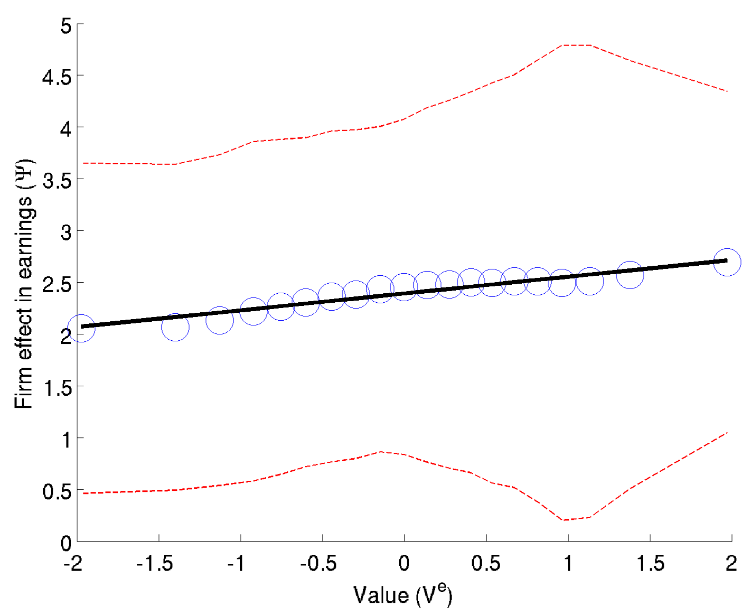
Notes: This figure considers the sample of workers who switch annual dominant jobs between consecutive years. I have sorted the job changers into 20 bins on the basis of the change in the firm effects. The circles plot the bin means. The slope of firm-level pay is estimated using the earnings changes of the stayers. The solid line plots the best-fitting line estimated based on the micro-data.

Figure 10: Relationship between values and earnings

(a) Sector level



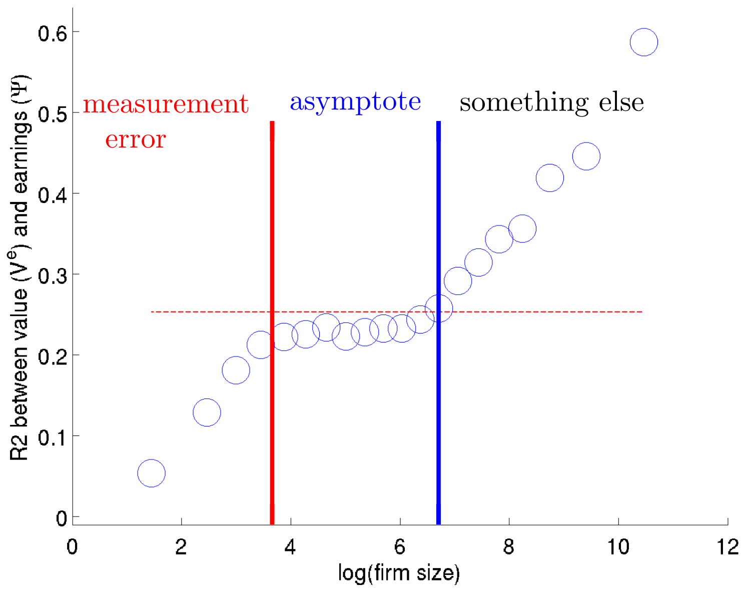
(b) Overall



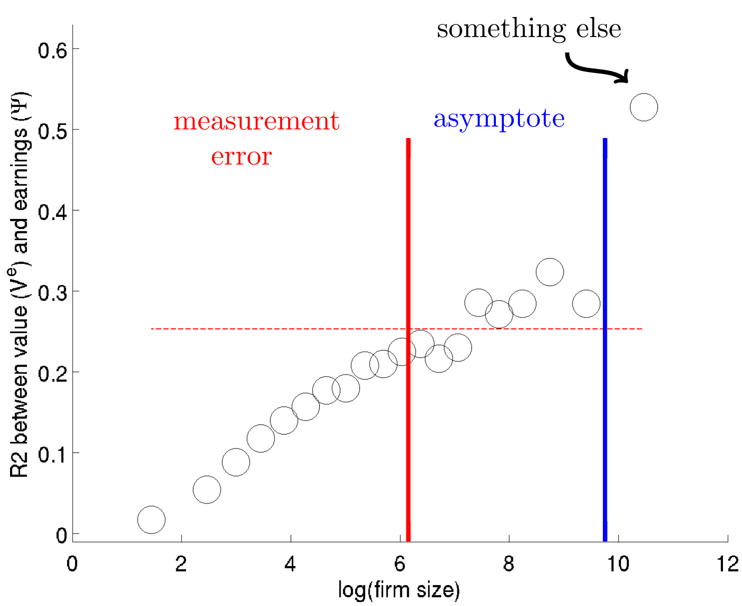
Notes: The top panel of this figure plots the sector-level means of the earnings and values. The solid black line plots the regression line run at the sector level and weighting by the number of person-years represented by each sector. The R^2 is 0.45. The bottom panel sorts firms on the basis of firm-level values. The circles plot 20 bins with the same number of person-years, while the solid line plots the regression line estimated on the firm-level data. The red line shows plus and minus one standard deviation of the firm-level earnings within each value bin. The R^2 is 0.25.

Figure 11: Relationship between R^2 and firm size

(a) All



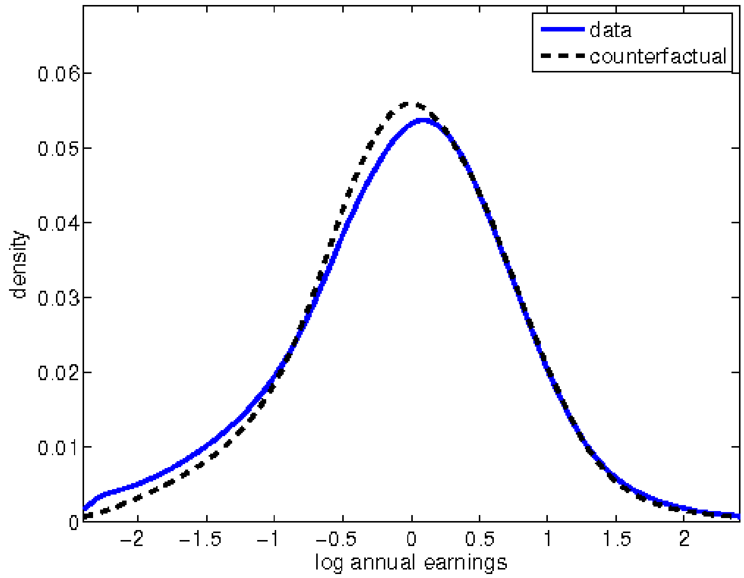
(b) Residual



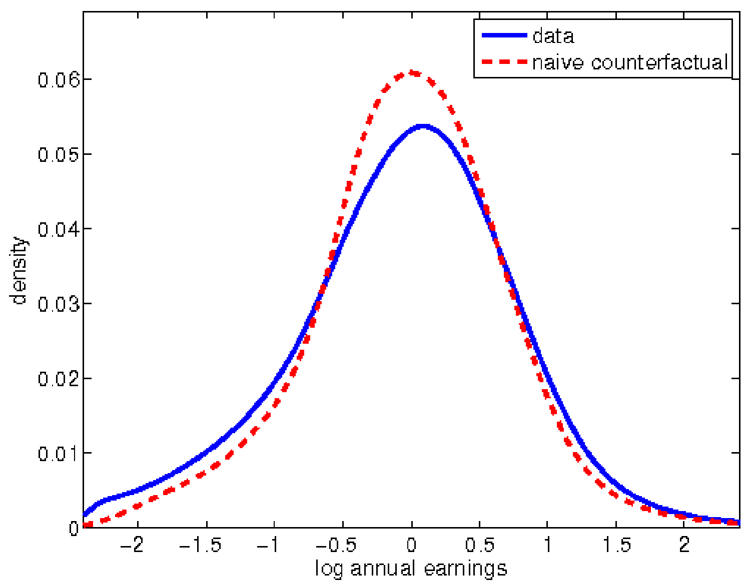
Notes: This figure sorts firms on the basis of firm size into 20 bins with the same number of person-years. The upper panel plots the R^2 between the "raw" firm-level values and earnings computed bin by bin. The bottom panel plots the R^2 on the residual firm-level values and earnings, where I have removed the county, industry and size means from the earnings and the values.

Figure 12: Counterfactual inequality

(a) Counterfactual



(b) "Naive" counterfactual



Notes: The top panel of this figure plots the distribution of income in the data and in a counterfactual where I use my estimates of the firm-level values to price out compensating differentials. The bottom panel considers a "naive" counterfactual where I deflate all the firm components of earnings by a constant fraction and then recompute income.

A Appendix: Omitted Proofs

Proof of Result 1

Notational/definitional preliminaries: This follows the presentation in Minc (1988) of standard graph theory definitions. Let M be a matrix, where entry M_{ij} measures flows from employer j to employer i. Note that all entries in M are by construction nonnegative: the entries are either zeros, or positive values. Let $\mathcal{E}$ be a set (of employers) labelled from $1...n$. Let A be a set of ordered pairs of elements of $\mathcal{E}$. The pair $D = (\mathcal{E}, A)$ is a *directed* graph. $\mathcal{E}$ is the set of *vertices*, and the elements of A are the *arcs* of D, which represent *directed* flows between employers. A sequence of arcs $(i, t_1)(t_1, t_2)...(t_{m-2}, t_{m-1})(t_{m-1}, j)$ is a *path connecting* j to i. The *adjacency* matrix of a directed graph is the $(0, 1)$ matrix whose (i, j) entry is 1 if and only if (i, j) is an arc of D. An adjacency matrix is associated with a nonnegative matrix M if it has the same zero pattern as M. The directed graph is *strongly connected* if for any pair of distinct vertices i and j there is a path in D connecting i to j and j to i. The directed graph is *connected* if for any pair of distinct vertices i and j there is a path in D connecting i to j or a path connecting j to i.

Proof. Observe that if M is strongly connected, then every column sum is nonzero so that the adjacency matrix associated with M is the same as the adjacency matrix associated with $S^{-1}M$.

By Minc (1988), chapter 4, theorem 3.2,] a nonnegative matrix is irreducible if and only if the associated directed graph is strongly connected. By Minc (1988), chapter 1, theorem 4.4, an irreducible matrix has exactly one eigenvector in E^n (the simplex). If M represents a set of strongly connected firms then these two theorems (often jointly called the Perron-Frobenius theorem) guarantee the existence of a unique solution of the form:

$$S^{-1}Mexp(\tilde{V}) = \lambda exp(\tilde{V}),$$

where all the entries in $exp(\tilde{V})$ are of the same sign and when we take the positive version, $\lambda > 0$.

All that remains to show is that $\lambda = 1$.

Consider the j^{th} row of $S^{-1}Mexp(\tilde{V}) = \lambda exp(\tilde{V})$. Let e_j be the basis vector; that is, it is a zero vector with 1 in the j^{th} row.

$$[S^{-1}Mexp(\tilde{V})]_j = [\lambda exp(\tilde{V})]_j, \tag{A1}$$

$$\frac{e_j^T Mexp(\tilde{V})}{||Me_j||_1} = \lambda e_j^T exp(\tilde{V}), \tag{A2}$$

where $|| \cdot ||_1$ is the l_1 norm of a matrix so for an arbitrary matrix A we have $||A||_1 = \sum_i \sum_j |a_{ij}|$. Note that $||Me_j||_1$ is a scalar.

Because M is a nonnegative matrix, we can rewrite the l_1 norm as a dot product with a vector of ones. Specifically, let $\mathbf{1}$ be a column vector of 1s:

$$||Me_j||_1 = \mathbf{1}^T Me_j. \tag{A3}$$

Rearrange:

$$\frac{e_j^T Mexp(\tilde{V})}{||Me_j||_1} = \lambda e_j^T exp(\tilde{V}), \tag{A4}$$

$$\frac{e_j^T Mexp(\tilde{V})}{\mathbf{1}^T Me_j} = \lambda e_j^T exp(\tilde{V}), \tag{A5}$$

$$e_j^T Mexp(\tilde{V}) = \lambda \mathbf{1}^T Me_j e_j^T exp(\tilde{V}). \tag{A6}$$

Now sum over the rows:

$$\sum_j e_j^T Mexp(\tilde{V}) = \sum_j \lambda \mathbf{1}^T Me_j e_j^T exp(\tilde{V}), \tag{A7}$$

$$\sum_j e_j^T Mexp(\tilde{V}) = \lambda \sum_j \mathbf{1}^T Me_j e_j^T exp(\tilde{V}), \tag{A8}$$

$$\mathbf{1}^T Mexp(\tilde{V}) = \lambda \sum_j \mathbf{1}^T Me_j e_j^T exp(\tilde{V}), \tag{A9}$$

$$\mathbf{1}^T Mexp(\tilde{V}) = \lambda \mathbf{1}^T M \sum_j e_j e_j^T exp(\tilde{V}), \tag{A10}$$

$$\mathbf{1}^T Mexp(\tilde{V}) = \lambda \mathbf{1}^T Mexp(\tilde{V}). \tag{A11}$$

Hence, $\lambda = 1$. □

Proof of Result 2

Proof. The proof shows that the diagonal elements cancel out. First, use the identity from (11):

$$exp(\tilde{V}_i) \sum_{j' \in \mathcal{E}+n} M_{j'i} = \sum_{j \in \mathcal{E}+n} M_{ij} exp(\tilde{V}_j).$$

Expand to write the diagonal elements explicitly:

$$exp(\tilde{V}_i) M_{ii} + exp(\tilde{V}_i) \sum_{j' \in \mathcal{E}+n \setminus \{i\}} M_{j'i} = \sum_{j \neq i \in \mathcal{E}+n \setminus \{i\}} M_{ij} exp(\tilde{V}_j) + exp(\tilde{V}_i) M_{ii}.$$

Then cancel the diagonal terms to show that (11) holds with arbitrary diagonal elements:

$$exp(\tilde{V}_i) \sum_{j' \in \mathcal{E}+n \setminus \{i\}} M_{j'i} = \sum_{j \neq i \in \mathcal{E}+n \setminus \{i\}} M_{ij} exp(\tilde{V}_j).$$

□

Proof of Result 3

Preliminaries: It is helpful to first have explicit expressions for a number of quantities. Write the R^2 between Ψ and V in terms of the known variable V and the unknown variable a:

$$R^2 = \frac{Cov(\Psi, V)^2}{Var(\Psi)Var(V)} \tag{A12}$$

$$= \frac{Cov(\Psi, \alpha(\Psi + a))^2}{Var(\Psi)Var(\alpha(\Psi + a))} \tag{A13}$$

$$= \frac{\alpha^2 Cov(\Psi, (\Psi + a))^2}{\alpha^2 Var(\Psi)Var((\Psi + a))} \tag{A14}$$

$$= \frac{[Var(\Psi) + Cov(\Psi, a)]^2}{Var(\Psi)[Var(\Psi) + Var(a) + 2Cov(\Psi, a)]}. \tag{A15}$$

It is also helpful to write $Var(a)$ in terms of one unknown quantity by rearranging equation (A15):

$$R^2[Var(\Psi)^2 + Var(\Psi)Var(a) + 2Var(\Psi)Cov(\Psi, a)] = Var(\Psi)^2 + 2Var(\Psi)Cov(\Psi, a) + Cov(\Psi, a)^2, \tag{A16}$$

$$R^2 Var(\Psi)Var(a) = (1 - R^2)Var(\Psi)^2 + 2(1 - R^2)Var(\Psi)Cov(\Psi, a) + Cov(\Psi, a)^2, \tag{A17}$$

$$Var(a) = \frac{(1 - R^2)Var(\Psi)^2 + 2(1 - R^2)Var(\Psi)Cov(\Psi, a) + Cov(\Psi, a)^2}{R^2 Var(\Psi)}. \tag{A18}$$

The following is a useful expression for $Corr(\Psi, a)$:

$$Corr(\Psi, a) = \frac{Cov(\Psi, a)}{\sqrt{Var(a)Var(\Psi)}}, \tag{A19}$$

$$= \frac{Cov(\Psi, a)}{\sqrt{\frac{(1-R^2)Var(\Psi)^2+2(1-R^2)Var(\Psi)Cov(\Psi,a)+Cov(\Psi,a)^2}{R^2 Var(\Psi)} Var(\Psi)}}, \tag{A20}$$

$$= \sqrt{R^2} \frac{Cov(\Psi, a)}{\sqrt{(1 - R^2)Var(\Psi)^2 + 2(1 - R^2)Var(\Psi)Cov(\Psi, a) + Cov(\Psi, a)^2}}. \tag{A21}$$

Proof. **A lower bound on $Var(a)$:** To minimize $Var(a)$, start with the expression for $Var(a)$ (equation (A18)) in terms of $Cov(\Psi, a)$ and take the first order condition with respect to $Cov(\Psi, a)$:

$$\frac{\partial Var(a)}{\partial Cov(\Psi, a)} = \frac{2(1 - R^2)Var(\Psi) + 2Cov(\Psi, a)}{R^2 Var(\Psi)} \tag{A22}$$

$$0 = \frac{2(1 - R^2)Var(\Psi) + 2Cov(\Psi, a)}{R^2 Var(\Psi)} \tag{A23}$$

$$Cov(\Psi, a) = -(1 - R^2)Var(\Psi). \tag{A24}$$

The second order condition is $\frac{2}{R^2 Var(\Psi)}$, which is positive. Substitute this into the expression for

$Var(a)$ (equation (A18)) to get that the minimum value is given by:

$$Var(a) = \frac{(1-R^2)Var(\Psi)^2 + 2(1-R^2)Var(\Psi)(-(1-R^2)Var(\Psi)) + (-(1-R^2)Var(\Psi))^2}{R^2 Var(\Psi)} \tag{A25}$$

$$= Var(\Psi)(1-R^2). \tag{A26}$$

Compute the correlation between Ψ and a at this lower bound:

$$Corr(\Psi, a) = \frac{Cov(\Psi, a)}{\sqrt{Var(a)Var(\Psi)}} \tag{A27}$$

$$= \frac{-(1-R^2)Var(\Psi)}{\sqrt{Var(\Psi)(1-R^2)Var(\Psi)}} \tag{A28}$$

$$= -\sqrt{1-R^2}. \tag{A29}$$

And compute the variance of utility in log dollar units:

$$Var(\Psi + a) = Var(\Psi) + Var(a) + 2Cov(\Psi, a) \tag{A30}$$

$$= Var(\Psi) + Var(\Psi)(1-R^2) - 2(1-R^2)Var(\Psi) \tag{A31}$$

$$= R^2 Var(\Psi). \tag{A32}$$

An upper bound on $Var(a)$: Take the limit of the expression for for $Var(a)$ (equation (A18)):

$$\lim_{Cov(\Psi,a)\to\infty} \frac{(1-R^2)Var(\Psi)^2 + 2(1-R^2)Var(\Psi)Cov(\Psi,a) + Cov(\Psi,a)^2}{R^2 Var(\Psi)} = \infty. \tag{A33}$$

Note that this implies that $Var(a)$ goes to infinity with the square of $Cov(\Psi, a)$, which is why the R^2 expression remains finite.

What is $Corr(\Psi, a)$ in this case?

$$\lim_{Cov(\Psi,a)\to\infty} Corr(\Psi, a) = \lim_{Cov(\Psi,a)\to\infty} \sqrt{R^2} \frac{Cov(\Psi, a)}{\sqrt{(1-R^2)Var(\Psi)^2 + 2(1-R^2)Var(\Psi)Cov(\Psi,a) + Cov(\Psi,a)^2}} \tag{A34}$$

$$= \sqrt{R^2} \tag{A35}$$

$$= Corr(\Psi, V). \tag{A36}$$

And:

$$Var(\Psi + a) \to \infty. \tag{A37}$$

Rosen vs. Mortensen amenities: To decompose the a term into *Rosen* and *Mortensen* amenities, note that the values for the *Rosen* amenities correspond to the lower bounds in these results.

The willingness to pay for a *Rosen* amenity is one. Consider $\hat{\beta}$ estimated from $\Psi = \beta a_{\text{Rosen}}$.

Then $\hat{\beta} = \frac{Cov(\Psi, a_{\text{Rosen}})}{Var(a_{\text{Rosen}})} = \frac{-(1-R^2)Var(\Psi)}{(1-R^2)Var(\Psi)} = -1.$

$\square$

B Appendix: Description of estimating the model

With N employers, the model depends on the following parameters:

- $N \times \{f_i, g_i, V^e(v_i), \rho_i, \delta_i\} = 5N,$
- $\{W, \lambda_1, V^n\} = 3,$

for a total of $5N + 3$ parameters. In the model, there are three additional parameters, λ_0, U, and β. It turns out that I do not need to know these parameters to estimate the parts of the model that I want to estimate. (If I wanted to estimate b or v_i I would need to know these three parameters).

I use $5N + 4$ moments. In particular, $N \times \{f_i^o, g_i, \tilde{V}_i, \rho_i, \delta_i\} = 5N$, and W is also observed (f_i^o is the "observed" share of hires from nonemployment and is defined more formally in Section B.2). The three remaining moments are $\tilde{V}_n$, the probability of making an endogenous $employer-to-employer$ transition (equation (A50)), and a moment that relates the employers that hire workers from nonemployment to the value of nonemployment (equation (A43)). Intuitively, the value of nonemployment, V^n, is increasing in the probability of making a nonemployment-to-employment transition, while the probability of an outside offer λ_1 is increasing in the probability of an endogenous $employer-to-employer$ transition. The overidentification comes in the fact that $\tilde{V}_n$ also contains information about the value of nonemployment.

I now present the steps for estimating the model, which outlines both the equations I need to determine the parameters as well as the solution algorithm. There are two groups of steps: calibration steps and moment matching steps.

B.1 Data and calibration steps

Step 1: The relative size of employers (g_i) and the number of workers (W) are summary statistics of the data. This gives $N + 1$ parameters.
Step 2: Get a firm-specific estimate of the δ_i and the ρ_i. Implement the method discussed in section 1.2.1 to measure displaced workers. This gives the matrix of mobility that reflects preferences, or M. Aggregating the displaced mobility by source relative to the number of workers gives the probability of job destruction shocks (δ_i) and the probability of reallocation shocks ($(1 - \delta_i)\rho_i$). This step gives $2N$ parameters.

B.2 Moment matching steps

Define f_i^o to be the share of workers hired from nonemployment that are hired by firm i, or:

$$f_i^o \equiv \frac{f_i \frac{exp(V^e(v_i))}{exp(V^e(v_i))+exp(V^n)}}{\sum_{i' \in \mathcal{E}} f_{i'} \frac{exp(V^e(v_{i'}))}{exp(V^e(v_{i'}))+exp(V^n)}} \frac{\lambda_0 U}{\lambda_0 U} = \frac{M_{in}}{\sum_{j \in \mathcal{E}} M_{jn}}. \tag{A38}$$

The denominator is the share of offers accepted by nonemployed workers. This equation only identifies f_i up to scale.[48] Hence, I use the natural normalization $\sum_{i \in \mathcal{E}} f_i = 1$.

Define C_1 to be the share of offers that are accepted from nonemployment, or:

$$C_1 \equiv \sum_{i' \in \mathcal{E}} f_{i'} \frac{exp(V^e(v_{i'}))}{exp(V^e(v_{i'})) + exp(V^n)}, \tag{A39}$$

so that f_i^o can be written more compactly as

$$f_i^o = \frac{f_i \frac{exp(V^e(v_i))}{exp(V^e(v_i)) + exp(V^n)}}{C_1}. \tag{A40}$$

Use a grid-search to find a value for λ_1 (the arrival rate of offers) that minimizes the gap between the probability of an employer-to-employer transition in the data and the model. The reason to use grid-search is that the function from a guess of λ_1 to a new value of λ_1 implied by the following steps is not a contraction mapping (nor is it guaranteed that the model can exactly reproduce the employer-to-employer transition probability in the data).

Step 1: Solve the following equations, where I maintain the convention of data or variables whose values are known by a given step are on the left-hand side, while unknowns are on the right-hand side. In the following equation, g_i and f_i^o are known directly from data, δ_i and ρ_i are estimated based on the displaced workers,[49] and $\tilde{V}_i$ is estimated based on the matrix of moves across firms:

$$\frac{g_i exp(\tilde{V}_i)}{f_i^o}(1-\delta_i)(1-\rho_i) = g_i \frac{f_i exp(V^e(v_i))}{g_i(1-\delta_i)(1-\rho_i)} C_1 \frac{1}{f_i} \frac{exp(V^e(v_i)) + exp(V^n)}{exp(V^e(v_i))}(1-\delta_i)(1-\rho_i) \tag{A41}$$

$$= C_1[exp(V^e(v_i)) + exp(V^n)]. \tag{A42}$$

In the next equation, W is data, M_{in} is known from the matrix of moves, λ_1 is from the calibration step, and $exp(\tilde{V}_n)$ is estimated from the matrix of moves:

$$\frac{1}{1-\lambda_1}\frac{1}{W}\sum_{i \in \mathcal{E}} M_{in} exp(\tilde{V}_n) = \frac{1}{1-\lambda_1}\frac{1}{W}\sum_{i \in \mathcal{E}} \lambda_0 U f_i \frac{exp(V^e(v_i))}{exp(V^n) + exp(V^e(v_i))} \frac{(1-\lambda_1) W exp(V^n)}{\lambda_0 U} \tag{A43}$$

$$= \frac{1}{1-\lambda_1}(1-\lambda_1) exp(V^n) \sum_{i \in \mathcal{E}} f_i \frac{exp(V^e(v_i))}{exp(V^n) + exp(V^e(v_i))} \tag{A44}$$

$$= exp(V^n) C_1. \tag{A45}$$

Step 2: Given a value of λ_1, do the following:

- Combining equations (A42) and (A45), give the following two terms: $C_1 exp(V^e(v_i))$ and $C_1 exp(V^n)$.

[48]For all α, $\frac{f_i \frac{exp(V^e(\ln v_i))}{exp(V^e(\ln v_i)) + exp(V^u)}}{\sum_{i' \in \mathcal{E}} f_{i'} \frac{exp(V^e(\ln v_{i'}))}{exp(V^e(\ln v_{i'})) + exp(V^u)}} = \frac{\alpha f_i \frac{exp(V^e(\ln v_i))}{exp(V^e(\ln v_i)) + exp(V^u)}}{\sum_{i' \in \mathcal{E}} \alpha f_{i'} \frac{exp(V^e(\ln v_{i'}))}{exp(V^e(\ln v_{i'})) + exp(V^u)}}$.

[49]In practice, I used the person-year-weighted average value of $(1-\delta_i)(1-\rho_i)$.

- Rewrite Equation (A39) by multiplying by $\frac{C_1}{C_1}$ and rearranging:

$$f_i^o = f_i \frac{exp(V^e(v_i))}{exp(V^e(v_i)) + exp(V^n)} \frac{1}{C_1} \tag{A46}$$

$$= f_i \frac{C_1 exp(V^e(v_i))}{C_1 exp(V^e(v_i)) + C_1 exp(V^n)} \frac{1}{C_1} \tag{A47}$$

$$f_i^o \frac{C_1 exp(V^e(v_i)) + C_1 exp(V^n)}{C_1 exp(V^e(v_i))} = \frac{f_i}{C_1}. \tag{A48}$$

  In this equation, the terms on the left-hand side are known from step 1, so this step gives $\frac{f_i}{C_1}$.

- Now that $\frac{f_i}{C_1}$ is known, solve for C_1 by using the normalization $\sum_{i \in \mathcal{E}} f_i = 1$ (note that C_1 contains f_i, so the scale transformation cancels out.[50]).

$$\sum_{i \in \mathcal{E}} \frac{f_i}{C_1} = \frac{\sum_{i \in \mathcal{E}} f_i}{C_1} = \frac{1}{C_1}. \tag{A49}$$

  Now that C_1 is known and from equation (A48) $\frac{f_i}{C_1}$ is known, it is possible to solve for f_i, since $\frac{f_i}{C_1}$ is known from equation (A48).

- Knowledge of C_1 gives $exp(V^n)$ and $exp(V^e(v_i))$, via equations (A42) and (A45).

This step produces the following $2N + 1$ parameters: $\{f_i, exp(V^e(v_i)), exp(V^n)\}$. In combination with the calibration step, this gives $5N + 3$ parameters.
Step 3: Given the parameters of the model, compute the probability of an "endogenous" employer-to-employer transition:

$$\lambda_1 \sum_i g_i(1 - \delta_i)(1 - \rho_i) \sum_j f_j \frac{exp(V_j^e)}{exp(V_j^e) + exp(V_i^e)}. \tag{A50}$$

To make this computationally feasible, group firms into $1,000$ categories on the basis of the firm values (V^e).

C Appendix: Constructing datasets

C.1 Annual dataset

I follow Abowd, Lengermann, and McKinney (2003) to construct the dataset to estimate the earnings decomposition. I depart from them to define employment in a way that is consistent with how employment is defined to construct employer-to-employer flows, to follow more recent literature in imposing age restrictions, and to follow more recent literature in dropping jobs with very low earnings.

For the purposes of estimating the earnings decomposition, the *annual dominant employer* is the employer from which the worker had the highest earnings in the calendar year. This employer is

[50] Define $\hat{f}_i = \alpha f_i$ and let $\hat{C}_1$ be the C_1 constructed using $\hat{f}_i$. Then $\frac{\hat{f}_i}{\hat{C}_1} = \frac{f_i}{C_1} = \frac{\alpha f_i}{\sum_{i' \in \mathcal{E}} \alpha f_{i'} \frac{exp(V^e(v_{i'}))}{exp(V^e(v_{i'})) + exp(V^n)}}$

chosen from the employers from which the worker had received earnings for two or more consecutive quarters within the calendar year; the reason to make this restriction is to allow me to code transitions between employers as employer-to-employer or employer-to-nonemployment-to-employer.[51] In this set of jobs, the annual dominant employer is the one with the highest total earnings in the calendar year.

To construct annualized earnings, for each quarter within a year I first identify the nature of the worker's attachment to the employer. Specifically, code quarter t of earnings into one of the following two mutually exclusive categories: full-quarter (if earnings from the employer are in quarters $t-1$, t and $t+1$) or continuous (if earnings are in quarters $t-1$ and t or in t and $t+1$). Annualize these earnings as follows. First, if the worker had any quarters of full-quarter earnings, take the average of these quarters and multiply by 4 to get an annualized salary. Second, if a the worker did not have full-quarter earnings and has any quarters of continuous earnings, take the average of these and multiply by 8 to get an annualized salary. The justification for this procedure is that if a worker is present in only two consecutive quarters and if employment duration is uniformly distributed then on average the earnings represent $\frac{1}{2}$ a quarter's work, while if a worker is present in both adjacent quarters then the earnings reflect a full quarter's work.[52] Then take the log of these earnings.

I then make two additional sample restrictions. First, I keep workers aged 18-61 (on December 31st of the year), inclusive. This is an attempt to avoid issues with retirement. This age restriction is similar to, e.g., Card, Heining, and Kline (2013) (20-60 in Germany) and Taber and Vejlin (2013) (19-55 in Denmark), though Abowd, Lengermann, and McKinney (2003) do not report imposing any age restriction. Second, following Card, Heining, and Kline (2013) I drop observations with *annualized* earnings of less than \$3,250 in 2011:IV dollars.[53]

I use this dataset to construct three employer-level characteristics. The first one is average annual employment (based on employment as defined earlier in this section). The second one is average age of these workers. Finally, I construct the average log annualized earnings of workers. I also include industry (four-digit NAICS codes), location (county), and whether the employer is a multi-unit.

I now summarize how the various sample restrictions affect the same size. Table A1 in Appendix H shows that there are about 650 million person-employer-years before imposing an earnings test, 614 million after imposing an earnings test, and 505 million after going down to one observation per person per year. This means that after dropping the low-earnings jobs, there are an average of 1.2 employers per person per year.[54] Table A2 shows the distribution of the number of jobs per

[51] This eliminates quarters of employment that Abowd, Lengermann, and McKinney (2003, pg. 15-16) term "discontinuous," that is, where a worker is observed in neither adjacent quarter. Abowd, Lengermann, and McKinney (2003, pg. 15-16) report that such discontinuous quarters of employment accounted for 5 percent of person-year observations in their final dataset. Second, it eliminates "continuous" quarters of employment where the first quarter of the match is quarter *IV* within the year, and the second quarter is quarter *I* of the following year. Under the assumption that continuous quarters are uniformly distributed within the year, this eliminates $\frac{1}{8}$ of continuous workers. Abowd, Lengermann, and McKinney (2003, pg. 15-16) report that continuous quarters account for 11 percent of observations in their final dataset, so this eliminates about 1.4 percent of observations.

[52] In the small number of cases where a worker had forward-looking continuous employment in quarter IV as well as another quarter of continuous employment at the same employer, I included this quarter in the earnings calculation.

[53] Card, Heining, and Kline (2013) drop daily wages of less than 10 euros. 10 euros $\times \approx 1.3$ euros per dollar $\times 250$ days per year= 3,250.

[54] For Germany, Card, Heining, and Kline (2013, Appendix Table 1a, row 5) find 1.10 employers per person per year, and this number is stable from 1985 through 2009.

year in row 2 of Table A1.

Table A3 shows that on the full annual dataset, 91% share of person-year observations are full-quarter and 9% are continuous.[55] Table A4 shows the distribution of the number of years per person. About 40% of the people are in the dataset for all 7 years, and only 13% are in the dataset for only a single year. Table A5 shows that there is a substantial amount of mobility in this sample: half of the workers have two or more employers. Table A6 shows that about 10% of person-employer matches (or 30% of person-years) last for the entire span of my data. However, almost half of matches (20% of person-years) only last for a single year.

C.2 Quarterly dataset

I build on ideas developed in Bjelland et al. (2011) and Hyatt et al. (2014). Specifically, the procedure of restricting to jobs with two quarters of earnings and using overlapping quarters of earnings to label an employer-to-employer transition comes from Bjelland et al. (2011, pg. 496, equation 2). The idea of using earnings in the two quarters to select the dominant job is found in Hyatt et al. (2014, pg. 3).

For the purposes of measuring flows, the *quarterly dominant employer* in quarter t is the employer from which the worker had the highest earnings summing over quarter t and quarter $t-1$. This job is chosen from among the employers where the worker had positive earnings in *both* quarter t *and* quarter $t-1$. To count as employment, the earnings must pass the the same earnings test as for the annual dataset.[56]

For the person-quarters that remain after the earnings test, the goal is to select a single employer—the quarterly dominant employer. The quarterly dominant employer is the employer from which the worker has the most total earnings summing across $t-1$ and t. There is one exception to this selection rule. If a worker has earnings from her *annual* dominant employer in quarters $t-1$ and t, then this employer is the quarterly dominant employer regardless of whether it is the employer with the most total earnings summing across $t-1$ and t. The reason for prioritizing the annual dominant job is that I want to use this quarterly dataset to code transitions between annual dominant jobs so it is important that they appear in the quarterly dataset.

If a worker has different quarterly dominant employers in quarter t and quarter $t+1$, then this worker had earnings from both employers in quarter t and I label the worker as having undergone an employer-to-employer transition in quarter t. If a worker has no dominant employer in quarter $t+1$, then, with one exception highlighted a little later in this section, I consider that worker to have been nonemployed in quarter $t=1$, so I label the transition from the quarter t dominant employer as a transition into nonemployment.[57].[58]

I depart from prior work to address the possibility that workers move on the seam between

[55] Abowd, Lengermann, and McKinney (2003, pg. 15-16) find 84% are full-quarter, 11% are continuous, and 5% are discontinuous.

[56] Sum together the two quarters of earnings and multiply by 4. If the earnings are below \$3,250, then drop the person-employer match. Multiplying by 4 is justified if one assumes that each quarter is a continuous quarter of employment. The assumption that this is a continuous quarter of employment does lead to more jobs being included than the annual dataset; specifically, if a job is actually full-quarter, then the annualized earnings treating it as full quarter can be lower than the annualized earnings assuming it is a continuous quarter.

[57] Similarly, Burgess, Lane, and Stevens (2000) drop matches that only last a single quarter.

[58] This definition of a transition into nonemployment will pick up very few recalls as employer-to-nonemployment-to-employer transitions. The reason is that even if a worker is nonemployed awaiting recall for 13 weeks, the probability that I record a quarter with zero earnings from her employer is less than 10% ($\frac{1}{13}$).

two quarters (Hyatt and McEntarfer (2012) emphasize that on some outcomes these transitions look like employer-to-employer moves). To make this concrete, suppose that I observe a worker at firm A in quarter $t-2$ and $t-1$ and at firm B in t and $t+1$. Then the definitions developed previously say that in quarter $t-1$ firm A is the dominant employer and in quarter $t+1$ firm B is the dominant employer. But in quarter t the worker had no dominant employer because it was not the second consecutive quarter of any employment relationship. So the transition from A to B was an employer-to-nonemployment-to-employer transition. It might be, however, that the worker's last day at A was the last day of quarter $t-1$ and her first day at B was the first day of quarter t, so this was actually an employer-to-employer transition. The way I attempt to capture these transitions is to use the stability of earnings across quarters to suggest that a worker was probably employed for the full quarter in both quarters. So, if the earnings from firm A in quarters $t-2$ and $t-1$ are within 5% of each other (using quarter $t-1$ earnings as the denominator), then this employer is the dominant employer in quarter t. This then allows me code the transition from A to B as employer-to-employer. Table A7 shows that this correction accounts for 3.5% of the employer-to-employer transitions in my dataset.

The final result is a dataset that at the quarterly level says where the person was employed and, if this is a new job, says whether the worker came to this job directly from another job, or had an intervening spell of nonemployment.

C.3 Using the quarterly dataset to construct exogenous weights

I use the quarterly dataset to construct the exogenous weights. I proceed in the following steps:

1. Using the definition of employed explained in the previous section, construct employer size in each quarter.
2. Compute quarter-to-quarter employer growth rates (the growth rate in quarter t is the change in employment from t to $t+1$).
3. Label worker separations from the employer in quarter t as either employer-to-employer or employer-to-nonemployment using the definitions in the previous section.
4. Compute the probability of each separation type (employer-to-employer and employer-to-nonemployment) within size-demographic-growth rate bin.
5. Finally, the exogenoust weight is one minus the probability of the separation divided by the probability of the separation occuring at an expanding employer.[59]

I use the following employer size bins by number of workers in that quarter: 1-4, 5-9, 10-24, 25-49, 50-99, 100-249, 250+. I create a different number of growth bins by each employer size. Specifically, I use the following number of growth bins per size category: 2, 3, 5, 9, 11, 16, 26. One bin is always the expanding employers, and the remaining bins are equally-weighted by person-quarter bins among contracting employers. I create 40 distinct group workers by demographic characteristics denoted by d (20 equally-sized age bins for men, and 20 for women).

[59]In the case where this is less than 0, I set this to 0.

C.4 Combining the quarterly and annual datasets

The goal of combining the datasets is twofold: first, to use the detail of the quarterly dataset to label each transition between annual dominant employers as an employer-to-employer or an employer-to-nonemployment-to-employer transition and second, to find out the growth rate at the annual dominant employer in the quarter that the worker separated in order to construct the endogenous weight as discussed in section 1.2.1.

To label the transition as employer-to-employer or employer-to-nonemployment-to-employer, I proceed as follows. First, identify consecutive observations where a worker has a different annual dominant employer; to be concrete, suppose that the worker's annual dominant employer is A in 2002 and B in 2003.[60] Second, look at the quarterly dataset and find the last quarter that the worker is employed at A (this might be in 2002 or 2003). Third, look at the quarterly dataset and find the first quarter that the worker is employed at B (this might be in 2002 or 2003). If the last quarter at A and first quarter at B are adjacent, then there was an overlapping quarter of earnings and I label this an employer-to-employer transition. If not, then typically I label this an employer-to-nonemployment-to-employer transition. The exception to labelling the transition an employer-to-nonemployment-to-employer transition is if the worker made an employer-to-employer move through some third (and possibly fourth or fifth) employer en route to moving from A to B. Suppose, for example, that the worker makes the following transitions (where EE is employer-to-employer and ENE is employer-to-nonemployment-to-employer): $A \xrightarrow{EE} C \xrightarrow{EE} B$. Because the worker only made employer-to-employer transitions between A and B, I label this an employer-to-employer transition between annual dominant employers. Alternatively, suppose that I observe $A \xrightarrow{EE} C \xrightarrow{ENE} B$. Then I label the transition between annual dominant employers an employer-to-nonemployment-to-employer transition.

To assign the endogenous weights, I proceed as follows. First, I use the quarterly dataset to identify the quarter in which the worker separated from her annual dominant employer. Second, I use the quarterly dataset to identify whether the worker separated in an employer-to-employer or an employer-to-nonemployment way. Third, I use the quarterly dataset to measure how much the employer grew in the quarter the worker was separating; i.e., if quarter t is the last quarter the worker was employed, then I compute the change in employment at the employer from quarter t to $t + 1$. Finally, I compute the firm size in quarter t and worker demographics to merge on the relevant endogenous weight using a) firm characteristics, b) worker characteristics, c) firm growth rate, and d) nature of separation (employer-to-employer or employer-to-nonemployment) as merging variables. A transition from nonemployment always gets an endogenous weight of 1.[61]

If a worker never has another employer, then I do not attempt to label this transition. For example, if a worker has a dominant employer in 2006 and no dominant employer in 2007, then I do not record a separation in 2006. The reason is that this could occur for any number of reasons: 1) a worker ages out of my age range, 2) a worker moves out of my states, or 3) a worker leaves the labor force. For the purposes of computing transition probabilities, I remove these observations

[60] It is possible that a worker only appears in the annual dataset in nonconsecutive years—say, 2002 and 2004. In this case, the procedure ends up labelling the transition an employer-to-nonemployment-to-employer.

[61] In the case of multiple transitions between annual dominant jobs, I proceed as follows. In a case such as $A \xrightarrow{EE} C \xrightarrow{EE} B$, I compute the endogenous weight for the $A \to C$ and $C \to B$ transitions and take the geometric average. In a case such as $A \xrightarrow{EE} C \xrightarrow{ENE} B$, I take the geometric average of the $A \to C$ and $C \to$ nonemployment transitions for the endogenous weight on the separation from A to nonemployment; I then assign an endogenous weight of 1 to the nonemployment-to-employer transition.

from the denominator (that is, where the last year of a dominant employer is not the final year in the dataset). So the denominator for separation probabilities removes the last year where the worker has a dominant employer whether this is the last year in the dataset or before then (e.g., for a worker I see in 2004 and 2005, I do not count the 2005 separation in my separation probabilities). Similarly, when I compute g—share of employment—I do not include the final worker-specific year (rather than taking out the last calendar year).

To compute δ and ρ, I compute the probability that each transition was exogenous (one minus the endogenous weight). I then sum up the exogenous transitions over all transitions in the annual dataset and compute the relevant probabilities.

C.5 Constructing model-relevant objects

Total employment (W) and employer share of total employment (g): To use a common notion of employer size across all calculations, in the interval 2001-2007 (inclusive), I use all data except for the final person-year observation. This is so that I can compute separation probabilities for all person-years used to measure g. This means I use data from the period 2001-2006 to create the measure of employer size, but I might not count a particular person-year in g if this person never appears again.

Number of hires from nonemployment, and share of hires from nonemployment (f^{no}): To alleviate concerns that hires from nonemployment simply reflect migrants or labor market entrants, I only count workers as hired from nonemployment if I previously saw them employed in my data, and I labelled their previous transition an employment-to-nonemployment transition. Hence, I use hiring data from the period 2002-2007 (inclusive), except that I omit hires where the person was never previously employed in my data.

Exogenous job destruction and job reallocation shocks (δ, ρ): For each transition recorded in the annual dataset, I assign it a probability of being endogenous from the quarterly dataset (based on the worker's age and gender, the employer size, and whether the employer was growing or shrinking (and by how much) in the quarter). I then take the total number of separations of each kind ($employer-to-employer$ and $employer-to-nonemployment$) and compare it to the sum of the endogenous transition probabilities. The total number of exogenous $employer-to-employer$ transitions divided by total employment (W) is $(1-\delta)\rho$;[62] the number of exogenous $employer-to-nonemployment$ transitions divided by W is δ.

D Appendix: Computational details

Solving for the Abowd, Kramarz, and Margolis (1999) decomposition can only be done in the connected set of firms. Similarly, the model can only be estimated in the strongly connected set of firms.

To estimate the Abowd, Kramarz, and Margolis (1999) decomposition, I built on the code provided by Card, Heining, and Kline (2013). To identify the strongly connected set of firms, I use David Gleich's open source Matlab implementation of a depth-first search algorithm as part of the package Matlab BGL. To estimate the decomposition, I use Matlab's built-in preconditioned conjugate gradient function (pcg), with an incomplete Cholesky preconditioner and a tolerance of 0.01. For an extensive discussion of conjugate gradient, as well as the benefits of preconditioning,

[62] The $1-\delta$ appears because of timing assumptions in the model.

see Trefethen and Bau (1997) (especially Lectures 38 and 40). Despite the fact that my dataset is larger than that used by Card, Heining, and Kline (2013), I do not resort to a two-step estimation procedure of estimating the firm effects on the sample of movers. At least with my data and resources, the computational bottleneck was computing $X'X$. To get around this, I split X into 3 pieces and computed $X'X$ in 9 pieces.

To estimate the eigenvector, I use Matlab's built-in eigenvector solver that allows the researcher to specify the number of eigenvectors to solve for (eigs, rather than eig) (My own implementation of the power method yielded numerically identical answers.)

E Appendix: Selection-correcting the earnings

I selection-correct the earnings equation by combining the proportionality assumption and the results of the search model. That is, I add the expectation of the error term from the search model to the earnings equation. In the first period of a worker's employment relationship, this expectation depends on the identity of her prior firm in her first year at each firm. That is, suppose a worker moves from firm 2 to firm 1 then $\mathbb{E}[\iota_1 | V_1^e + \iota_1 > V_2^e + \iota_2] = \mathbb{E}[\iota_1 | \iota_1 - \iota_2 > V_2^e - V_1^e]$. In the second and subsequent years, this selection term for a worker at employer i is

$$\mathbb{E}[\iota | V_i^e, \text{not move}] = \frac{\sum_{\mathcal{E}-i,n} Pr(\text{offer from j and not move}) \mathbb{E}[\iota | \text{offer from j and not move}]}{\sum_{\mathcal{E}-i,n} Pr(\text{offer from j and not move})}. \quad \text{(A51)}$$

For a worker at i, these terms—when involving other firms—are:

$$Pr(\text{offer from j and not move}) = \lambda_1 f_j \frac{exp(V_i^e)}{exp(V_j^e) + exp(V_i^e)}, \quad \text{(A52)}$$

$$\mathbb{E}[\iota | \text{offer from j and not move}] = \gamma - \log\left(\frac{exp(V_i^e)}{exp(V_j^e) + exp(V_i^e)}\right). \quad \text{(A53)}$$

For a worker at i, these terms are (when involving nonemployment):

$$Pr(\text{offer from nonemp and not move}) = (1 - \lambda_1) \frac{exp(V_i^e)}{exp(V^n) + exp(V_i^e)}, \quad \text{(A54)}$$

$$\mathbb{E}[\iota | \text{offer from nonemp and not move}] = \gamma - \log\left(\frac{exp(V_i^e)}{exp(V^n) + exp(V_i^e)}\right). \quad \text{(A55)}$$

In implementation there are a couple issues. First, for the first year that a worker appears in the dataset I do not know which selection correction term to apply; that is, it might be that the worker showed up from another firm, or it might be that the worker had already been there. To address this, I assume that all such observations are in the second or subsequent years of the employment relationship. Second, there are firms that I cannot estimate the revealed value of, even though I can estimate the value of the firm in the earnings equation (these are firms in the strongly connected set for which I cannot estimate either f or g). For the purposes of the selection correction, I assume that $\frac{g}{f} = 1$ and therefore use the mobility relevant value. Third, to speed up computation, I discretize the firms into $1,000$ equally-sized (in terms of person-years) bins and use the bin means to compute the selection correction.

F Appendix: Measurement error

This appendix states formal conditions under which the grouping strategy leads to a consistent estimate of the correlation between firm-level values and earnings. I state results with a single grouping characteristic. The extension to grouping sequentially is straightforward.

F.1 Preliminaries

Consider one grouping characteristic, g, which might be location, industry, or size. Formally, let $\{\Omega_1, ...\Omega_G\}$ be mutually exclusive sets of firms. This grouping partitions the set of firms. Earnings at firm i are:

$$\Psi_i = \Psi_g + \tilde{\Psi}_i + \epsilon_i^{\Psi}, \tag{A56}$$

where Ψ_g is the group-level component, $\tilde{\Psi}_i$ is the firm-level component, and ϵ_i^{Ψ} is mean zero measurement error. Implicit in this notation is the fact that each firm i belongs to a unique group g. Similarly, the value at firm i is given by (where I suppress the e subscript from the text for simplicity):

$$V_i = V_g + \tilde{V}_i + \epsilon_i^{V}. \tag{A57}$$

I now state assumptions on how the terms relate.

Assumptions: *The assumptions are about the mean value of the non-group-level components:*

1. $\mathbb{E}[\tilde{\Psi}_i + \epsilon_i^{\Psi} | i \in \Omega_g] = 0.$
2. $\mathbb{E}[\tilde{V}_i + \epsilon_i^{V} | i \in \Omega_g] = 0.$

The economic content of these assumptions are that grouping characteristics are exogenous. These assumptions would be violated if firms see Ψ_i and then decide which group to choose, since g would be related to the error term. The statistical content of these assumptions are relatively mild. For example, they allow for the variance of the measurement error to depend on the grouping characteristic (i.e., when I group by size, it allows for the smaller firms to have higher variance). It also allows the measurement error to be correlated with the firm-specific component.

F.2 Estimation with one characteristic

With one characteristic, I can estimate the group-level components by the group-level means and then compute the relationship between the two.

For earnings,

$$\lim_{N_g \to \infty} \hat{\mathbb{E}}[\Psi_g] = \lim_{N_g \to \infty} \frac{1}{N_g} \sum_{i \in \Omega_g} \Psi_i \tag{A58}$$

$$= \lim_{N_g \to \infty} \frac{1}{N_g} \sum_{i \in \Omega_g} \left[\Psi_g + \tilde{\Psi}_i + \epsilon_i^{\Psi} \right] \tag{A59}$$

$$= \mathbb{E}[\Psi_g + \tilde{\Psi}_i + \epsilon_i^{\Psi} | i \in \Omega_g] \tag{A60}$$

$$= \Psi_g + \mathbb{E}[\tilde{\Psi}_i + \epsilon_i^{\Psi} | i \in \Omega_g] \tag{A61}$$

$$= \Psi_g, \tag{A62}$$

where the second line is a definition, the third is a law of large numbers, the fourth is a result of Ψ_g being nonstochastic once g is fixed, and the last line is by assumption.

Similarly, I can estimate V_g by the group level mean.

Then as the number of firms within each group grows large, I get a consistent estimate of $Corr(\Psi, V)$.

The limitation of this strategy is that if the grouping characteristics capture relatively little of the variance of the Ψ and V, then this is not particularly informative.

G Appendix: Subgroup results

This appendix reestimates the model on subgroups defined by age and gender. The central finding of the paper is robust within each subgroup. The assumption of homogeneity in the baseline results do not do too much damage to the data, since splitting the data along these dimensions of observable heterogeneity does not change the main finding.

I split the sample by men and women, as well as into "young" (18-34) and "old" workers (35-61). (I choose the age split so that each age range contains about half of the employer-to-employer transitions in the data.) By subgroup, I reestimate the whole model, the earnings decomposition and the comparison between them.

Table A11 in Appendix H shows that firms matter in explaining earnings inequality *within* each subgroup. It reports sample sizes and variance decompositions by subgroup in the set of firms strongly connected by employer-to-employer transitions made by that subgroup. For ease of comparison, I reproduce in the first column of Table A11 column (3) of Table 1, which reports the decomposition for the whole sample. Table A11 shows that firms explain a similar amount of the variance of earnings among men as among women, and this is very similar to the overall number. That said, there are more interesting patterns by age: firms explain more of the variance among young workers than among old workers.

Panel A of Table A12 in Appendix H shows that compensating differentials are an important part of the explanation for the role of firms in earnings inequality within each subgroup. The table reports the explanatory power of compensating differentials within each subgroup. The numbers are constructed in an identical way to the bottom panel of Table 4. For ease of comparison, I reproduce in the first row the numbers for the overall sample. Compensating differentials are slightly more important for explaining between-women inequality than between-men inequality. As with earnings inequality, the more striking difference is that compensating differentials are much more important for old workers than for young workers.[63]

Panel B of Table A12 reports the results of a conceptually distinct robustness exercise and shows that aggregating across industries and locations does not do too much damage to the data. One might be concerned that by treating all states in my sample as an integrated labor market I do important damage to the data. To assuage this concern, I reestimate the firm-level accepted-offer-relevant values ($\tilde{V}$) using only accepted offers within a particular state. State by state I then compute the correlation with the benchmark accepted-offer-relevant-firm-level values that also use the across-state moves. Finally, I aggregate the correlations across all states weighting by the number of person-years. Panel B of Table A12 shows that the correlation between the two measures is 0.97. That is, the accepted offers across states do not dramatically affect the estimates (but using these accepted offers allows me to compare utilities across counties).

At the firm level, one might also prefer to focus on within-industry moves based on the idea that perhaps across sector moves are not well described by a search model. The second column of Panel B in Table A12 reports a sector-by-sector analysis that is identical to the state-by-state analysis. I find a correlation for the within-sector accepted-offer-relevant values and the benchmark accepted-offer-relevant-values of 0.76. Of course, I rely on precisely these across-sector moves to value sectors as reported in Figure 10a.

H Appendix: Additional tables and figures

Table A1: Constructing Sample of Dominant Jobs

	Number (1)	Unique People (2)	Unique Employers (3)
Person-employer-year pre-earnings test	650,288,000	108,002,000	6,688,000
Person-employer-year post-earnings test	613,341,000	105,921,000	6,511,000
Person-years	504,945,000	105,921,000	6,155,000

Notes: All counts are rounded to the nearest thousand. Row 2 divided by row 3 is 1.215. The first row shows the total number of person-year-employer observations that are continuous quarter or full-quarter among workers in the relevant age range. The second row shows the number of person-year-employer observations where the person's dominant job in the particular year passes an earnings test. The third row goes down to the unique employer that provides the worker's "dominant" job, or the employer from which the worker made the most in the calendar year.

[63] Understanding why patterns look different for young and old workers is an interesting topic for future research.

Table A2: Distribution of jobs per person per year

	Number of person-years
1	$413,553,000$
2	$77,735,000$
3	$11,611,000$
4+	$2,047,000$

Notes: All counts are rounded to the nearest thousand. This table deconstructs the gap between row 2 and row 3 in Table A1. The column sum is row 3 in Table A1. This shows among workers in the sample of workers with dominant jobs the distribution of the number of continuous and full-quarter jobs in a year.

Table A3: Type of earnings in the annual dominant job dataset

Type of earnings	Number of person-years
Full quarter	$458,017,000$
Continuous quarter	$46,928,000$
Continuous quarter share	0.093

Notes: All counts are rounded to the nearest thousand. The column sum is the number of person-years in row 3 in Table A1. A worker is employed full-quarter in quarter t if she has earnings from her employer in quarter t and quarters $t-1$ and $t+1$. A worker is comployed in a continuous quarter way in quarter t if she has earnings from her employer in quarter t and quarter $t-1$ or quarter $t+1$.

Table A4: Number of years per person

	Number of people	Share
1	14,041,000	0.133
2	11,422,000	0.108
3	9,873,000	0.093
4	9,111,000	0.086
5	8,963,000	0.085
6	10,396,000	0.098
7	42,115,000	0.398

Notes: All counts are rounded to the nearest thousand. The column sum is the number of unique people in row 3 in Table A1.

Table A5: Dominant employers per person

Number of dominant employers	Number of people	Share of people
1	52,938,000	0.500
2	27,228,000	0.257
3	14,945,000	0.141
4	7,157,000	0.068
5	2,764,000	0.026
6	771,000	0.007
7	118,000	0.001

Notes: All counts are rounded to the nearest thousand. The column sum is the number of unique people in row 3 in Table A1.

Table A6: Number of years per match

Years per match	Matches (person-employers)	Share of matches	Share of person-years
1	93,327,000	0.466	0.185
2	39,176,000	0.196	0.155
3	19,842,000	0.099	0.118
4	12,295,000	0.061	0.097
5	8,573,000	0.043	0.085
6	6,745,000	0.034	0.080
7	20,175,000	0.101	0.280

Notes: All counts are rounded to the nearest thousand. The column sum in the first column is the number of matches and is approximately 200,000,000, which is between the number of unique people and the number of person-years. The next column shows the distribution by share of matches. The last column shows the distribution of person-years.

Table A7: Composition of separations in the quarterly dataset

Type of transition	Definition	Number
employer-to-nonemployment	Standard	131,621,000
employer-to-employer	Standard	76,152,000
employer-to-employer	New	2,680,000
employer-to-employer transition share		0.375
New definition share		0.035
Total separations		210,453,000

Notes: All counts are rounded to the nearest thousand. The dataset is the quarterly dataset, so it includes some workers not in the annual dataset. The standard definition uses overlapping quarters to measure employer-to-employer transitions. The new definition uses stability of earnings to measure employer-to-employer transitions.

Table A8: Workers frequently move across industries and locations

Moves with different...	EE and ENE	EE (endog)	EE (all)	Single-Unit Only?
State	0.12	0.10	0.10	
County	0.65	0.64	0.63	
Sector	0.65	0.61	0.60	
4-digit industry	0.81	0.79	0.78	
State	0.12	0.10	0.10	X
County	0.56	0.53	0.52	X
Sector	0.60	0.56	0.56	X
4-digit industry	0.77	0.74	0.73	X

Notes: This table reports the share of annual moves that cross geographic and industry boundaries. The bottom panel reports the moves that are between single-unit employers. EE moves are employer-to-employer moves and ENE moves are employer-to-nonemployment-to-employment.

Table A9: Variance decompositions of earnings and utility

Dummy set	Log Earnings	Firm-level earnings (Ψ)	Firm-level utility (V^e)
State	0.03	0.06	0.02
County	0.10	0.12	0.05
Sector	0.40	0.37	0.42
3-digit industry	0.51	0.46	0.49
4-digit industry	0.57	0.52	0.56
Size groups	0.03	0.03	0.03
		Single-units only	
State	0.03	0.05	0.01
County	0.12	0.13	0.04
4-digit industry	0.51	0.45	0.40
Size groups	0.06	0.06	0.08

Notes: The sample is column (4) in Table 1.

Table A10: Relationship between alternative firm effects

	S. Conn	Selec. Corr.	EE Movers	EE Receive	EE Send	S. Conn (w/slope)	Slope
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) S. Connected by EE	1	1.00	0.96	0.91	0.91	1.00	0.02
(2) Selection Corrected		1	0.96	0.91	0.91	1.00	0.02
(3) EE Movers only			1	0.95	0.94	0.95	0.03
(4) EE Receive				1	0.85	0.91	0.07
(5) EE Send					1	0.91	-0.03
(6) S. Connected by EE (with slope)						1	-0.03
(7) Slope							1

Notes: This table shows correlations between firm effects estimated using different samples and different identifying variation. EE stands for employer-to-employer. Column (1) is the benchmark firm effect estimated in the set of firms strongly connected by employer-to-employer transitions, which is identified by both employer-to-employer and employer-to-nonemployment-to-employer movers. Column (2) adds the selection correction from the structural model to the mobility equation. Column (3) estimates firm effects using only workers who make employer-to-employer transitions and only the person-years at the firms that these workers either leave or join. Hence, firm effects are identified using only employer-to-employer moves. Column (4) and (5) are estimated in a single regression. These firm effects are identified by the workers making employer-to-employer transitions, but allows for separate effects at the firm-level for a sending and receiving firm. These provide an additional test of the symmetry assumption reflected in the Abowd, Kramarz and Margolis (1999) decomposition. Columns (6) and (7) are estimated in a single regression. Column (6) is the firm effect, and column (7) is the slope effect. I can only estimate a slope in earnings at firms that have workers that stay for two or more years. This restriction limits the sample to 98.4% of firms representing 99.9% of person-years.

Table A11: Variance decomposition of earnings by subgroup

	Strongly connected by employer-to-employer				
	All	Men	Women	Old (35-61)	Young (18-34)
	(1)	(2)	(3)	(4)	(5)
Sample size					
People-years	470,387,000	233,475,000	223,113,000	271,736,000	181,222,000
People	100,547,000	50,474,000	47,918,000	58,402,000	49,053,000
Employers	1,971,000	1,279,000	1,237,000	1,260,000	1,353,000
Summary statistics					
Mean log earnings	10.45	10.62	10.27	10.66	10.14
Variance of log earnings	0.69	0.69	0.61	0.62	0.61
Share of variance of earnings explained by each parameter set					
Employers	0.22	0.22	0.23	0.20	0.27
People	0.55	0.55	0.54	0.71	0.63
Covariates	0.11	0.13	0.11	-0.01	-0.04
Overall fit of AKM decomposition					
R^2	0.88	0.89	0.87	0.90	0.85
Adj. R^2	0.85	0.86	0.83	0.87	0.79

Notes: Sample counts are rounded to the nearest thousand. Column (1) reproduces column (3) from Table 1. The data is at an annual frequency. There is one observation per person per year. The observation is the job from which a person made the most money, but only if she made at least $3,250 (in $2011). The table includes person-years in which on December 31 of the year the person was 18-61 (inclusive). AKM stands for Abowd, Kramarz, and Margolis (1999).

Table A12: Why do some firms pay so much and some so little? Subgroups

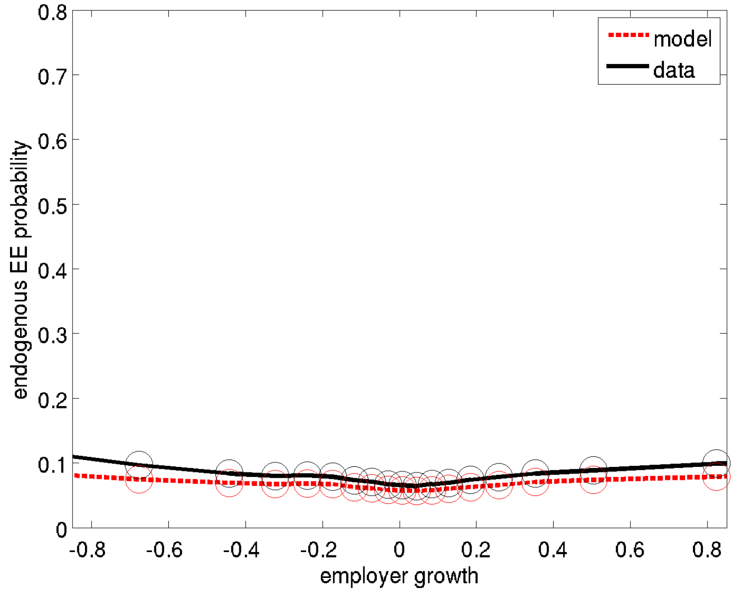
Panel A. Compensating differentials' share		
	Group	All
(1)	All	0.25
(2)	Men	0.31
(3)	Women	0.21
(4)	Old (35-61)	0.10
(5)	Young (18-34)	0.36

B. Correlations by firm characteristics:		
	State	Sector
Corr($\tilde{V}$,$\tilde{V}$ x by x)	0.97	0.76

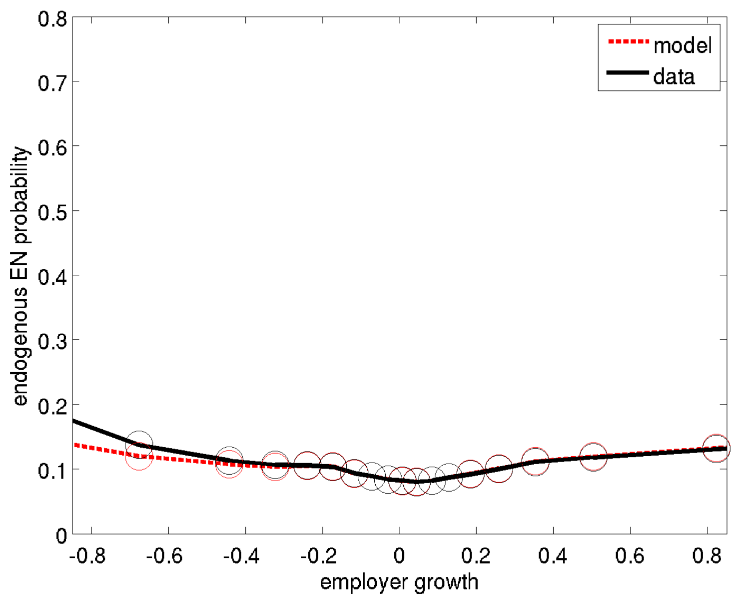
Notes: Panel A reports results analogous to Table 4. Row (1) reproduces results from the first row in column (4) in Table 4. The state-by-state exercise in Panel B contains 94.1% of the employers in column (3) of Table 1, which represent 99.3% of the person-years. The sector-by-sector exercise in Panel B contains 45.4% of the employers and 89.2% of the person-years.

Figure A1: Endogenous employer-to-employer and employer-to-nonemployment probabilities by firm growth rate

(a) Employer-to-employer

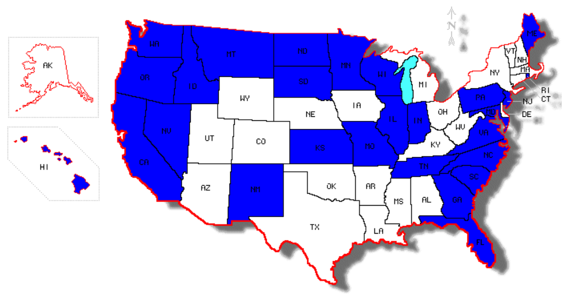


(b) Employer-to-nonemployment



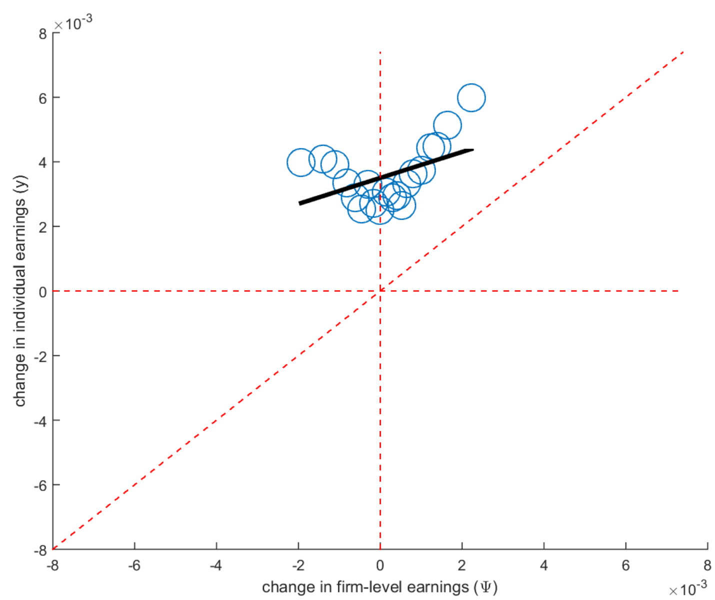
Notes: These figures plot the model-implied firm-specific endogenous employer-to-nonemployment and employer-to-employer probabilities, as well as these probabilities in the data as a function of firm-growth. Firm growth is growth from 2001 through 2007. To construct the figure, I sort firms into 20 equal person-sized year bins on the basis of firm-growth. To be consistent with the axes in figure 1a, this figure only displays seventeen bins.

Figure A2: States used in analysis



Notes: The states in blue are used in the analysis.

Figure A3: Change in firm effect does not predict magnitude of earnings change in a matching model



Notes: This figure is based on simulating the example production function in Eeckhout and Kircher (2011) and is constructed in a manner analogous to Figure 8.

I Appendix: Simulating models using individual data

I.1 Hall and Mueller (2013)

I simulate using the parameter values in the $\kappa = 0$ column of Hall and Mueller (2013, pg. 20, Table 2). To simplify matters, I remove the standard deviation of personal productivity and the reference value of nonwage value (σ_x and $\bar{n}$). The only relevant equation is then the mass balance equation:

$$G(v) = \frac{\lambda F(v) u}{(1-u)[\lambda(1-F(v)) + s]}, \tag{A63}$$

where G is the distribution of job value in the employed distribution, F is the distribution of job values in the accepted offer distribution, u is the unemployment rate, s is the job destruction rate, and λ is the arrival rate of offers on and off the job. $v = y + n$, or job value is the sum of earnings and nonpecuniary characteristics. The parameter values I use are in the following table.

Parameter	Explanation	Value	Notes
μ_y	mean of offers	0.37	(or 2.75-2.38)
σ_y	stdev of wage	0.304	in F
σ_n	stdev of non-wage value	0.882	in F
λ	offer arrival rate on/off job	0.058	
s	job destruction rate	0.0041	
$u = \frac{s}{\lambda+s}$	u/e rate	0.0660	

The key parameters are σ_y and σ_n. The Hall and Mueller (2013) estimates imply more variance in the value of nonpecuniary characteristics than earnings. I consider a million draws from the offer distribution and use equation (A63) to compute the steady state distribution. I then compute the R^2 between y and v in G.

I.2 Sullivan and To (2014)

The key mass balance equation in Sullivan and To (2014) is:

$$G(v) = \frac{\lambda_u F(v) Pr(v > U^*) u + \lambda_{le} F(v) Pr(v > U^*)(1-u)}{[\lambda_e(1-F(v)) Pr(v > U^*)(1-u) + (1-u)s + \lambda_{le} F(v) Pr(v > U^*)(1-u)]}, \tag{A64}$$

where $Pr(v > U^*)$ is the probability that the offer exceeds the reservation utility, λ_u is arrival probability of an offer when unemployed, and λ_e is the arrival probability of an offer when employed, λ_{le} is the reallocation shock (and I have changed some notation). Sullivan and To (2014) allow for unobserved heterogeneity and fit three types. The following table reports the values I use and is taken from Sullivan and To (2014, pg. 489, Table 2, specification 1). (The bottom two rows are computed as a function of the rest of the table.)

Parameter	Explanation	Type 1	Type 2	Type 3
σ_y	stdev of wage in F	0.3435	0.3435	0.3435
σ_n	stdev of non-wage value in F	0.3908	0.3908	0.3908
μ_y	mean wage of offers	1.1774	1.7252	2.1766
U^*	reservation utility	1.8163	1.8948	1.9869
λ_u	offer while unemp	0.9198	0.6299	0.1421
λ_e	offer while emp	0.4214	0.5348	0.0365
λ_{le}	reallocation	0.2545	0.0214	0.0016
s	job destruction rate	0.0905	0.0529	0.0345
$Pr(v > U^*)$	prob. of accepting an offer	0.1098	0.3726	0.6416
$u = \frac{s}{\lambda_u Pr(v>U^*)+s}$	u/e rate	0.4726	0.1839	0.2745