Ranking firms using revealed preference and other essays about labor markets

by

Isaac Sorkin

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2015

Doctoral Committee:

       Professor Matthew D. Shapiro, Chair
       Professor Daniel A. Ackerberg
       Professor John Bound
       Assistant Professor Joshua K. Hausman

For Mary

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Ranking firms using revealed preference and other essays about labor markets
by
Isaac Sorkin

Chair: Professor Matthew D. Shapiro

This dissertation contains essays on three questions about the labor market. Chapter 1 considers the question: why do some firms pay so much and some so little? Firms account for a substantial portion of earnings inequality. Although the standard explanation is that there are search frictions that support an equilibrium with rents, this chapter finds that compensating differentials for nonpecuniary characteristics are at least as important. To reach this finding, this chapter develops a structural search model and estimates it on U.S. administrative data. The model analyzes the revealed preference information in the labor market: specifically, how workers move between the 1.5 million firms in the data. Evidence for compensating differentials is workers systematically moving towards lower-paying firms, while evidence for rents is workers systematically moving towards higher-paying firms. With on the order of 1.5 million parameters, standard estimation approaches are infeasible and so the chapter develops a new estimation approach that is feasible on such big data.

Chapter 2 considers the question: why do men and women work at different firms? Men work for higher-paying firms than women. The chapter builds on chapter 1 to consider two explanations for why men and women work in different firms. First, men and women might search from different offer distributions. Second, men and women might have different rankings of firms. Estimation finds that the main explanation for why men and women are sorted is that women search from a lower-paying offer distribution than men. Indeed, men and women are estimated to have quite similar rankings of firms.

Chapter 3 considers the question: what are there long-run effects of the minimum wage? An empirical consensus suggests that there are small employment effects of

minimum wage increases. This chapter argues that these are short-run elasticities. Long-run elasticities, which may differ from short-run elasticities, are more policy relevant. This chapter develops a dynamic industry equilibrium model of labor demand. The model makes two points. First, long-run regressions have been misinterpreted because even if the short- and long-run employment elasticities differ, standard methods would not detect a difference using U.S. variation. Second, the model offers a reconciliation of the small estimated short-run employment effects with the commonly found pass-through of minimum wage increases to product prices.

CHAPTER I

Ranking firms using revealed preference

This paper takes a new approach to estimating the value of working at a particular firm that relies only on quantities, not on earnings. Specifically, I exploit the intuitive notion that workers move towards firms with higher value. I use this revealed preference approach combined with a standard search-theoretic model to estimate the value of working at essentially each firm in the United States. This approach imposes sufficient structure to map the 1.5 million by 1.5 million matrix of worker flows across all firms in the economy into estimates of firm value.

I then combine the revealed-preference based estimate of firm values with earnings data to decompose the variance of firm-level earnings into a rents and compensating differentials component. Conditional on person fixed effects, firms account for over 20% of the variance of earnings (e.g. Abowd, Kramarz, and Margolis (1999) and Card, Heining, and Kline (2013)).[1] There are two main explanations for why: rents and compensating differentials. Recent literature has focused on the rents explanation (e.g. Postel-Vinay and Robin (2002)). In a decentralized market, frictions prevent workers from moving to—or bidding away the rents at—the high-paying firms. An alternative explanation is compensating differentials (e.g. Rosen (1986)). Firms differ not only in how much they pay, but also in nonpecuniary characteristics. According to the compensating differentials explanation, people do not want to move to the higher paying firms because the higher pay compensates for the variation in nonpecuniary characteristics. Disentangling these explanations in labor market equilibrium can be quite difficult, but the revealed preference approach allows me to do so. Intuitively, rents explain the variation in firm-level earnings to the extent that the higher paying firms are the higher value firms, while compensating differentials explain variation in

---

[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.

firm-level earnings to the extent that higher paying firms are not the higher value firm.

Using matched U.S. employer-employee data, I find that both rents and compensating differentials explanations are operative, but compensating differentials are more important than rents. The quantitative importance of compensating differentials stands in contrast to a conventional wisdom that they are relatively unimportant in explaining the structure of earnings. Compensating differentials allows a benchmark search model to reproduce the extent of earnings dispersion while also yielding a plausible value of nonemployment. The presence of compensating differentials increases earnings inequality. If amenities were removed and earnings changed to compensate, then the variance of earnings would fall.

In the first part of the paper, I write down—and develop the tools to estimate—a simple model of the labor market that contains both the rents explanation and the compensating differentials explanation. The model is a benchmark partial equilibrium utility-posting model in the spirit of Burdett and Mortensen (1998). The non-standard ingredient in the model is that firms post a utility offer that consists of both earnings and a nonpecuniary bundle, rather than just posting earnings. The rents explanation is contained in the model because there is the possibility of equilibrium 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.

To estimate the levels of utility that firms post, I use only quantity information. I estimate the model in two steps. In the first step, I isolate the employer-to-employer and employer-to-nonemployment transitions that plausibly reveal preferences between employers (or are due to a modeled idiosyncratic shock) by using information about what the worker's coworkers were doing at the time of the separation. The idea is that if an unusually high share of other workers were also separating when the worker separates, then—in the spirit of the displaced worker literature (Jacobson, LaLonde, and Sullivan (1993))—a firm-level shock caused the workers to leave, and there is a high probability that any particular separation was involuntary. In contrast, if turnover levels look normal when the worker left, then this separation was endogenous in the sense that any shock must have been idiosyncratic to the worker.

In the second step, I measure the central tendency of worker flows in the labor market. The core piece of revealed preference information is the number of workers who go from firm A to B and from B to A, for all pairs of firms in the economy, as well as between each employer and nonemployment. I record these flows in a 1.5

million by 1.5 million matrix, where one cell is the number of workers who go from A to 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 the 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 suitably 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.

I then unravel this central tendency of mobility to recover the value of working at each firm. The model emphasizes two factors in addition to the value of a firm that affects the central tendency of mobility. First, a large firm will naturally have more workers moving away from it than a small firm. I can 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 estimate the offer distribution by 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.

I estimate the model on the U.S. Census Bureau's Longitudinal Employer House-hold Dynamics (LEHD) dataset. The model fits the worker flow data well: it reproduces both the probability of workers leaving each type of firm and the structure of worker flows. The model generates considerable heterogeneity across firms in the probability of an employer-to-nonemployment quit that is also present in the data. To illustrate this heterogeneity, I sort firms on the based of their estimated value. From the bottom 5% to the top 5% of firms, employer-to-nonemployment quits vary by a factor of about 10. Similarly, the model generates heterogeneity in the probability an employer-to-employer quit that is also present in the data. From the bottom to the top firms, these vary by a factor of 5. Finally, the model reproduces the detailed structure of employer-to-employer transitions: not only which firms workers leave, but also which firms they join.

The estimates provide a new source of evidence in favor of the frictional view of the labor market: all firms do not offer the same level of utility. In a neoclassical world, the idiosyncratic utility draw would drive all choices, whereas in a frictional world the common values would. Both explanations are operative: these two distributions have about equal dispersion.

In the second part of the paper, I estimate the earnings that firm 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)). This decomposition contains a firm effect in earnings that is estimated based on workers who switch firms.

To identify compensating differentials, I rely on a revealed preference argument. The feature of the data that most directly says that high-paying firms are not always high-utility firms is that there are some pairs of firms for which I observe people systematically moving from the higher-paying firm to the lower-paying firm. By revealed preference, this observation implies that there must be good nonearnings characteristics at the lower-paying firm and thus compensating differentials are operative. At the firm level, the revealed preference argument sets a high bar for finding compensating differentials. To find compensating differentials, the model needs to uncover patterns of moves to lower-paying firms that cannot be explained by idiosyncratic utility shocks. An idiosyncratic reason such as sharing a hobby with the boss might explain why one worker would move to the lower-paying firm, but the model only finds evidence for compensating differentials if all workers share this preference for the lower-paying firm. For example, all workers love the boss at the lower-paying firm.

Formally, I combine the firm-level utilities and earnings to measure the relative role of compensating differentials and rents by proving an identification result about measuring compensating differentials in frictional markets. 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 a theoretical tradition in the search literature that nonpecuniary characteristics might be positively correlated with earnings (Hwang, Mortensen, and Reed (1998) and Lang and Majumdar (2004)). In addition to the nonpecuniary characteristics that are compensating differentials, there is a separate component of nonpecuniary characteristics that are positively correlated with earnings; for example, some high-paying firms might also offer great benefits. But the variance of this second component is not identified. Thus, an arbitrarily large share of nonpecuniary characteristics might be positively correlated with earnings.

Compensating differentials account for about two-thirds of the variance of firm-level earnings, while rents account for the remaining third. The estimated ranking of sectors is intuitively plausible, as is the implied distribution of nonpecuniary characteristics. For example, education has good nonpecuniary characteristics, while firms in many blue-collar sectors that we associate with physically challenging work have bad nonpecuniary characteristics. The central finding that compensating differential are at least as important as frictions at explaining firm-level earnings is robust when I re-estimate the model across subgroups defined by age and gender.

The finding that rents, or frictions, do not explain all firm-level earnings dispersion is complementary to the argument in Hornstein, Krusell, and Violante (2011). The finding of a quantitatively large role for compensating differentials contrasts with Hornstein, Krusell, and Violante (2011) who—reflecting a consensus in the literature—argue (pg. 2883) that compensating differentials "do not show too much promise" in explaining earnings dispersion. As discussed further in section 1.6, I also find an empirically reasonable value unemployment, and thus pass a crucial test proposed by Hornstein, Krusell, and Violante (2011).

Finally, the model estimates imply that if the estimated nonpecuniary characteristics were removed and earnings changed to reflect this—so that the firm-level component of pay dispersion only reflected pure rents—then earnings inequality would decline. The effect of removing nonpecuniary characteristics and compensating workers on earnings inequality is theoretically ambiguous and depends on the correlation between earnings potential and nonpecuniary characteristics. In this counterfactual, earnings inequality as measured by the variance of earnings would decline. This reduction comes mainly from the lower tail of the income distribution shifting up.

The approach developed in this paper builds on several different strands of the search literature. To allow a distinction between utility and earnings, I focus on how workers move across firms. The dominant strand of the structural search literature seeks instead to understand the distribution of earnings. Bagger and Lentz (2014) also emphasize patterns in worker reallocation across firms, though they do not allow for nonpecuniary characteristics and do not exploit the complete structure of employer-to-employer moves in estimation. Similarly, Moscarini and Postel-Vinay (2014), Haltiwanger, Hyatt, and McEntarfer (2014) and Kahn and McEntarfer (2014) explore this feature of the data. The basic two-step estimation approach is similar in spirit to the two-step procedure in Postel-Vinay and Robin (2002).

The core identification idea that voluntary 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 relates variation in amenities to compensating differentials. In section 1.6, I discuss the these papers in more detail.[2]

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; e.g. Kline (2008) and Artuc, Chaudhuri, and McLaren (2010).

1.1 Ranking firms using revealed preference

1.1.1 A model with utility-posting firms

This section writes down 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 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 a firm posts is given by:

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

The potential trade-off between earnings and nonpecuniary characteristics allows the model to contain both the compensating differentials explanation and the rents explanation. The key innovation in this paper relative to the structural search literature is to directly estimate the V^e. Second, 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 A to 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. This restrictive form of preference heterogeneity is prob-

---

[2]There is also literature, e.g. Dey and Flinn (2005), Bonhomme and Jolivet (2009) and Aizawa and Fang (2013), which estimates the value of specific observable amenities in a search environment. See section 1.6 for further discussion of this approach.

ably 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 key additional assumption in the model is what Hall and Mueller (2013) term the proportionality-to-productivity hypothesis. Persistent heterogeneity only enters the model through a worker specific constant which shifts the flow payoff to all employers as well as nonemployment. 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. By exogenous moves, I mean moves that are related to a firm-level shock and would typically be thought of as involuntary. I identify these moves by building on the displaced worker literature, which aims to capture mass layoff events. By endogenous moves, I mean moves that result from maximizing decisions in the model. From the worker perspective, some of these moves would be perceived as involuntary because they would be in response to a negative idiosyncratic draw at the current firm. For example, Mortensen and Pissarides (1994) is a model in which all separations are endogenous, but many of them are unpleasant for workers.

The following Bellman equation summarizes this verbal discussion of the model.

A worker at employer i has the following value function:[3]

$$
\underbrace{V^e(v_i)}_{\text{value of being at i}} = \underbrace{v_i}_{\text{flow payoff}} + \underbrace{\beta}_{\text{discounter}} \mathbb{E}\Bigg\{ \delta_i \underbrace{\int_{\iota_1} \{V^n + \iota_1\}dI}_{\text{exogenous job destruction}}
$$

$$
+ \rho_i(1-\delta_i) \underbrace{\int_{V^e(v')} \int_{\iota_2} \{V^e(v') + \iota_2\}dId\tilde{F}}_{\text{exogenous employer-to-employer (reallocation)}}
$$

$$
+ \underbrace{(1-\rho_i)(1-\delta_i)}_{\text{no exogenous shocks}}
$$

$$
\times \Bigg[\underbrace{\underbrace{\lambda_1}_{\text{offer } V^e(v')} \int \int_{\iota_3} \int_{\iota_4} \max\{\underbrace{V^e(v'') + \iota_3}_{\text{accept}}, \underbrace{V^e(v_i) + \iota_4}_{\text{reject}}\}dIdIdF}_{\text{endogenous employer-to-employer}}
$$

$$
+ \underbrace{(1-\lambda_1)}_{\text{no offer}} \int_{\iota_5} \int_{\iota_6} \max\{\underbrace{V^n + \iota_5}_{\text{accept}}, \underbrace{V^e(v_i) + \iota_6}_{\text{reject}}\}dIdI \Bigg] \Bigg\}. \qquad (1.1)
$$
$$
\underbrace{\phantom{+ (1-\lambda_1) \int_{\iota_5} \int_{\iota_6} \max\{V^n + \iota_5, V^e(v_i) + \iota_6\}dIdI}}_{\text{endogenous job destruction}}
$$

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

---

[3] 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).

equation:

$$
\underbrace{V^n}_{\text{value of being nonemployed}} = \underbrace{b}_{\text{flow payoff}}
$$

$$
+ \underbrace{\beta}_{\text{discounter}} \mathbb{E}\Bigg\{ \underbrace{\lambda_0 \int\limits_{\text{offer } V^e(v')} \int\limits_{\iota_7} \int\limits_{\iota_8} \max\{\underbrace{V^e(v') + \iota_7}_{\text{accept}}, \underbrace{V^n + \iota_8}_{\text{reject}}\} dI dI dF}_{\text{endogenous nonemployment-to-employment}}
$$

$$
+ \underbrace{(1 - \lambda_0) \int\limits_{\iota_9} \{V^n + \iota_9\} dI}_{\text{no offer}} \Bigg\}. \tag{1.2}
$$

Reading from left to right, an unemployed worker receives a total value of nonemployment of b, which includes both unemployment benefits as well as the value of non market 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.

1.1.2 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 estimation of the utilities.

The estimation procedure is nonparametric 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 1.9 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 nonemployment) in terms of parameters of the model. With this value in hand, the model reduces to an overidentified (by one equation) system of equations. I can then unravel this central tendency into the underlying structural parameters.

1.1.2.1 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.

When an employer contracts by a lot, the displaced worker literature claims the employer contraction caused the worker separations. To see this graphically, consider Figure 1.1. The figure depicts employer-to-employer separation probabilities as a function of employer growth (and is inspired by Davis, Faberman, and Haltiwanger (2012, Figure 6), which uses employer-side survey evidence on the reason for separation; Flaaen, Shapiro, and Sorkin (2015) displays a similar hockey-stick shape for a worker-side survey). 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 labels all separations to the left of this as caused by the employer contracting and all separations to the right as not.

The continuous notion of displacement assigns a displacement probability as the excess probability of separation at a contracting employer relative to at an expanding employer. To take a simple numerical example, suppose 7% of the workforce typically separates each year (and the firm typically hires 7% of its workforce), and in one year we observe 36% of the workforce separate. Assuming hiring stays constant, the employer contracts by 29%. The Jacobson, LaLonde, and Sullivan (1993) approach would label none of these separations as displacements. But $81\% \approx \frac{36\%-7\%}{36\%}$ of the separations only happened because of a presumed firm-level shock and so are displacements. Conversely, $\approx 19\%$ of the separations were endogenous and would have happened even in the absence of a firm-level shock.

I use the displacement probabilities to assign displacement, or exogenous, weights to each transition in the data. To continue with the example, suppose all the transitions are employer-to-employer. Each of these employer-to-employer transitions counts as 81% of an exogenous move, and 19% of an endogenous move. 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.[4]

1.1.2.2 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 non-displaced workers flowing to employer i from employer j. In the model, workers receive one offer at a time and so only ever make binary choices. Adopting the standard continuum assumption in discrete choice models, such flows from employer j to employer i are given by

$$\underbrace{M_{ij}}_{\text{j to i flows}} = \underbrace{g_j W}_{\text{\# of workers at j}} \underbrace{(1 - \delta_j)(1 - \rho_j)}_{\text{no shocks}} \underbrace{\lambda_1 f_i \Pr(i \succ j)}_{\text{get offer from i and accept}}, \tag{1.3}$$

where W is the number of employed workers. To interpret this equation, 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 gives rise to 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:

$$\frac{M_{ij}}{M_{ji}} = \frac{f_i g_j (1 - \delta_j)(1 - \rho_j) \Pr(i \succ j)}{f_j g_i (1 - \delta_i)(1 - \rho_i) \Pr(j \succ i)}. \tag{1.4}$$

It is helpful to use the type I extreme value distribution to simplify $\Pr(i \succ j)$. Specif-

---

[4]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 unemployment to estimate exogenous shocks (i.e. Postel-Vinay and Robin (2002, pg. 2319)), or uses the exogenous shocks to explain unexplained mobility (for example, 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 worker's choices.

ically, the type I extreme value distribution implies the differences in the error terms are distributed logistic so that

$$\Pr(i \succ j) = \underbrace{\frac{exp(V^e(v_i))}{exp(V^e(v_i)) + exp(V^e(v_j))}}_{\text{logistic (difference of type I errors)}} \Rightarrow \frac{\Pr(i \succ j)}{\Pr(j \succ i)} = \frac{exp(V^e(v_i))}{exp(V^e(v_j))}. \qquad (1.5)$$

Combining (1.4) and (1.5) gives:

$$\underbrace{\frac{M_{ij}}{M_{ji}}}_{\text{relative flows}} = \underbrace{\frac{f_i}{f_j}}_{\text{relative offers}} \times \underbrace{\frac{g_j}{g_i}}_{\text{relative size}} \times \underbrace{\frac{(1-\delta_j)(1-\rho_j)}{(1-\delta_i)(1-\rho_i)}}_{\text{relative shocks}} \times \underbrace{\frac{exp(V^e(v_i))}{exp(V^e(v_j))}}_{\text{relative values}}. \qquad (1.6)$$

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.[5] Now introduce notation which defines the flow-relevant firm-level value that summarizes the determinants of relative flows. Define

$$\underbrace{exp(\tilde{V}_i)}_{\text{flow-relevant value}} \equiv \underbrace{\frac{f_i exp(V^e(v_i))}{g_i(1-\delta_i)(1-\rho_i)}}_{\text{offer} \times \text{value / effective size}}. \qquad (1.7)$$

$exp(\tilde{V}_i)$ 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:

$$\frac{M_{ni}}{M_{in}} = \frac{(1-\lambda_1)g_i W(1-\delta_i)(1-\rho_i)\Pr(n \succ i)}{\lambda_0 f_i U \Pr(i \succ n)} = \frac{exp(\tilde{V}_n)}{exp(\tilde{V}_i)}, \qquad (1.8)$$

where

$$\underbrace{exp(\tilde{V}_n)}_{\text{flow-relevant value}} = \underbrace{\frac{(1-\lambda_1)W exp(V^n)}{\lambda_0 U}}_{\text{offers} \times \text{values / effect size}}. \qquad (1.9)$$

Reading across, the offers to nonemployed workers occur when a worker does not get an outside offer, and the effective size of the nonemployed pool for valuation purposes

---

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

is the number of workers who have offers.

Appendix 1.9 shows how to take the employer-level value defined in equations (1.7) and (1.9) and infer the underlying values of employers and nonemployment. From the calibration step, I know $\{g_i, \delta_i, \rho_i, W\}$. The key remaining object to estimate 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-employment 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.

To estimate the firm-level value, I now show that the model implies a set of linear restrictions. Let $\mathcal{E}$ be the set of employers and n be the nonemployment state. Combining (1.6) and (1.7) gives relative flows between employers in terms of $exp(\tilde{V}_i)$:

$$\underbrace{\frac{M_{ij}}{M_{ji}}}_{\text{relative flows}} = \underbrace{\frac{exp(\tilde{V}_i)}{exp(\tilde{V}_j)}}_{\text{relative flow-relevant values}} . \tag{1.10}$$

Cross-multiplying (1.10) 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 (1.10) goes through for all employers (as well as nonemployment). Summing across all employers on both sides gives:

$$\underbrace{\sum_{j \in \mathcal{E}+n} M_{ij}}_{\text{\# of hires by i}} \underbrace{exp(\tilde{V}_j)}_{\text{flow-relevant value}} = \underbrace{\sum_{j \in \mathcal{E}+n} M_{ji}}_{\text{\# of exits from i}} \underbrace{exp(\tilde{V}_i)}_{\text{flow-relevant value}} . \tag{1.11}$$

Divide by the summand on the right hand side:

$$\frac{\overbrace{\sum_{j \in \mathcal{E}+n} M_{ij}exp(\tilde{V}_j)}^{\text{value weighted hires}}}{\underbrace{\sum_{j \in \mathcal{E}+n} M_{ji}}_{\text{exits}}} = \underbrace{exp(\tilde{V}_i)}_{\text{flow-relevant value}} . \tag{1.12}$$

Equation (1.12) implies one linear restriction per firm (and one for nonemployment).

The equation generates a recursive definition of employer quality: the quality of an employers depends on the quality of employers it hires from, which in turn depends on the quality of employers it hires from. In 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 (1.12). Specifically, define a square matrix S which is all zeros off-diagonal, and the ith diagonal entry is $S_{ii} = \sum_{j \in \mathcal{E}+n} M_{ji}$. Then letting $exp(\tilde{V})$ be the $|\mathcal{E}+n| \times 1$ vector that contains the firm-level $exp(\tilde{V}_i)$ and $exp(\tilde{V}_n)$

$$\underbrace{S^{-1}M}_{\text{normalized flows}} \quad \underbrace{exp(\tilde{V})}_{\text{flow-relevant values}} \quad = \quad \underbrace{exp(\tilde{V})}_{\text{flow-relevant values}} \quad . \tag{1.13}$$

This equation allows me to solve for $exp(\tilde{V})$. Intuitively, $exp(\tilde{V})$ is the fixed point of the function $S^{-1}M : \mathbb{R}^{|\mathcal{E}+n|} \to \mathbb{R}^{|\mathcal{E}+n|}$. 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.[6] Moreover, in order for the values to be interpretable I need that the $exp(\tilde{V})$ vector is all positive so that the $\tilde{V}$ is defined (the log of 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 1.8 (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

---

[6] In many other contexts the (top) eigenvector of matrices have been shown to have interesting economic content. Some examples include: the Leontief inverse of the input-output table, 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).

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 (1.12). 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 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. S divides the ith *row* of M by the ith column sum of M. In other applications, i.e. Pinski and Narin (1976), Page et al. (1998) (Google's PageRank) and Palacios-Huerta and Volij (2004), the normalizing matrix instead divides the ith *column* of M by the ith 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 1.8. □

1.2   Matched employer-employee data

This section describes the U.S. Census Bureau's Longitudinal Employer Household Dynamics (LEHD) data, which is a quarterly data that is constructed from

unempoyment insurance records.[7] The LEHD is matched employer-employee data and so allows me to follow workers across firms.

1.2.1 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.[8] Second, only employers that are covered by the unemployment insurance system appear in the dataset.[9] 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.[10] Thus, variation in hours as well as 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), the assumption is that large groups of workers moving from employer A to 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 employer B is a relabelling of employer A, or else employer B acquired employer A, and so I do not count this as an employer-to-employer transition.

---

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

[8]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 that 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.") On the other hand, 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.

[9]This restriction results in the exclusion of certain sectors of the economy. In particular, small nonprofits (those employing fewer than four workers), domestic, self-employed, some agricultural workers and federal government (but not state and local government) 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`.

[10]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)).

I pool data from 27 states from the fourth quarter of 2000 to the first quarter of 2008.[11] 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.

1.2.2 Dataset construction

I treat the model as 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.[12] I also restrict attention to workers aged 18-61 (inclusive) and, following Card, Heining, and Kline (2013), require that the annualized earnings exceed \$3250 (in 2011 dollars). With an annualized 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-nonemployer-to-employer. The transition is coded as employer-to-nonemployer-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 displacement weight (see section 1.1.2.1). See appendix 1.10 for details on dataset construction, and appendix 1.11 for a discussion of computation.

---

[11]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 appendix Figure 1.15 for a map.

[12]Reduction to one observation per person per year is common. See e.g. 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. (Forthcoming) also reduce to one such observation per year.

1.2.3 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 1.1, 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 (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 from column (1) to (4) exist for 7 years, they had 1.5 people per year on average. The second portion of Table 1.1 shows that the mean and variance of earnings are quite stable as I lose person-years and employers.

1.3 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.

1.3.1 Exogenous probabilities

The data are very similar to the stylized Figure 1.1 that I used to motivate and explain the method to identify displaced workers probabilistically.

Figure 1.2 is consistent with the assumption in Section 1.1.2.1 (and Figure 1.1) that separation probabilities are flat as employers expand. Figure 1.2a shows the overall employer-to-employer and employer-to-nonemployment probabilities as a function of employer growth. Looking at the right hand side of the graph, the separation probabilities are reasonably flat.

The left hand side of Figure 1.2a supports the view that employer-level contractions cause separations. The figure shows that both the employer-to-employer and the employer-to-nonemployment probabilities rise as employers contract. This suggests that the contraction causes the separation.[13] Interestingly, the employer-to-

---

[13]This finding is not quite mechanical: it could be that employers contract by cutting hiring. In some sense it is surprising that there is not more evidence of this margin of adjustment.

nonemployment probability rises much more rapidly than the employer-to-employer probability.

The inference from administrative data that workers are being laid-off 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 employer-reported reasons for separations as a function of firm growth rates. They find patterns of employer-reports of quits and layoffs that are similar to the employer-to-employer line and employer-to-nonemployment lines in Figure 1.2a. Flaaen, Shapiro, and Sorkin (2015) compute worker-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 separations rise most rapidly.

Turning to the bottom panel, Figure 1.2b displays the displacement weight attached to an employer-to-employer and employer-to-nonemployment transition as a function of the employer growth rate. The displacement probabilities rise rapidly as the employer contracts. (The displacement probability among expanding employers is always zero by construction.)

Figure 1.3 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 model estimates of 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.[14] 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.

1.3.2 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

---

[14]These pictures are constructed slightly differently than Figure 1.2a. Figure 1.2a uses quarterly growth rates because to construct the displacement weights I want to know what was happening in the quarter the worker separating. In contrast, Figure 1.3 uses the growth rates from 2001 to 2007 because in estimating the model I assume that the economy is in steady state from 2001 to 2007 and so the firm value is constant over the period. Hence, to assign a single growth per employer I use the overall growth rate.

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, it is worth keeping in mind that 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 transitions between annual dominant employers are employer-to-employer. For workers in column (1) of Table 1.1, I code each transition between annual dominant jobs as employer-to-employer or employer-to-nonemployment-to-employer. The top rows of Table 2.3 shows that the annual separation probability is 25% and only 40% of these are employer-to-employer.[15]

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, this exogenous share is 48%.[16]

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, 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.[17]

---

[15]To compute a probability, the denominator is the number of person-years less one per worker. The level of transitions is slightly lower—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.

[16]This exogenous share is between the estimates of Sullivan and To (2014, pg. 482, Table 1) that about 15% of employer-to-employer transitions are reallocation and 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. Unlike Moscarini and Postel-Vinay (2014), I do not use worker outcomes to separate employer-to-employer transitions into exogenous and endogenous.

[17]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 the separation might

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 displacement weights and the separation probabilities, gives an annual displaced employer-to-nonemployment, or job destruction, rate (δ) of 0.05.[18] Combining the displacement weights and the separation probabilities, gives an annual displaced employer-to-employer, or reallocation, rate (ρ) or 0.03.[19] The remaining rows of Table 2.3 contain the other model parameters.[20]

1.3.3 The slippery ladder

Figure 1.4 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 variation depicted in Figure 1.2. The figure shows that job destruction and reallocation shocks are more likely at the worst firms. 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 (2014).

1.3.4 Assessing the fit of the search model

The search 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.

---

be labelled by the firm as a layoff.

[18]This annual job destruction rate is quite similar to Bagger and Lentz (2014, pg. 34) who estimate the annual job destruction rate for the low type (which dominates their data) at 0.063. On the other hand, 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 iid way to annual values.

[19]The reallocation rate is quite a bit lower than the estimate of Bagger and Lentz (2014) of 0.106, but again, I do not use outcomes to code the reallocation rate.

[20]I estimate the arrival rate of offers on the job, λ_1, by matching the level of non-displaced 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.)

I first consider whether the model can match both the slope of the probability of making an employer-to-employer transitions 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 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 search model is able to match the probability of endogenous employer-to-employer transitions across employers. Figure 1.5a 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 equal-person-year-size bins. For each firm, I then compute the model-implied probability of an endogenous employer-to-employer separation. Within 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% 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 1.5b 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 (i.e. one from an expanding firm) varies by a factor of about 10 from the bottom 5% to the top 5% of firms, which nearly exactly matches the patterns 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 $\tilde{V}_A > \tilde{V}_C$.[21] The local information is contained in binary comparisons. A binary comparison of employer A and C occurs when I observe accepted offers from A to C

---

[21]That is, I focus on the information in the accepted offers to parallel the binary comparisons. I use the values defined in equation (1.7).

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 1.6. In this figure, C wins on the binary comparison over A ($2 > 1$). But when the accepted offers related to 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 broadly 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).[22] I find that the 90 percent confidence interval under the random null is [49.67%, 50.28%]. Under the null that the model is the data generating process the 90 percent confidence interval is [75.38%, 75.56%].[23] 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.

1.3.5   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 1.4). Strikingly, however, there is more dispersion when measured in terms of earnings. This section also shows that patterns of left-shifts in the offer distribution predicted by search theory maintain in both value and earnings space.

Figure 1.10a 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

---

[22]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 eqully appealling.

[23]I repeat the procedure described in footnote 22 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.

(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 idiosyncratic draws are similar. Looking only at 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 1.10b 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). Comparing the two panels, what is clear is that the common component is relatively more important in terms of earnings than in terms of values. This feature says that the systematic patterns in mobility are weaker than the systematic patterns in earnings changes.

Figure 1.11 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). Note that the earnings result is not mechanical since the earnings information is not used in the estimation of the search model. In addition, the figures compare where workers who are hired from nonemployment and on reallocation shocks are hired. 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.

1.4   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 the task of estimating a notion of the earnings posted by each employer. In section 1.5, I show how to combine the values and earnings to decompose the variation in firm-level earnings into a rents and compensating differentials component.

1.4.1 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):

$$\underbrace{y_{wt}}_{\text{log earnings}} = \underbrace{\alpha_w}_{\text{person effect}} + \underbrace{\Psi_{\boldsymbol{J}(w,t)}}_{\text{firm effect}} + \underbrace{x'_{wt}\beta}_{\text{covariates}} + \underbrace{r_{wt}}_{\text{residual/error term}} , \tag{1.14}$$

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.[24]

Once I use the search model to selection-correct the equation, the firm effects in equation (2.1), Ψ, are the model-consistent notion of earnings, where the Ψ is the same firm-level earnings as discussed in section 1.1. The firm effects are identified by workers who switch firms. As such, the firm effects remove a time-invariant worker effect and so captures the earnings at a firm shared by all workers.[25] 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 (2.1) using movers requires that workers do not move on the basis of the error term. To remedy this inconsistency, I selection-correct equation (2.1) by inserting the expected value of the idiosyncratic utility draw calculated from the search model. Appendix Table 1.16 shows that this does not affect the firm effects. See appendix 1.12 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:

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

$$\tag{1.15}$$

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

$$\frac{\text{Cov}(\Psi_{\boldsymbol{J}(w,t)}, y_{wt})}{\text{Var}(y_{wt})}. \tag{1.16}$$

---

[24]Because I only use 7 years of data, the linear terms in the age-wage profile are highly correlated with the person fixed effects and so, following Card, Heining, and Kline (2013), are omitted.

[25]One implementation difference from what is standard is that typically researchers use both employer-to-employer and employer-to-nonemployment-to-employer movers to estimate the firm effects. Table 1.16 shows that restricting to just employer-to-employer transitions barely changes the firm effects (the correlation with the benchmark is 0.96).

Firms play an important role in earnings determination. The third portion of Table 1.1 performs the Abowd, Kramarz, and Margolis (1999) decomposition (equation (1.15)). About 22-3% of the variance of earnings is explained by employer-level heterogeneity.[26]

1.4.2 Earnings cuts are an important feature of the data

This section shows that seemingly voluntary earnings cuts 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 pursuit of higher-pay explains some employer-to-employer moves. In section 1.5, I directly relate the average direction of mobility in the labor market to the average pattern in earnings changes.

Earnings cuts are widespread in the data.[27] Table 1.3 shows that 42% of changes between annual dominant employers take earnings cuts. Even restricting to transitions that are most likely to be voluntary—employer-to-employer, weighting separations from expanding employers more heavily—37% of transitions take earnings cuts.

The firm effects capture the probability of an earnings cuts. Table 1.3 shows that 52% of the most likely to be voluntary moves to lower-firm-effects firms have earnings cuts, while only 26% of the moves to higher-paying firms have earnings cuts. Figure 1.7 shows this fact graphically. Figure 1.7a plots the change in firm effects against the probability of taking an earnings cut on an employer-to-employer transition, while panels 1.7b show this for employer-to-nonemployment-to-employer transitions. The

---

[26]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) also shows that changing the sample does not affect the decomposition.

[27]Earnings cuts are widespread in any dataset. See, for example, Jolivet, Postel-Vinay, and Robin (2006). Tjaden and Wellschmied (2014) argue that such earnings cuts are explained by reallocation shocks.

probability of an earnings cut is 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 $\beta_1 = 1$.[28] The x-axis of Figure 1.8 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 we would expect to see 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 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, and this result is not predicted by models where mobility is on the basis of comparative advantage. A natural concern is that the firm effects summarize the information in the individual earnings changes and so the fit apparent in Figure 1.8 is mechanical. But this fit does not arise on data simulated from models where mobility is on the basis of comparative advantage in two ways. First, in models where mobility is on the basis of comparative advantage (i.e. matching models) there are no earnings cuts, and so the individual level earnings changes always lie above the x-axis. Second, in models where mobility is on the basis of comparative advantage 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). Appendix Figure 1.16 plots the analogous figure to Figure 1.8 with data simulated from the example production

---

[28]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 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, while on people 18-34 and 35-61 it is 0.86.

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.

The earnings cuts captured by moving to lower-firm effects firms are not offset by future earnings increases. One explanation for earnings cuts is that workers accept earnings 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 1.9 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 is calibrated to match the share of earnings cuts in the data. In the simulated data, the correlation between the firm effects in intercept and slope is -0.90. This result implies that earnings cuts are not explained by the possibility of future earnings increases at the same firm.

1.5 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.

1.5.1 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:

$$\underbrace{V_i^e}_{\text{forward-looking values}} = \underbrace{\omega}_{\text{value per log dollar}} \times (\underbrace{\Psi_i}_{\text{earnings}} + \underbrace{a_i}_{\text{amenities/nonpecuniary}}), \qquad (1.17)$$

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 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 to avoid 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).

The following result characterizes what can be identified using the patterns of utilities and earnings that firms post identifies, and is the basis of how I measure compensating differentials and frictions.

Result 3. *Suppose the utility function is given by equation (1.17), and $\{V_i^e\}_{i\in\mathcal{E}}$ and $\{\Psi_i\}_{i\in\mathcal{E}}$ are known. Then*

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

where $R^2 = Corr(V^e, \Psi)^2$. In these limits:

$$Var(\Psi + a) \in [Var(\Psi)R^2, \infty),$$
$$Corr(\Psi, a) \in [-\sqrt{1 - R^2}, \sqrt{R^2}], \ and$$
$$Cov(\Psi, a) \in [-(1 - R^2)Var(\Psi), \infty].$$

Proof. See Appendix 1.8. □

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.

To understand why the second term in the decomposition is compensating differentials, it is helpful to consider the infeasible decomposition of the nonpecuniary part of firm value into two components: $a = a_{\text{rosen}} + a_{\text{mortensen}}$.[29] The a_{rosen} component reflects compensating differentials. It is the lower bound of the variance of nonpecuniary characteristics, $Var(a_{\text{rosen}}) = Var(\Psi)(1-R^2)$. A regression between earnings and a_{rosen} finds a willingness to pay to *avoid* one unit of this component of nonpecuniary characteristics of 1 log dollar.[30] The $a_{\text{mortensen}}$ component reflects the possibility that some firms offer not only high levels of pay but also, on average, high levels of nonpecuniary characteristics; for example, a high-paying firm might also offer great benefits and a nice work environment. The correlation of $a_{\text{mortensen}}$ with Ψ is at the upper bound given in the result, $Corr(\Psi, a_{\text{lang}}) = \sqrt{R^2} = Corr(V^e, \Psi)$. But the variance of $a_{mortensen}$ is not identified: it lies in $[0, +\infty]$.

Because the variance of $a_{mortensen}$ 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.[31] And third, I can only provide a (informative) lower bound on the variance of utility offered by firms.[32]

---

[29] 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 on average benefits are positively correlated with earnings.

[30] Consider $\hat{\beta}$ estimated from $\Psi = \beta a_{\text{rosen}}$. Then $\hat{\beta} = \frac{Cov(\Psi, a_{\text{rosen}})}{Var(a_{\text{rosen}})} = \frac{-(1-R^2)Var(\Psi)}{(1-R^2)Var(\Psi)} = -1$.

[31] This fact also shows that measuring the correlation between earnings and nonpecuniary characteristics is not informative about the role of compensating differentials in the labor market.

[32] One might hope that I could use knowledge of the variance of utility to pin down $Var(a_{\text{mortensen}})$. The reason I cannot is because I do not know how the units on utility relate to the units on earnings, or α. To see why this is a problem, note the variance in utility in log dollar units is given by $\frac{Var(V)}{\alpha^2}$. For a given empirical value of $Var(V)$, the variance in log dollar units can be made arbitrarily large by sending $\alpha \to 0$. (The result limits how arbitrarily small I can make this variance.)

1.5.2 Sectoral-level evidence

This section discusses the utilities and earnings that firms post aggregated to the sector level.[33] At this level of aggregation, we can assess the intuitive plausibility of the results of the model.

Table 1.4 shows that there are similarities between these rankings, which is evidence of rents. Column (1) ranks sectors on the basis of the utilities and column (2) ranks sectors on the basis of earnings. 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.

On the other hand, there are also differences between these rankings, which is evidence for compensating differentials. The most striking sector is education. It is one of the best sectors in terms of utility and one of the worst in terms of earnings. This implies the presence of good nonearnings characteristics in education. Similarly, public administration is much higher ranked in terms of utility than earnings. In contrast, traditionally blue-collar and male sectors—mining, manufacturing, and construction—tend to be ranked higher in terms of earnings than utility, which implies the presence of bad nonpecuniary characteristics.[34]

Figure 1.12a shows the quantitative version of the alignment between sector-level values and earnings. It shows the scatterplot of the sectoral-level values and earnings, as well as the best-fitting line. The R^2—on an employment-weighted basis—is 0.45. Thus, 45% of the inter-sectoral wage structure is rents and 55% is compensating differentials.

1.5.3 First-pass answer

Section 1.5.1 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 earnings and utility. The first row of Table 1.5 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 1.12b shows that the relationship between the firm-level values and earnings is approximately linear. It shows a binned scatterplot of the firm-level values

---

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

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

and earnings as well as the line of best fit on the firm-by-firm data, where I have grouped firms into twenty equally-employment-weighted bins (vigintiles). The approximate linearity of the relationship implies that the correlation (or R^2) captures the relationship and richer ways of relating the values and earnings would not yield a tighter relationship.[35]

The figure also illustrates the identification result. The upward slope in the line of best fit is the variation in earnings that is reflected in the variation in values and reflects rents. The figure also shows the variation in firm-level earnings holding utility constant: it plots plus and minus one standard deviation bands of the firm-level earnings within a firm-level utility bin. This variation in earnings is not—in equilibrium—reflected in utility and so must compensate for amenities: the firms above the line have relatively bad amenities and the firms below the line have relatively good amenities.

1.5.4 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 variable, 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 my estimates of the firm-level earnings and values have converged. This means that at bigger firms the firm-level values and earnings are more precisely estimated.

Figure 1.13a 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, where each bin represents

---

[35]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; 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 predicts mobility patterns). It is also important to note that the transition matrix—and $\tilde{V}$—only contains information in accepted offers.

the same number of person-years (so that the bins on the left hand side of the graph have many more firms than bins on the right hand side). Within each bin, I compute the correlation between the values and earnings. Consistent with the law of large numbers logic, the R^2 rises rapidly as I consider larger firms and then is flat among firms with between about 50 and 1800 workers per years (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 based on common characteristics. Because the quantity I am interested in is the R^2 in a regression, I need to account for measurement error in the left-hand side and right-hand side variable simultaneously. I group firms by characteristics which predict both earnings and nonpecuniary characteristics, and then ask how related these group-level averages are. 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. See appendix 1.13 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.[36] 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 sorted on the basis of firm size discussed above.

Table 1.15 shows that these groupings together explain about 60% of the variance of firm-level earnings.[37] Unconditionally, about 12% of the variance of earnings is at the county-level, 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

---

[36]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).

[37]Table 1.14 shows that there are many employer-to-employer transitions across these boundaries.

values.[38]

Table 1.5 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 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 and 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 rents and $0.12 \times 0.96 \approx 0.11$ is 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 and so I present bounds. This component accounts for about 40% of the variance of firm-level earnings. Figure 1.13b shows how the R^2 varies by firm size bin having removed the common component. As expected, because I have extracted much of the signal in the values and earnings, the R^2 are 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 and so the R^2 reflects the true relationship. To get a lower bound, I take the maximum value of the R^2 in the picture (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

---

[38]The table 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.

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

Appendix 1.14 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. The appendix also shows that aggregating across industries and locations does not do too much damage to the data.

1.6 Discussion and implications

1.6.1 Compensating differentials

This paper finds that compensating differentials explain about 15% of the variance of individual-level earnings in the U.S. economy.[39] While one should interpret this point estimate cautiously given the numerous strong assumptions it took to reach it, the take-away of this analysis is that compensating differentials are important. 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 sectoral-level analysis.

The importance of compensating differentials contrasts with the conventional wisdom summarized by Hornstein, Krusell, and Violante (2011, pg. 2883) that compensating differentials do "not show too much promise" in explaining earnings dispersion. Their pessimism about compensating differentials comes from the vast literature estimating willingness to pay for particular amenities—e.g. risk of death, risk of separation, or health insurance. With a few exceptions, this literature has not found robust evidence that compensating differentials are important. An important assumption in the one-amenity-at-a-time approach, however, is that firms compensate workers for particular amenities through pay, rather than through variation in other amenities. The fact that this literature has studied more than one amenity suggests that there is scope for this assumption not to hold. 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

---

[39]This number comes from the following calculation. According to Table 1.1, firms account for 22% of the variance of earnings. According to Table 1.5, 69% of the variance of firm-level earnings is compensating differentials. And $0.22 \times 0.69 = 0.15$.

number of voluntary earnings cuts had not been interpreted as evidence of the importance of compensating differentials. The difference is that this paper operates at the firm-level rather than the individual level. The intuition of the identification result is that compensating differentials are 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.[40] I take the preferred estimates from these papers and compute the implied relationship between behavior and pay.[41] 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.[42] 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 differentials, and this likely reflects idiosyncratic factors.

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

---

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

[41]I take the preferred estimates from these papers 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 1.16 for details).

[42]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 1.18, 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).

paper answers a different question related to the role of nonpecuniary characteristics than this paper. The calculation in this paper 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 the question: "what would the variance of earnings be if people only valued money?"[43] 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.

1.6.2 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 that the model uncovers 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 firms in explaining mobility.

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. This tradition 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

---

[43]They 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.

wage 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 in the labor market.

1.6.3 Ability of search models to match earnings dispersion

This paper finds that search frictions by themselves have a hard time explaining firm-level earnings dispersion. This message is consistent with Hornstein, Krusell, and Violante (2011). They argue that benchmark search models have a hard time rationalizing the extent of earnings dispersion—measured as the residual in a Mincerian regression using Census data—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 that rents do not explain all the firm-level dispersion in earnings is the finding of systematic patterns of employed workers making employer-to-employer transitions to lower-paying firms. As such, 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.

Unlike Hornstein, Krusell, and Violante (2011), this paper argues that an important part of why frictions do not explain all earnings dispersion is because compensating differentials are quantitatively large.

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 unemployment relative to the average value of a job. One statistic 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.

1.6.4 Inequality

This section combines the implications of the estimates of compensating differentials and rents by considering 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 and so earnings inequality—or the variance of earnings—decreases. On the other hand, to see how eliminating compensating differentials could increase inequality, consider the Professor (of economics) and the Miner. Even given their large differences in non-earnings compensation, the Professor is still better paid than the Miner. Equalizing the pleasantness of their work leads the difference in their earnings to be even more dramatic and so earnings inequality increases.

To measure the effect on inequality of pricing out a_{rosen}, I use the identification result in section 1.5.1. V^e represents dispersion in utility and so 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.[44] Then I recompute the variance of individual-level earnings. Because of concerns about measurement error, I do this in stages where I sequentially price out the location piece, the industry and etc.

Table 1.6 shows that at my central estimate, removing amenities and compensating workers would reduce inequality. About half of the effect occurs at the industry-level.

Figure 1.14 shows that this counterfactual has surprising impacts on the structure of earnings. The figure plots the actual distribution of earnings, and the counterfactual distribution at my central estimate. Inequality is reduced primarily by shifting

---

[44]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 $\sqrt{\frac{Var(\Psi)}{Var(V^e)}} R^2$.

in the lower tail of the distribution. This is not what we would expect from the patterns of sorting of workers to firms. Figure 1.14b shows a naive counterfactual. To compute the naive counterfactual, I multiply the firm effects by $1 - R^2$ and recompute the variance of earnings. The naive counterfactual shifts in both the lower and the upper tail of the income distribution. As documented in the last row of Table 1.6, this has a much larger effect on the variance of earnings than what I estimate. The reason for this surprise is that while the earnings potential of workers is positively correlated with rents, it is negatively correlated with a_{rosen} amenities.

1.7 Conclusion

This paper exploits the intuitive notion that workers will move towards firms with higher value to develop a method to estimate a revealed value of essentially every firm in the U.S. economy. It then shows how to combine 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 that compensating differentials and rents are both quantitatively important explanations, but compensating differentials are more important. The intuition for the importance of compensating differentials is that there are systematic patterns of moves towards lower-paying firms. The rankings of sectors are plausible, as is the implied distribution of compensating differentials. The main finding is robust across a number of subgroups.

The ideas and methods in this paper potentially have many applications. The empirical methodology could be used to study a variety of other questions. In Sorkin (2015b) 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 1.14 suggests that there are interesting differences by age. The firm-level moment also suggests a simple way of studying the direction of reallocation in the labor market, a question which has recently received attention due to the work of Moscarini and Postel-Vinay (2013). Figure 1.10a suggests a new way to assess the extent of frictions in labor markets; it might be that frictions differ by geography (i.e. one benefit of density might be more competitive labor markets). Finally, 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)).

Table 1.1: Summary statistics and the variance of earnings

	All	Connected by EE and ENE	S. Connected by EE	S. Connected by EE (restrictions)
	(1)	(2)	(3)	(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

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 \$3250 (\$2011). The table includes person-years in which on December 31 of the year the person was 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 EE transition, and hire a worker from nonemployment.

Table 1.2: 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

All probabilities and parameters are annual. The sample for the transition probabilities is column (1) of Table 1.1. 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 1.1. 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 1.3: Earnings cuts are common and correlated at the firm level

Pr($y \downarrow$)	All	EN	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

The pay of a firm is defined by its firm effect. This table summarizes moves where a worker had a different dominant employer in consecutive years.

Table 1.4: Ranking sectors

Utility (largest to smallest) (1)	Earnings (largest to smallest) (2)
Utilities	Mining
Public Admin.	Utilities
Education	Information
Mining	Manufacturing
Management	Prof./Scientific/Technical Services
Manufacturing	Management
Finance and Insurance	Finance and Insurance
Wholesale Trade	Wholesale Trade
Information	Construction
Prof./Scientific/Technical Services	Public Admin.
Transport/Warehousing	Transport/Warehousing
Health Care	Health Care
Construction	Real Estate
Agriculture	Agriculture
Other services	Admin/Support/Waste Management
Real Estate	Education
Arts/Entertainment/Recreation	Other Services
Retail Trade	Retail Trade
Admin/Support/Waste Management	Arts/Entertainment/Recreation
Hotels/Restaurants	Hotels/Restaurants

This table uses the sample of firms in column (4) of table 1.1.

Table 1.5: Why do some firms pay so much and some so little?

	Share of variance of firm-level earnings	Share of (1) explained by firm-level utility	Share of (1) explained by:	
			Rents	Comp. diff.
	(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	$(1-.12)\times0.54 = 0.47$	0.45	0.21	0.26
Size	$(1-.12-.47)\times0.04 =$ 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

This table uses firm effects estimated in the set of firms strongly connected by EE transitions. Column (1) reports an order-dependent decomposition of the firms effect 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 6 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. While the next row reports 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 1.13b; version 3 assumes that it is the maximum value in figure 1.13b. Column (3) = Column (1) × Column (2) and is the share of the variance of firm-level earnings that is related to utility. Column (4) = Column (1) × (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 1.6: Implications for inequality

Panel A. Counterfactuals

Scenario	Variance of earnings	Change relative to data
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 less firm effects	Rents	Amenities
Earnings less firm effects	1	0.29	-0.07
Rents		1	0.08
Amenities			1

The sample is the person-years in column (4) of table 1.1.

Figure 1.1: A continuous measure of displaced workers
separation probability: employer-to-employer or employer-to-nonemployment
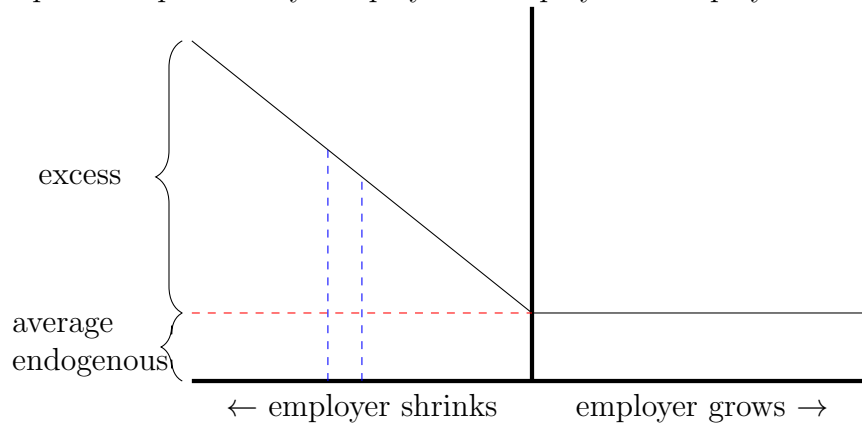
This figure illustrates how I assign displacement probabilities as a function of employer growth. $\Pr(\text{displaced}) = \frac{\text{excess}}{\text{average} + \text{excess}}$, while endogenous is the complement.

Figure 1.2: Constructing displacement rates

(a) Quarterly probabilities
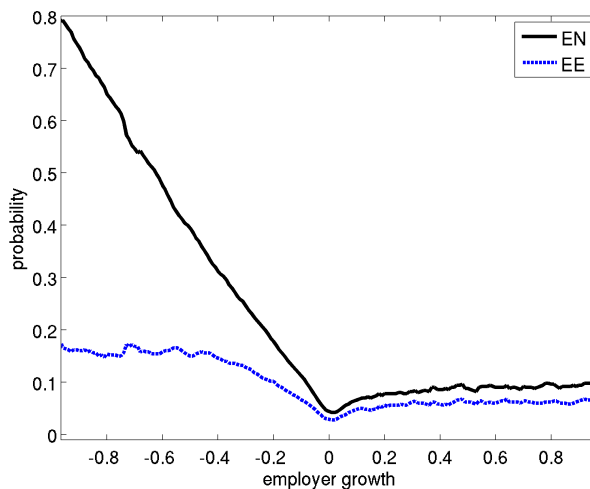
(b) Displacement weights
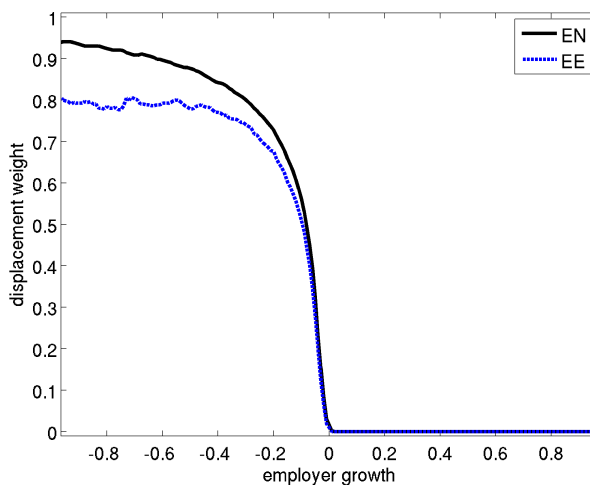
These figures shows the data used to construct the exogenous weights. The probabilities and growth rates are quarterly. The probabilities are computed in one percentage point wide bins of employer growth rates. The figure plots a five bin moving average. The exogenous weight is the probability of an employer-to-employer (or employer-to-nonemployment) transition in a given growth bin minus the average probability of an employer-to-employer (or employer-to-nonemployment) transition at an expanding employer divided by the probability of an employer-to-employer (or employer-to-nonemployment) transition in the growth bin. At expanding employer the exogenous weight is zero by construction.

Figure 1.3: Endogenous   employer-to-employer   and   employer-to-nonemployment
         probabilities by firm growth rate

(a) Employer-to-employer
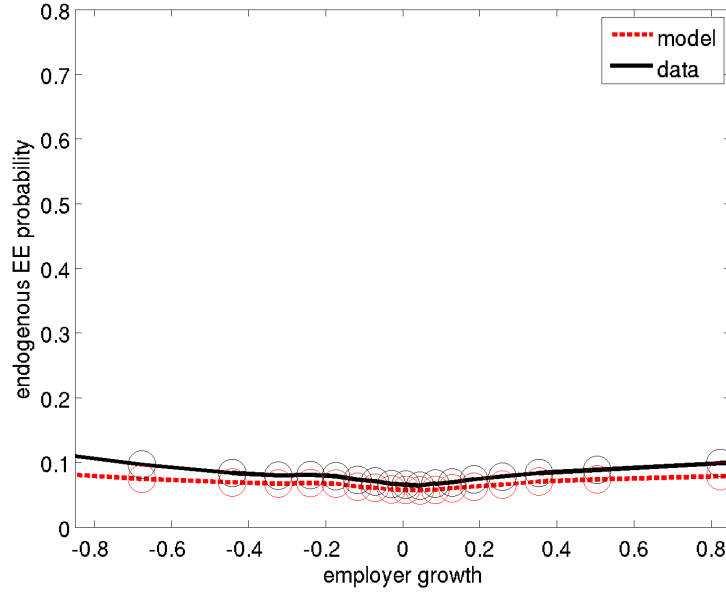
(b) Employer-to-nonemployment
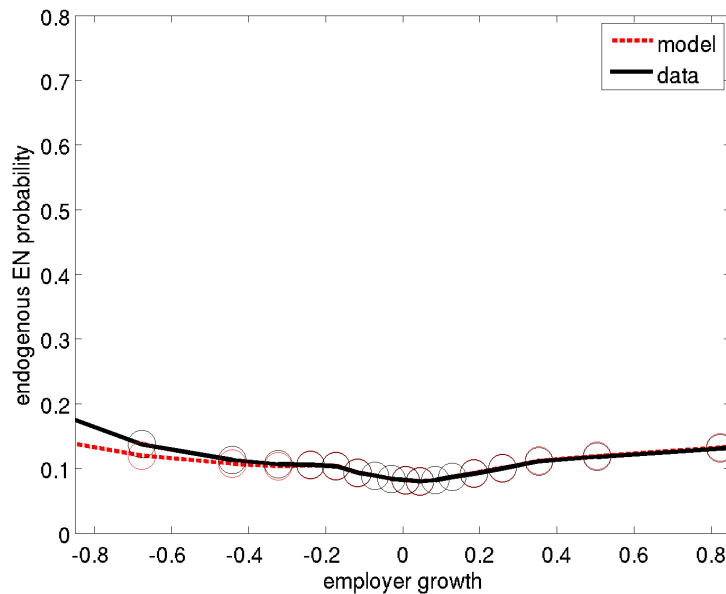
These figures plot the model-implied firm-specific endogenous employer-to-
nonemployment and employer-to-employer probabilities as well as these probabilities
in the data as a function of firm-growth. Firm growth is growth from 2001 to 2007.
To construct the figure, I sort firms into 20 equal person-sized year bins on the basis
of firm-growth. To be consistent with the axes in figure 1.2a this figure only displays
seventeen bins.

Figure 1.4: The "slippery ladder": exogenous shocks by firm value
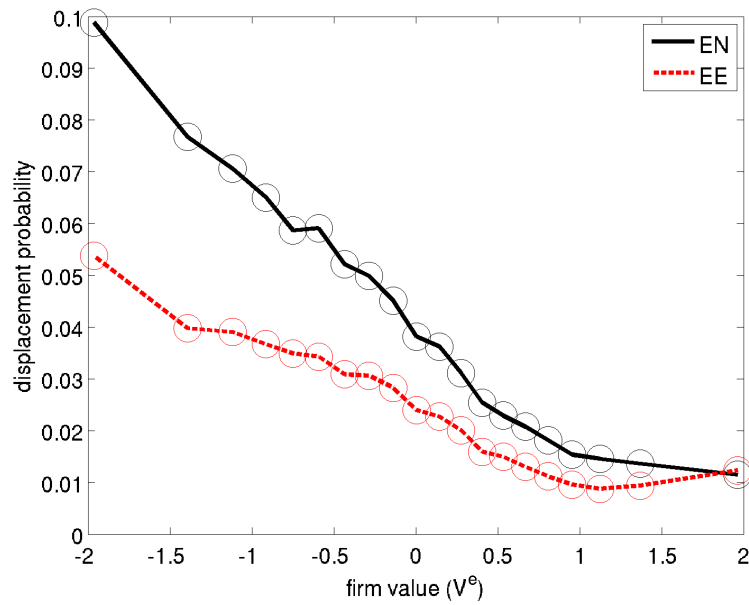
This figure sorts firms into 20 bins on the basis of firm-value. Within each bin, I compute the average probability of each kind of exogenous shock by summing across the exogenous weights in each transition, which are constructed using variation depicted in figure 1.2.

Figure 1.5: Endogenous employer-to-employer and employer-to-nonemployment probabilities by firm value

(a) Employer-to-employer
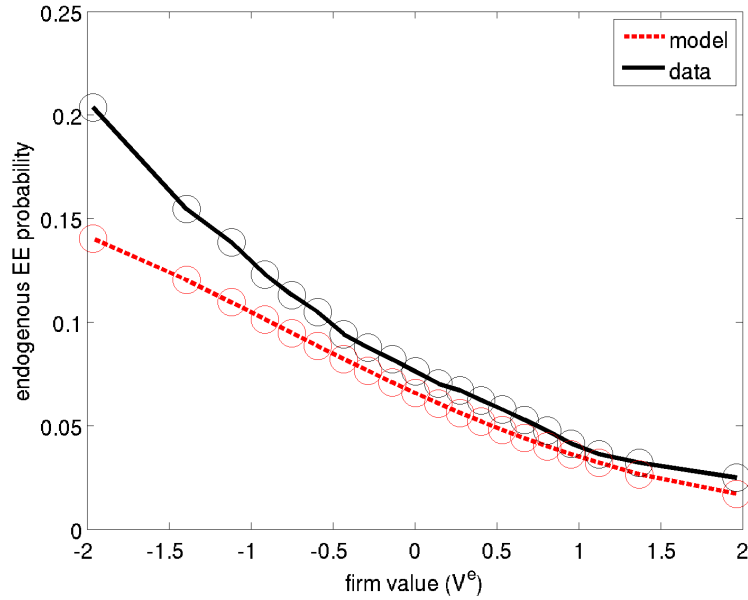
(b) Employer-to-nonemployment
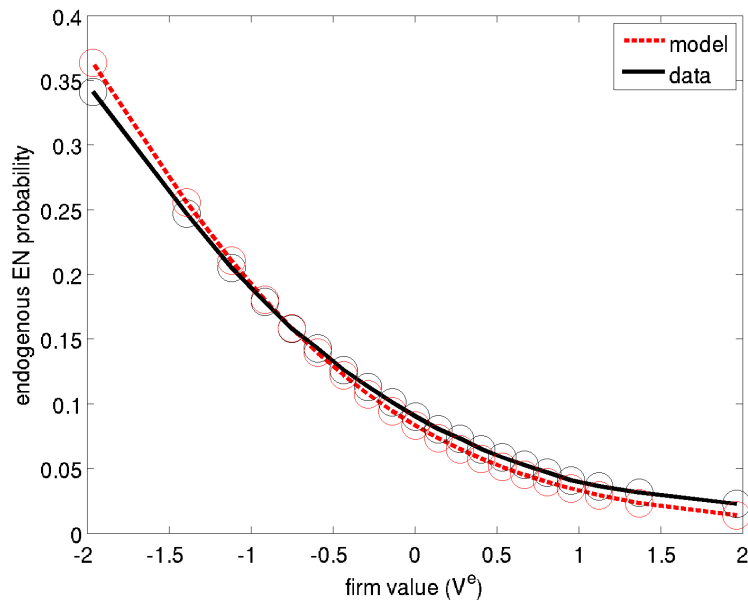
This figure sorts firms into 20 bins on the basis of firm-value. For each firm, I compute the model-implied probability of a voluntary employer-to-employer and employer-to-nonemployment transition. I then take the person-year-weighted average of the model and data within each bin.

Figure 1.6: The model returns a different answer than a binary comparison
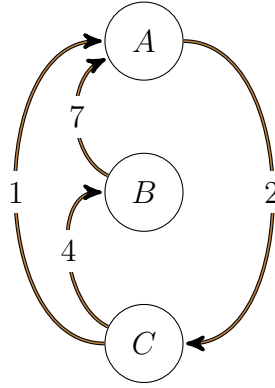
$$M = \begin{bmatrix} 0 & 7 & 1 \\ 0 & 0 & 4 \\ 2 & 0 & 0 \end{bmatrix}$$

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

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

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

Given M, the matrix S is defined as follows: entry $S_{ii} = \sum_j M_{ji}$ (the ith column sum), while the off-diagonal entries are zero. Hence $S^{-1}M$ divides the ith row of M by the ith 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: $exp(\tilde{V}) = S^{-1}M exp(\tilde{V})$.

Figure 1.7: Change in firm pay predicts probability of an earnings cut

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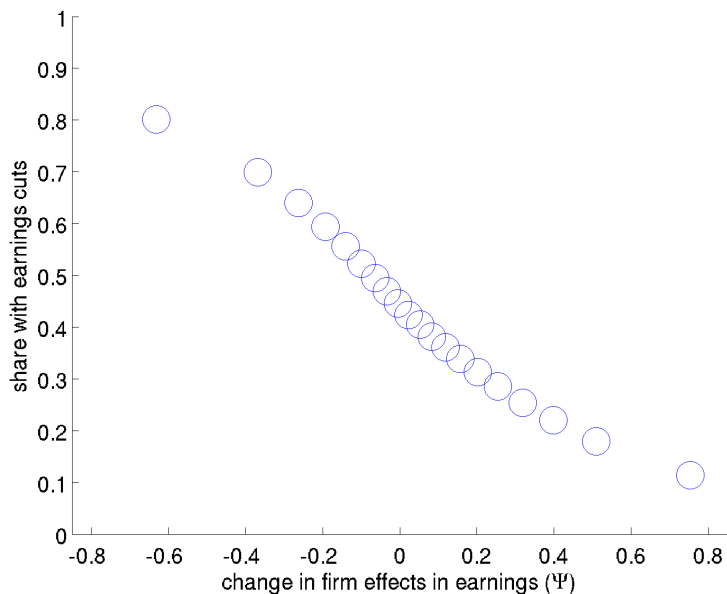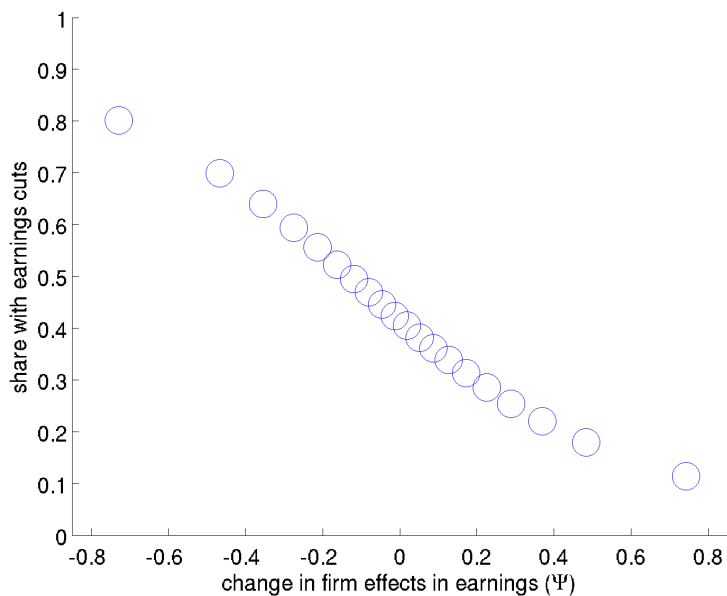
(b) Employer-to-nonemployment-to-employer

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 in the first year at the new job.

Figure 1.8: Change in firm pay predicts magnitude of earnings change

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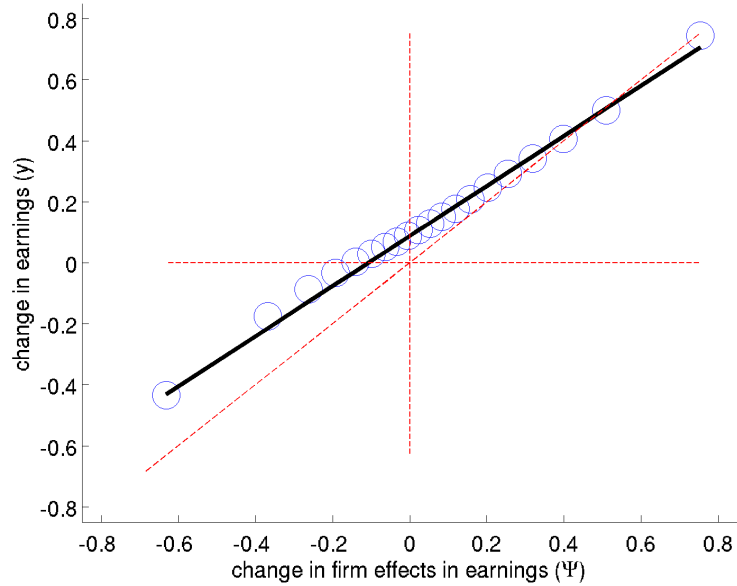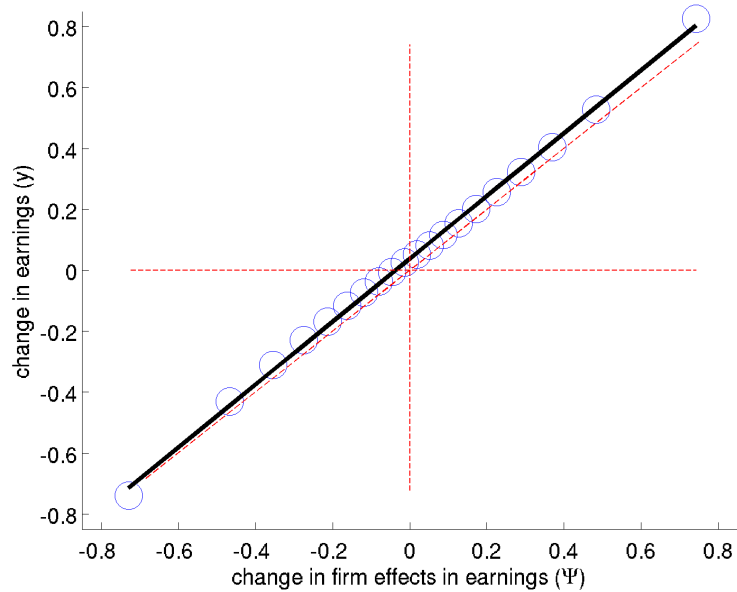
(b) Employer-to-nonemployment-to-employer

This figure considers the sample of workers who switch annual dominant jobs between consecutive years. The circles plot the bin means where I have sorted the job changers into 20 bins on the basis of the change in the firm effects. 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 1.9: Change in firm pay does not predict change in slope of earnings

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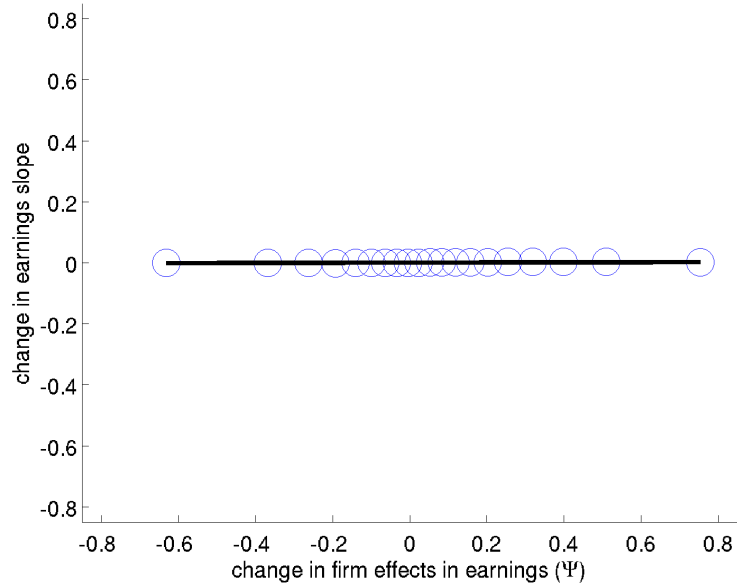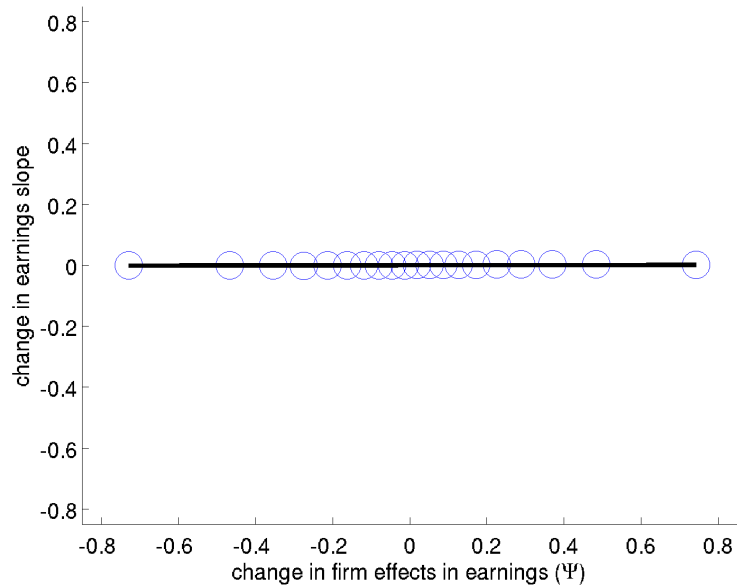
(b) Employer-to-nonemployment-to-employer

This figure considers the sample of workers who switch annual dominant jobs between consecutive years. The circles plot the bin means where I have sorted the job changers into 20 bins on the basis of the change in the firm effects. 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 1.10: Dispersion in the labor market

(a) Values
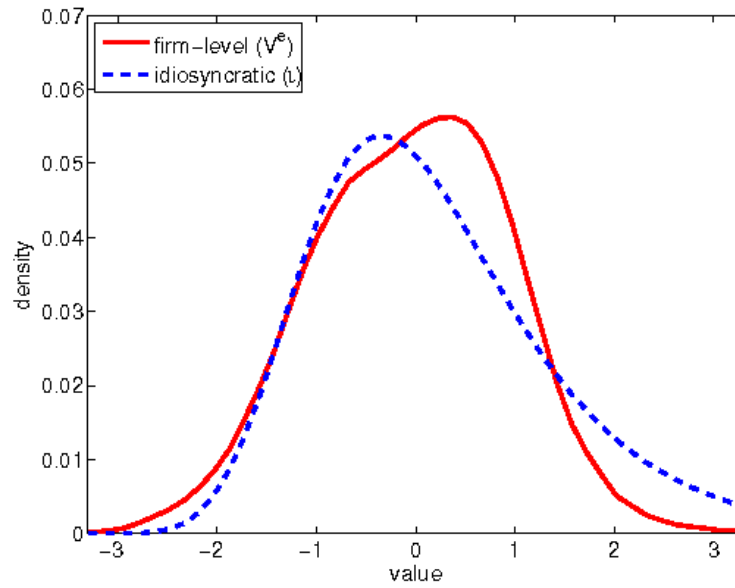
(b) Earnings
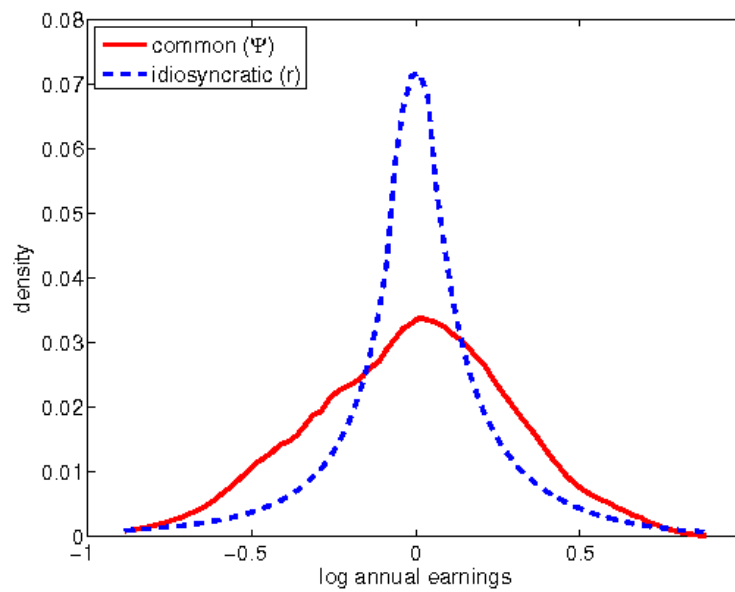
Both distributions are normalized so that the median value is zero.

Figure 1.11: Distributions

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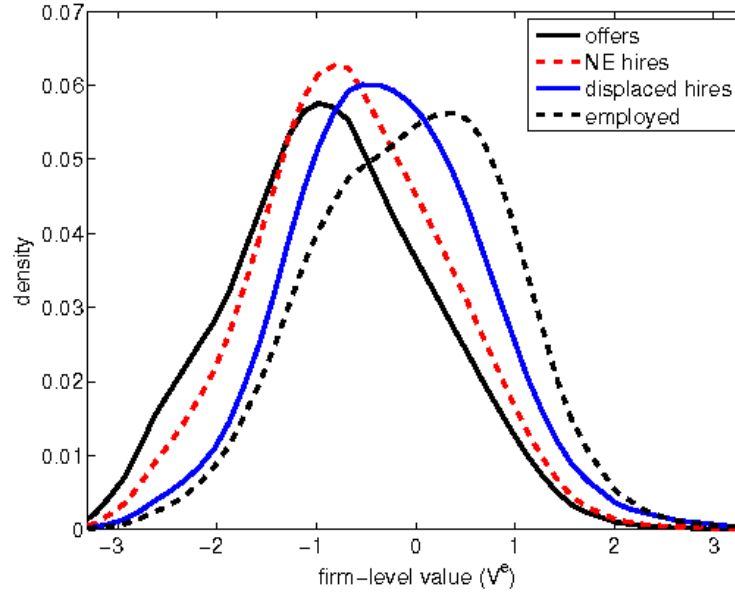
(b) Earnings
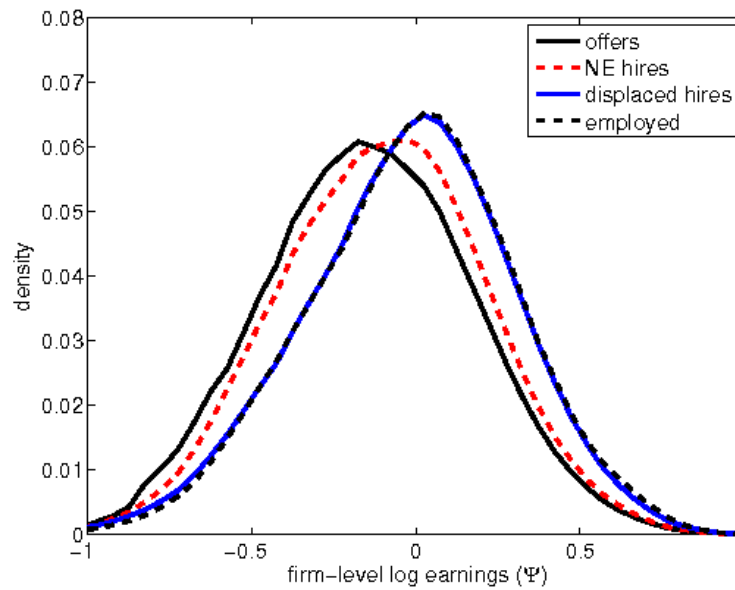
This figure plots the dispersion in the firm-level values (top panel) and earnings (bottom panel) in four distributions: the offer distribution, the distribution of where hires from nonemployment accept offers, where displaced workers making employer-to-employer transitions accept offers, and finally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

Figure 1.12: Relationship between values and earnings

(a) Sector level
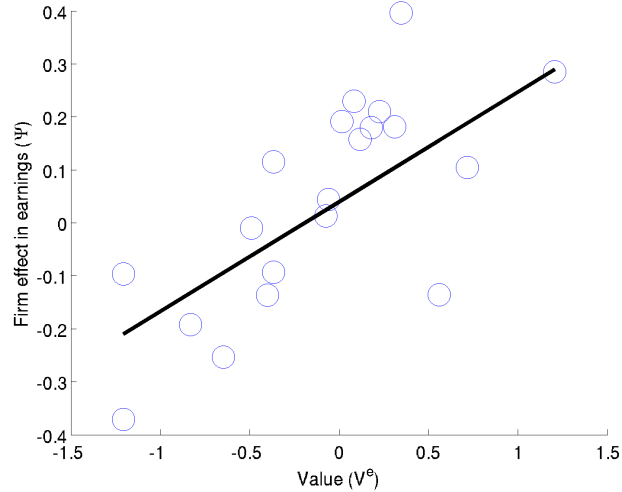
(b) Overall
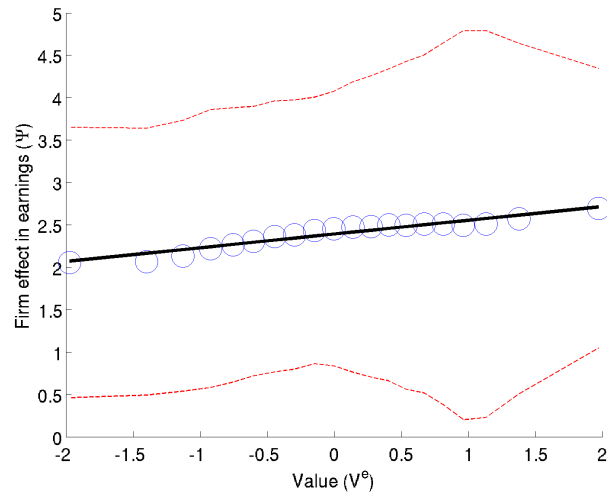
The top panel of this figure plots the sector-level means of the earnings and values. The solid black line plots the regression line run at the sector-level and weighting by the number of person-years represented by each sector. The R^2 is 0.45. The bottom panel sorts firms on the basis of firm-level values. The circles plot 20 equal-person-year-sized bins as well as the mean firm-level earnings, 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 1.13: Relationship between R^2 and firm size

(a) All
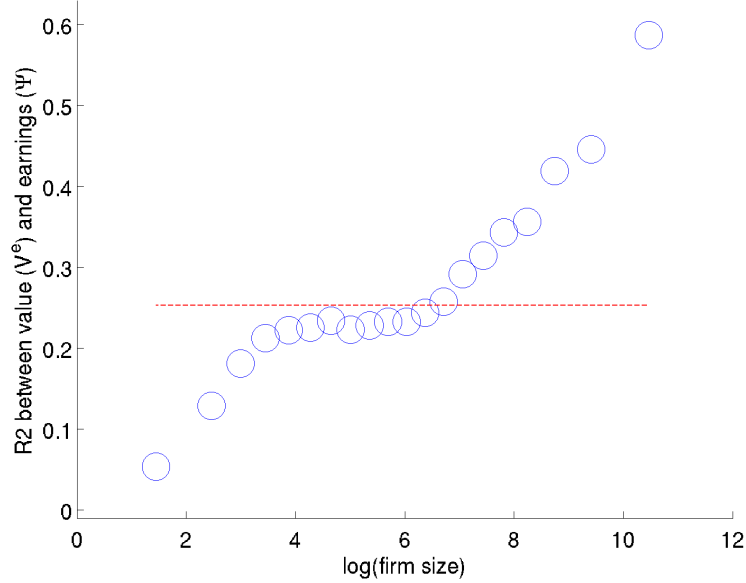
(b) Residual
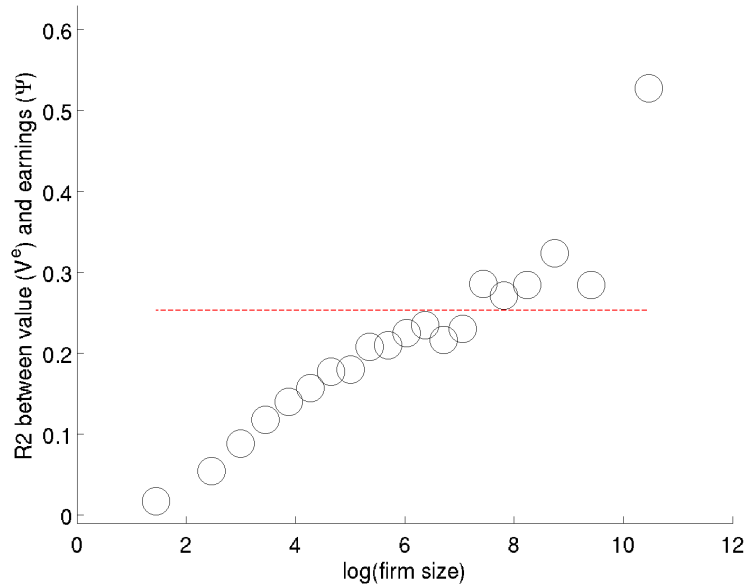
This figure sorts firms on the basis of firm size into 20 equal person-year-sized bins. The upper panel plots the R^2 between the "raw" firm-level values and earnings computed bin-by-bin. The bottom panel plots the R^2 on the residual firm-level values and earnings, where I have removed the county-, industry- and size- means from the earnings and the values.

Figure 1.14: Counterfactual inequality

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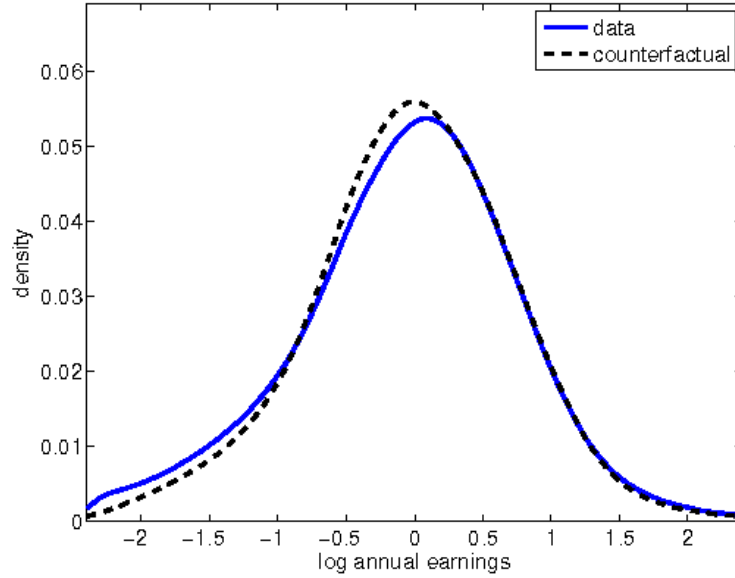
(b) "Naive"
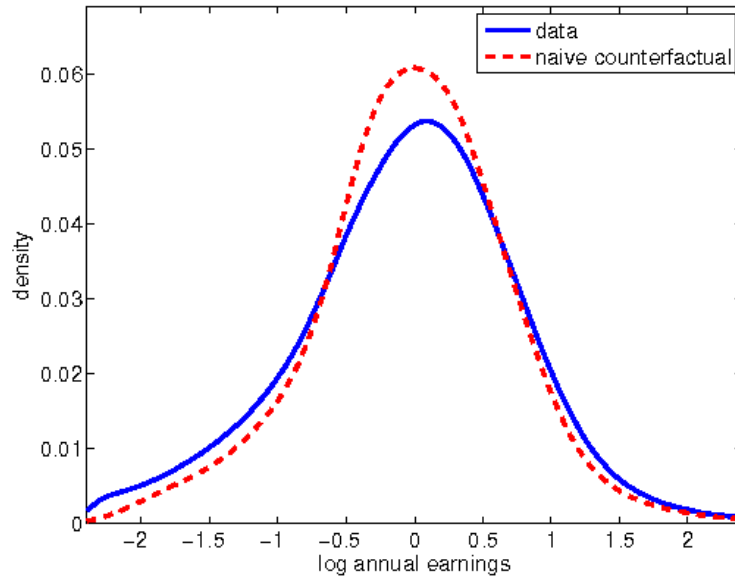
The top panel of this figure plots the distribution of income in the data, and in a counterfactual where I use my estimates of the firm-level values to price out compensating differentials. The bottom panel considers a "naive" counterfactual where I deflate all the firm components of earnings by a constant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

1.8 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 non-negative: 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 non-negative 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 non-zero 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 non-negative 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 jth 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 jth row.

$$[S^{-1}Mexp(\tilde{V})]_j = [\lambda exp(\tilde{V})]_j \tag{1.18}$$

$$\frac{e_j^T Mexp(\tilde{V})}{||Me_j||_1} = \lambda e_j^T exp(\tilde{V}), \tag{1.19}$$

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:

$$||Me_j||_1 = \mathbf{1}^T Me_j. \tag{1.20}$$

Rearrange:

$$\frac{e_j^T Mexp(\tilde{V})}{||Me_j||_1} = \lambda e_j^T exp(\tilde{V}) \tag{1.21}$$

$$\frac{e_j^T Mexp(\tilde{V})}{\mathbf{1}^T Me_j} = \lambda e_j^T exp(\tilde{V}) \tag{1.22}$$

$$e_j^T Mexp(\tilde{V}) = \lambda \mathbf{1}^T Me_j e_j^T exp(\tilde{V}). \tag{1.23}$$

Now sum over the rows:

$$\sum_j e_j^T Mexp(\tilde{V}) = \sum_j \lambda \mathbf{1}^T Me_j e_j^T exp(\tilde{V}) \tag{1.24}$$

$$\sum_j e_j^T Mexp(\tilde{V}) = \lambda \sum_j \mathbf{1}^T Me_j e_j^T exp(\tilde{V}) \tag{1.25}$$

$$\mathbf{1}^T Mexp(\tilde{V}) = \lambda \sum_j \mathbf{1}^T Me_j e_j^T exp(\tilde{V}) \tag{1.26}$$

$$\mathbf{1}^T Mexp(\tilde{V}) = \lambda \mathbf{1}^T M \sum_j e_j e_j^T exp(\tilde{V}) \tag{1.27}$$

$$\mathbf{1}^T Mexp(\tilde{V}) = \lambda \mathbf{1}^T Mexp(\tilde{V}). \tag{1.28}$$

Hence, $\lambda = 1$. $\qquad\square$

Proof of Result 2

Proof. The proof shows that the diagonal elements cancel out. First, use the identity from (1.11):

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

Expand to write the diagonal elements explicitly:

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

Then cancel the diagonal terms to show that (1.11) holds with arbitrary diagonal elements:

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

□

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:

$$R^2 = \frac{Cov(\Psi, V)^2}{Var(\Psi)Var(V)} \tag{1.29}$$

$$= \frac{Cov(\Psi, \alpha(\Psi + a))^2}{Var(\Psi)Var(\alpha(\Psi + a))} \tag{1.30}$$

$$= \frac{\alpha^2 Cov(\Psi, (\Psi + a))^2}{\alpha^2 Var(\Psi)Var((\Psi + a))} \tag{1.31}$$

$$= \frac{[Var(\Psi) + Cov(\Psi, a)]^2}{Var(\Psi)[Var(\Psi) + Var(a) + 2Cov(\Psi, a)]}. \tag{1.32}$$

It is also helpful to write $Var(a)$ in terms of one unknown by rearranging equation (1.32):

$$R^2[Var(\Psi)^2 + Var(\Psi)Var(a) + 2Var(\Psi)Cov(\Psi, a)] = Var(\Psi)^2 + 2Var(\Psi)Cov(\Psi, a) + Cov(\Psi, a)^2$$
(1.33)

$$R^2Var(\Psi)Var(a) = (1 - R^2)Var(\Psi)^2 + 2(1 - R^2)Var(\Psi)Cov(\Psi, a) + Cov(\Psi, a)^2$$
(1.34)

$$Var(a) = \frac{(1 - R^2)Var(\Psi)^2 + 2(1 - R^2)Var(\Psi)Cov(\Psi, a) + Cov(\Psi, a)^2}{R^2Var(\Psi)}.$$
(1.35)

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

$$Corr(\Psi, a) = \frac{Cov(\Psi, a)}{\sqrt{Var(a)Var(\Psi)}}$$
(1.36)

$$= \frac{Cov(\Psi, a)}{\sqrt{\frac{(1 - R^2)Var(\Psi)^2 + 2(1 - R^2)Var(\Psi)Cov(\Psi, a) + Cov(\Psi, a)^2}{R^2Var(\Psi)} Var(\Psi)}}$$
(1.37)

$$= \sqrt{R^2} \frac{Cov(\Psi, a)}{\sqrt{(1 - R^2)Var(\Psi)^2 + 2(1 - R^2)Var(\Psi)Cov(\Psi, a) + Cov(\Psi, a)^2}}.$$
(1.38)

Proof. **A lower bound on** $Var(a)$**:** To minimize $Var(a)$, start with the expression for $Var(a)$ (equation (1.35)) in terms of $Cov(\Psi, a)$ and take the first order condition with respect to $Cov(\Psi, a)$:

$$\frac{\partial Var(a)}{\partial Cov(\Psi, a)} = \frac{2(1 - R^2)Var(\Psi) + 2Cov(\Psi, a)}{R^2Var(\Psi)}$$
(1.39)

$$0 = \frac{2(1 - R^2)Var(\Psi) + 2Cov(\Psi, a)}{R^2Var(\Psi)}$$
(1.40)

$$Cov(\Psi, a) = -(1 - R^2)Var(\Psi).$$
(1.41)

The second order condition is, $\frac{2}{R^2Var(\Psi)}$ which is positive. Substitute this into the expression for $Var(a)$ (equation (1.35)) to get that the minimum value is given by:

$$Var(a) = \frac{(1 - R^2)Var(\Psi)^2 + 2(1 - R^2)Var(\Psi)(-(1 - R^2)Var(\Psi)) + (-(1 - R^2)Var(\Psi))^2}{R^2Var(\Psi)}$$
(1.42)

$$= Var(\Psi)(1 - R^2).$$
(1.43)

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

$$Corr(\Psi, a) = \frac{Cov(\Psi, a)}{\sqrt{Var(a)Var(\Psi)}} \tag{1.44}$$

$$= \frac{-(1 - R^2)Var(\Psi)}{\sqrt{Var(\Psi)(1 - R^2)Var(\Psi)}} \tag{1.45}$$

$$= -\sqrt{1 - R^2}. \tag{1.46}$$

And compute the variance of utility in log dollar units:

$$Var(\Psi + a) = Var(\Psi) + Var(a) + 2Cov(\Psi, a) \tag{1.47}$$

$$= Var(\Psi) + Var(\Psi)(1 - R^2) - 2(1 - R^2)Var(\Psi) \tag{1.48}$$

$$= R^2 Var(\Psi). \tag{1.49}$$

An upper bound on $Var(a)$**:** Take the limit of the expression for for $Var(a)$ (equation (1.35)):

$$\lim_{Cov(\Psi,a) \to \infty} \frac{(1 - R^2)Var(\Psi)^2 + 2(1 - R^2)Var(\Psi)Cov(\Psi, a) + Cov(\Psi, a)^2}{R^2 Var(\Psi)} = \infty. \tag{1.50}$$

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?

$$\lim_{Cov(\Psi,a) \to \infty} Corr(\Psi, a) = \lim_{Cov(\Psi,a) \to \infty} \sqrt{R^2} \frac{Cov(\Psi, a)}{\sqrt{(1 - R^2)Var(\Psi)^2 + 2(1 - R^2)Var(\Psi)Cov(\Psi, a) + Cov(\Psi}} \tag{1.51}$$

$$= \sqrt{R^2} \tag{1.52}$$

$$= Corr(\Psi, V). \tag{1.53}$$

And:

$$Var(\Psi + a) \to \infty. \tag{1.54}$$

$\square$

1.9 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 below). The three remaining moments are $\tilde{V}_n$, the probability of making an endogenous $employer-to-employer$ transition (equation (1.67)), and a moment that relates where workers go on nonemployment-to-employment transitions to the value of nonemployment (equation (1.60)). Intuitively, the value of nonemployment, V^n is increasing in the probability of making an 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 to 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.

1.9.1 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 1.1.2.1 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.

1.9.2 Moment matching steps

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

$$f_i^o \equiv \frac{f_i \frac{exp(V^e(v_i))}{exp(V^e(v_i))+exp(V^n)}}{\sum_{i' \in \mathcal{E}} f_{i'} \frac{exp(V^e(v_{i'}))}{exp(V^e(v_{i'}))+exp(V^n)}} \frac{\lambda_0 U}{\lambda_0 U} = \frac{M_{in}}{\sum_{j \in \mathcal{E}} M_{jn}}. \qquad (1.55)$$

The denominator is the share of offers accepted by unemployed workers. This equation only identifies f_i up to scale.[45] 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 unemployment, or:

$$C_1 \equiv \sum_{i' \in \mathcal{E}} f_{i'} \frac{exp(V^e(v_{i'}))}{exp(V^e(v_{i'})) + exp(V^n)} \qquad (1.56)$$

so that f_i^o can be written more compactly as:

$$f_i^o = \frac{f_i \frac{exp(V^e(v_i))}{exp(V^e(v_i))+exp(V^n)}}{C_1}. \qquad (1.57)$$

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 gridsearch 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, and unknowns 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,[46] and $\tilde{V}_i$ is estimated based on the matrix of moves across firms:

$$\frac{g_i exp(\tilde{V}_i)}{f_i^o}(1 - \delta_i)(1 - \rho_i) = g_i \frac{f_i exp(V^e(v_i))}{g_i(1 - \delta_i)(1 - \rho_i)} C_1 \frac{1}{f_i} \frac{exp(V^e(v_i)) + exp(V^n)}{exp(V^e(v_i))}(1 - \delta_i)(1 - \rho_i) \qquad (1.58)$$

$$= C_1[exp(V^e(v_i)) + exp(V^n)]. \qquad (1.59)$$

---

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

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

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:

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

$$(1.60)$$

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

$$(1.61)$$

$$= exp(V^n)C_1.$$

$$(1.62)$$

Step 2: Given a value of λ_1:

- Combining equations (1.59) and (1.62), give the following two terms: $C_1 exp(V^e(v_i))$ and $C_1 exp(V^n)$.

- Rewrite Equation (1.56) by multiplying by $\frac{C_1}{C_1}$ and rearranging:

$$f_i^o = f_i\frac{exp(V^e(v_i))}{exp(V^e(v_i))+exp(V^n)}\frac{1}{C_1} \qquad (1.63)$$

$$= f_i\frac{C_1exp(V^e(v_i))}{C_1exp(V^e(v_i))+C_1exp(V^n)}\frac{1}{C_1} \qquad (1.64)$$

$$f_i^o\frac{C_1exp(V^e(v_i))+C_1exp(V^n)}{C_1exp(V^e(v_i))} = \frac{f_i}{C_1}. \qquad (1.65)$$

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.[47])

$$\sum_{i\in\mathcal{E}}\frac{f_i}{C_1} = \frac{\sum_{i\in\mathcal{E}}f_i}{C_1} = \frac{1}{C_1}. \qquad (1.66)$$

Now that C_1 is known and from equation (1.65) $\frac{f_i}{C_1}$ is known, it is possible to solve for f_i since $\frac{f_i}{C_1}$ is known from equation (1.65).

---

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

- Knowledge of C_1 gives $exp(V^n)$ and $exp(V^e(v_i))$, via Equations (1.59) and (1.62).

This step produces the following $2N + 1$ parameters: $\{f_i, exp(V^e(v_i)), exp(V^n)\}$. In combination, this gives $5N + 3$ parameters.

Step 3: Given the parameters of the model, compute the probability of an "endogenous" employer-to-employer transition:

$$\lambda_1 \sum_i g_i(1 - \delta_i)(1 - \rho_i) \sum_j f_j \frac{exp(V_j^e)}{exp(V_j^e) + exp(V_i^e)}. \tag{1.67}$$

To make this computationally feasible, group firms into 1000 categories on the basis of the firm values (V^e).

1.10 Appendix: Constructing datasets

1.10.1 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 very low earnings jobs.

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 job is chosen from among the employers where the worker had two or more consecutive quarters of earnings within the calendar year; the reason to restrict to jobs with two or more consecutive quarters of earnings is to allow me to code transitions between employers as employer-to-employer or employer-to-nonemployment-to-employer in a way that I discuss below.[48] In this set of jobs, the

---

[48]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, 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, 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. In constructing the dataset, making these two restrictions slightly decreased the variance of earnings, and so slightly increased the fit of the AKM model.

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 first identify the nature of the workers 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 in $t-1$, t and $t+1$) or continuous (if earnings in $t-1$ and t, or 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 worth of work, while if a worker is present in both adjacent quarters then the earnings reflect a full quarters work.[49] Then take the log of these earnings.

I then make two additional sample restrictions. First, I keep workers between the ages of 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 \$3250 in 2011:IV dollars.[50]

I use this dataset to construct three employer-level characteristics. First, average annual employment (based on employment as defined above). Second, average age of these workers. Finally, I construct the average log annualized earnings of workers. I also merge on industry information (six digit NAICS codes), and information on whether the employer is a multi-unit.

I now summarize how the various sample restrictions affect the same size. Table 1.7 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.[51] Table

---

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

[50]They drop daily wages of less than 10 euros. 10 euros $\times \approx 1.3$ euros per dollar $\times 250$ days per year= 3250.

[51]For Germany, Card, Heining, and Kline (2013, Appendix Table 1a, row 5) find 1.10 employers

1.8 shows the distribution of the number of jobs-per-year in row 2 of Table 1.7.

Table 1.9 shows that on the full annual dataset, 91% share of person-year observations are full-quarter and 9% are continuous.[52] Table 1.10) 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 1.11 shows that there is a substantial amount of mobility in this sample: half of the workers have two or more employers. Table 1.12 shows that about 10% of person-employer matches (or 30% of person-years) last for the entire span of my data. On the other hand, almost half of matches (20% of person-years) only last for a single year.

1.10.2 Quarterly dataset

I build on ideas developed in Bjelland et al. (2011) and Hyatt et al. (2014). Specifically, restricting to jobs with two quarters of earnings and using overlapping quarters of earnings to label an employer-to-employer (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 $t-1$. 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.[53] 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 and so it is important that they appear in the quarterly dataset.

per person per year, and this number is stable from 1985 to 2009.

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

[53]Sum together the two quarters of earnings and multiply by 4. If the earnings are below $3250 then drop the person-employer match. Multiplying by 4 is assuming that each quarter is a continuous quarter of employment. By assuming that this is a continuous quarter of employer this includes more jobs 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 assuming it is continuous.

If a worker has different quarterly dominant employers in quarter t and $t+1$, 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 below, I consider that worker to have been nonemployed in quarter $t = 1$, and so I label the transition from the quarter t dominant employer as a transition into nonemployment.[54],[55]

I depart from prior work to address the possibility that workers move on the seam between 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 above 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 and 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. Concretely, if the earnings from firm A in quarters $t-2$ and $t-1$ are within 5% of eachother (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 1.13 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 says, if this is a new job, whether the worker came to this job directly from another job, or had an intervening spell of nonemployment.

1.10.3 Using the quarterly dataset to construct displacement weights

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

---

[54]Similarly, Burgess, Lane, and Stevens (2000), when they compute job flows drop matches that only last a single quarter.

[55]This definition will pick up very few recalls as employment-nonemployment-employment 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% ($\frac{1}{13}$).

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 displacement weight is one minus the probability of the separation divided by the probability of the separation occuring at an expanding employer.[56]

I use the following employer size bins: 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 equal-weighted by person-quarter bins among contracting employers. I create 40 distinct group workers by demographic characteristics denoted by d (20 equal sized age bins for men, and 20 for women).

1.10.4 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 (employer-to-employer) or an employer-to-nonemployment-to-employertransition. 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 1.1.2.1.

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

---

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

the worker's annual dominant employer is A in 2002 and B in 2003.[57] 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 in adjacent quarters, 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: $A \xrightarrow{employer-to-employer} C \xrightarrow{employer-to-employer} B$. 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 $A \xrightarrow{employer-to-employer} C \xrightarrow{employer-to-nonemployment-to-employer} B$. 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.[58]

---

[57]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.

[58]In the case of multiple transitions between annual dominant jobs, I proceed as follows. In a case like $A \xrightarrow{employer-to-employer} C \xrightarrow{employer-to-employer} B$ I compute the endogenous weight for the $A \to C$ and $C \to B$ transitions and take the geometric average. In a case like $A \xrightarrow{employer-to-employer} C \xrightarrow{employer-to-nonemployment-to-employer} B$ 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-employment transition.

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 3) a worker leaves the labor force. For the purposes of computing transition probabilities I remove these observations 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.

1.10.5   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 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 a worker 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 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 $(1 − \delta)\rho$;[59] the number of exogenous *employer − to − nonemployment* transitions divided by W is δ.

1.11 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, and the benefits of preconditioning, see Trefethen and Bau (1997) (especially Lectures 38 and 40). Despite the fact that my dataset is larger than that used by CHK, 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.)

1.12 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. I.e. 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 second and subsequent years,

---

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

this selection term for a worker at employer i is

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

(1.68)

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

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

(1.69)

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

(1.70)

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

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

(1.71)

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

(1.72)

In implementation there are a couple issues. First, 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 so use the mobility relevant value. Third, to speed computation time I discretize the firms into 1000 equal-sized (in terms of person-years) bins and use this to compute the selection correction.

1.13 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.

1.13.1 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:

$$\Psi_i = \Psi_g + \tilde{\Psi}_i + \epsilon_i^{\Psi} \tag{1.73}$$

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 belong. Similarly, the value at firm i is given by (where I suppress the e subscript from the text for simplicity):

$$V_i = V_g + \tilde{V}_i + \epsilon_i^{V}. \tag{1.74}$$

I now state assumptions on how the terms relate.

Assumption 1. *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 this assumption is that grouping characteristics are exogenous. This assumption 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 this assumption is relatively mild. For example, it allows 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.

1.13.2 Estimation with one characteristic

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

For earnings:

$$\lim_{N_g \to \infty} \hat{\mathbb{E}}[\Psi_g] = \lim_{N_g \to \infty} \frac{1}{N_g} \sum_{i \in \Omega_g} \Psi_i \tag{1.75}$$

$$= \lim_{N_g \to \infty} \frac{1}{N_g} \sum_{i \in \Omega_g} \left[\Psi_g + \tilde{\Psi}_i + \epsilon_i^\Psi \right] \tag{1.76}$$

$$= \mathbb{E}[\Psi_g + \tilde{\Psi}_i + \epsilon_i^\Psi | i \in \Omega_g] \tag{1.77}$$

$$= \Psi_g + \mathbb{E}[\tilde{\Psi}_i + \epsilon_i^\Psi | i \in \Omega_g] \tag{1.78}$$

$$= \Psi_g \tag{1.79}$$

where the second line is a definition, the third line is a law of large numbers, the fourth is because Ψ_g is non-stochastic once g is fixed, and the last line is by assumption.

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

Then as the number of firms within each group grows large we 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.

1.14 Appendix: Subgroup results

This appendix re-estimates 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, and 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 re-estimate the whole model, the earnings decomposition and the comparison between them.

Table 1.17 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, the first column reproduces column (3) of table 1.1, which reports the decomposition for the whole sample. The table 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. On the hand, there are more

interesting patterns by age: firms explain more of the variance among young workers than among older workers.

Panel A of table 1.18 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 1.5. For ease of comparison, the first row reproduces 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 older workers than for younger workers.[60]

Panel B performs 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 re-estimate 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 A of table 1.18 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 on the theory that perhaps across sector moves are not well-described by a search model. The second column of Panel F in table 1.18 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, table 1.4 relies on precisely these across sector moves to value sectors.

1.15   Appendix: Additional tables and figures

---

[60]Understanding why patterns look different for younger and older workers is an interesting topic for future research.

Table 1.7: 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	$504,945,000$	$105,921,000$	$6,155,000$

Note: 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 persons dominant job in the particular year passes an earnings test. The third row goes down to the unique employer that is the workers "dominant" job, or the employer from which the worker made the most in the calendar year.

Table 1.8: Distribution of jobs per person per year

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

Note: All counts are rounded to the nearest thousand. This table deconstructs the gap between row 2 and row 3 in Table 1.7. The column sum is row 3 in Table 1.7. 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 1.9: 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

Note: All counts are rounded to the nearest thousand. The column sum is the number of person-years in row 3 in Table 1.7. 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 1.10: 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

Note: All counts are rounded to the nearest thousand. The column sum is the number of unique people in row 3 in Table 1.7.

Table 1.11: 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

Note: All counts are rounded to the nearest thousand. The column sum is the number of unique people in row 3 in Table 1.7.

Table 1.12: 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

Note: 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, and so 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 1.13: 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

Note: All counts are rounded to the nearest thousand. The dataset is the quarterly dataset, and so 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 1.14: Workers frequently move across industries and locations

Moves with different...	employer-to-employer and employer-to-nonemployment	employer-to-employer
State	0.12	0.10
County	0.65	0.64
Sector	0.65	0.61
4 digit industry	0.81	0.79
State	0.12	0.10
County	0.56	0.53
Sector	0.60	0.56
4 digit industry	0.77	0.74

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.

Table 1.15: 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

The sample is column (4) in Table 1.1.

Table 1.16: Relationship between alternative firm effects

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

This table shows correlations between firm effects estimated using different samples and different identifying variation. 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. I can estimate separate sending and receiving effects for 99.4% of firms in my main sample which contain 99.9% of the person-years in my main sample. 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 constitutes 98.4% of firms representing 99.9% of person-years.

Table 1.17: 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

Sample counts are rounded to the nearest thousand. Column (1) reproduces column (3) from table 1.1. 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 $3250 ($2011). The table includes person-years in which on December 31 of the year the person was 18-61 (inclusive).

Figure 1.15: States used in analysis
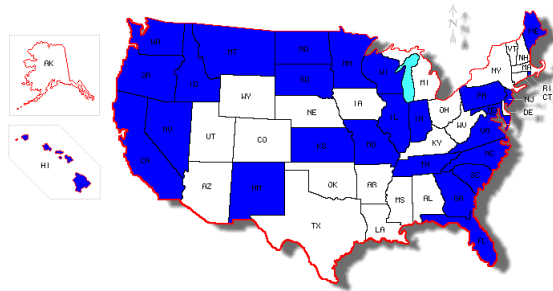
Note: Figure shows the states used in analysis.

Table 1.18: 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
$\text{Corr}(\tilde{V},\tilde{V}$ x by x)	0.97	0.76

Panel A reports results analogous to table 1.5. Row (1) reproduces results from the first row in column (4) in table 1.5. The state-by-state exercise in Panel F contains 94.1% of the SEINs in column (3) of table 1.1, which represent 99.3% of the person-years. The sector-by-sector exercise in Panel F contains 45.4% of the SEINs and 89.2% of the person-years.

Figure 1.16: Change in firm effect does not predict magnitude of earnings change in a matching model
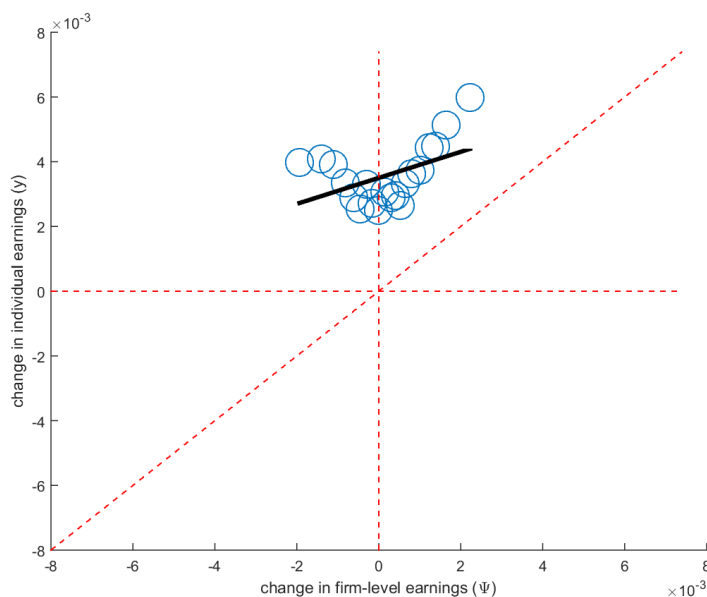
1.16 Appendix: Simulating models using individual data

1.16.1 Hall and Mueller (2013)

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

$$G(v) = \frac{\lambda F(v)u}{(1-u)[\lambda(1-F(v))+s]}. \tag{1.80}$$

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:

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 (1.80) to compute the steady state distribution. I then compute the R^2 between y and v in G.

1.16.2 Sullivan and To (2014)

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

$$G(v) = \frac{\lambda_u F(v)Pr(v > U^*)u + \lambda_{le}F(v)Pr(v > U^*)(1-u)}{[\lambda_e(1-F(v))Pr(v > U^*)(1-u) + (1-u)s + \lambda_{le}F(v)Pr(v > U^*)(1-u)]}, \tag{1.81}$$

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, Table 2, pg. 489, 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

CHAPTER II

Why do men and women work in different firms?

Men are more likely than women to work in both high-wage firms and high-wage industries. This sorting component accounts for 9 log points of the 36 log point (about 25%) gender earnings gap in the United States. A core interpretive issue is whether this sorting component reflects discrimination or differences in preferences.[1] The discrimination explanation is that women would like to work at the same firms or in the same industries as men, but are prevented from doing so. The preference explanation is that there are nonpecuniary characteristics that differ across the low- and high-paying firms and women value these nonpecuniary characteristics more than men. For example, women might highly value flexibility while men do not.

In this chapter, I build on chapter 1 to shed light on these two explanations. I begin by establishing the key fact that this paper is interested in understanding: men are at higher-paying firms than women in the United States, and more generally, that men and women are sorted in the labor market. Building on Card, Cardoso, and Kline (2015), I estimate the Abowd, Kramarz, and Margolis (1999) decomposition separately for men and women. This decomposition amounts to using the complete set of wage changes of workers who switch firms to estimate a firm effect in earnings. Estimating the decomposition separately for men and women allows me to construct gender-specific earnings premia at each firm. I reproduce the finding of Card, Cardoso, and Kline (2015) that the sorting of men and women across firms is quantitatively important; indeed, I find that the sorting component is more important in the U.S. than in their Portugese context.

I then turn to understanding why men and women are sorted. Following Sorkin (2015a), I write down a partial equilibrium utility-posting search model in the spirit

---

[1]See, for example, the Council of Economic Advisers (pg. 3, 2015) `https://www.whitehouse.gov/sites/default/files/docs/equal_pay_issue_brief_final.pdf` and Card, Cardoso, and Kline (2014, pg. 30-31).

of Burdett and Mortensen (1998) where the only nonstandard ingredient is that firms post utility offers which combine a wage and a nonpecuniary bundle, and there is a (transitory) idiosyncratic utility draw in each match, which explains why people might make different choices. The model estimates revealed values of employers by taking into account the network structure of accepted offers and rejected offers.

The model embeds versions of both the preference and discrimination explanations for why men and women are sorted. The preference explanation in the model is that men and women rank firms differently and so given the same set of opportunities would end up in different firms. The discrimination explanation in the model is that men and women receive a different set of offers, and so given the same preferences end up in different firms.

To separate the preference and discrimination explanations, I estimate the model separately by men and women. The implicit assumption in this exercise—as with standard earnings decomposition exercises—is that men and women operate in separate labor markets. This allows me to estimate separate offer distributions (opportunities), values (preferences), and earnings for men and women. In particular, I am able to estimate separate offer probabilities, values and earnings for men and women firm-by-firm. While the model relies on numerous strong assumptions, estimation is completely nonparametric along the dimensions that this paper is interested in. Namely, I impose no assumptions on the shape of the offer distribution, no assumptions on the distribution of employer values, and no assumptions on the relationship between men's and women's preferences or offer distributions.

My principal results are as follows. First, I find that men and women are systematically sorted in the labor market. About 60% of men's co-workers are men, while only about 40% of women's co-workers are men. This finding is quantitatively consistent with the results for the U.S. in Hellerstein, Neumark, and Mcinerney (2008, pg. 183). Second, I find that this sorting explains over 25% of the earnings gap (given the nature of the data, some of this might include differences in hours). This finding is robust to using men's earnings or women's earnings to compute how well-paying firms are. This extends the finding of Card, Cardoso, and Kline (2015) (for Portugal) that men are systematically employed in higher-paying firms than women.

I then use the model to interpret the reasons and consequences of sorting. The model points to differences in the offer distribution rather than differences in preferences to explain why men and women are sorted. Men's and women's preferences over firms are estimated to be highly correlated. For example, the overall estimated values have a correlation of 0.89. When I aggregate to the 4-digit industry level (which

explains 2/3s of the sorting component of the gender wage gap), the correlation is 0.98. But I estimate that women search from a lower-paying distribution. This result should be interpreted cautiously since—as I discuss further in the conclusion—the offer distribution is a reduced-form object that may itself contain revealed preference information (put differently, the fact that women are less likely to receive offers from high-paying firms may reflect the fact that they do not want to work at those firms, rather than the fact that they do not receive offers from these firms). The model also allows me to compare the distribution of employer values at the firms that employ men relative to the firms that employ women. Taking the model at face value, the estimates indicate that men and women are at approximately equal-valued firms regardless of whether it is men's or women's values that are used to value the firms.

2.1 Matched employer-employee data

To be able to study the extent to which men and women are sorted into firms requires matched employer-employee data. I use the U.S. Census Bureau's Longitudinal Employer Household Dynamics (LEHD) dataset. This quarterly dataset is constructed from employer Unemployment Insurance (UI) filings.[2]

This section draws on the discussion in Sorkin (2015a).

2.1.1 Data description

Being constructed from unemployment insurance records implies four features that 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.[3] This is desirable in the sense that a "smaller" notion of an employer means that working conditions are probably more similar. On the other hand, it means that I will not capture segregation within firms (i.e. men work in the executive office, and women work in the field). Second, only employers that are covered by the UI system appear in the dataset.[4] Overall, in 1994 the UI system covered about 96% of employment

---

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

[3] For employers that operate in multiple states, this understates true employer size. Similarly, it is possible for a given employer to have multiple UI 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."

[4] This restriction results in the exclusion of certain sectors of the economy. In particular, small nonprofits (those employing fewer than four workers), domestic, self-employed, some agricultural workers and federal government (but not state and local government) are excluded. For more

and 92.5% of wages and salaries (BLS (1997, pg. 42)). Third, the LEHD allows me to measure earnings, but not hours.[5] Thus, variation in these benefits as well as hours will be included in my measure of compensating differentials. This means that to the extent that women work in firms that tend to offer more part-time jobs, then this will be captured in the sorting component of the gender earnings gap (in contrast, Card, Cardoso, and Kline (2015) have hours in their Portuguese data). Finally, unlike the Portuguese data used by Card, Cardoso, and Kline (2015), the LEHD does not contain occupation data.

Being able to track employers over time is central to measuring employer-to-employer flows. The LEHD contains unique longitudinally-consistent employer identifiers. Administrative errors in the longitudinal linkages would lead to an overstatement of flows. In addition, business reorganizations, i.e. mergers and acquisitions or spinoffs, might lead to measured flows that are not economically relevant. Following Benedetto et al. (2007) I correct the longitudinal linkages 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 employer B is a relabelling of employer A, or else employer B acquired employer A, and so I do not count this as an employer-to-employer transition.

Like the dataset used by Topel and Ward (1992), the LEHD contains age, race and sex. I pool data from 27 states from the fourth quarter of 2000 to the first quarter of 2008.[6] Pooling data means that I keep track of flows between as well as within these states.

2.1.2 Dataset construction

To define a worker's employer, I reduce my dataset to one observation per person per year. The observation is the worker's *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 the worker had two quarters of employment at the

---

complete discussions see Kornfeld and Bloom (1999, pg. 173), BLS (1997, pg. 43) and `http://workforcesecurity.doleta.gov/unemploy/pdf/uilawcompar/2012/coverage.pdf`.

[5]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)).

[6]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 Appendix Figure 1.15 for a map.

employer and that the second quarter occurred in the calender year.[7] I also restrict attention to workers aged 18-61 (inclusive) and, following Card, Heining, and Kline (2013), require that the annualized earnings exceed \$3250 (in 2011 dollars). With an annualized 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.

See the appendices in Sorkin (2015a) for more details.

2.2 The sorting component of the gender earnings gap

This section shows how to establish the facts that this paper seeks to understand: women and men work at different firms and women work at lower-paying firms than men.

2.2.1 Sample and sorting

Table 2.1 describes the sample as well as the effects of various sample restrictions I make to be able to estimate the search model. The table shows that women make less than men, and women and men are sorted in the labor market.

Columns (1) - (3) contain information about men, while columns (4) - (6) contain information about women. Column (1) shows that I start with about 260 million men-years, 55 million men, and 4.7 million employers. In the set of firms strongly connected by EE moves of men, there are 50 million men and 1.3 million firms. Column (2) shows that firms account for about 22% of the variance of earnings between men. Column (3) shows that when I restrict attention to the set of firms where I can estimate firm effects and the model for both men and women, I have around 200 million men-years and 46 million men. Looking at the summary statistics, the means and variance of earnings do not change by very much when I restrict the sample.

For women, column (4) shows that I start with about 245 million women-years, 51 million men, and 4.6 million employers. In the set of firms strongly connected by EE moves of men, there are 48 million women and 1.2 million firms. Column (5) shows that firms account for about 23% of the variance of earnings between men. Column

---

[7]Reduction to one observation per person per year is common. See e.g. 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 AKM Bagger et al. (Forthcoming) also reduce to one such observation since in Danish administrative data the wage is only observed for the worker's employer in November. See Taber and Vejlin (2013) for a discussion of this point (they use Danish data but preserve more than one job per year, but only have one earnings observation per year).

(6) shows that when I restrict attention to the set of firms where I can estimate firm effects and the model for both men and women, I have around 200 million women-years and 46 million men. Looking at the summary statistics, the means and variance of earnings do not change by very much when I restrict the sample. It is interesting to note that the gender wage gap slightly decreases in the overlapping set of employers: from 0.35 log points to 0.33 log points.

The bottom panel of the table shows that men and women are sorted in the labor market. In the sample that I can estimate the model in (column (3)), 60% of men's co-workers are men, while 40% of women's co-workers are women. This degree of segregation is slightly less than what I find in the overall sample.

2.2.2 Measuring high-paying firms

One candidate explanation for why men and women work in different firms is that the firms that are high-paying for men are not high-paying for women. If high-paying firms for women were different than for men, then based solely on comparative advantage considerations it would not be surprising to see that men and women work in different firms. This section describes how to measure high-paying firms separately for men and women. The notion of a high-paying firm conditions on person fixed effects, and so addresses the possibility that high-paying reflects differences in human capital or other fixed characteristics.

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

$$\underbrace{y_{wt}}_{\text{log earnings}} = \underbrace{\alpha_w}_{\text{person effect}} + \underbrace{\Psi_{\boldsymbol{J}(w,t)}}_{\text{firm effect}} + \underbrace{x'_{wt}\beta}_{\text{covariates}} + \underbrace{r_{wt}}_{\text{residual/error term}}, \qquad (2.1)$$

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.[8]

By estimating equation (2.1) for men and women separately, I allow firms that are high-paying for men to be low-paying for women. Equation (2.1) is identified by workers who switch between firms.[9] It might be that men have systematically different

---

[8]Because I only use 7 years of data, the linear terms in the age-wage profile are highly correlated with the person fixed effects and so, following Card, Heining, and Kline (2013) are omitted.

[9]As in Sorkin (2015a), I can estimate equation (2.1) using EE and ENE transitions, or just EE transitions and it turns out that I get quantitatively very similar results.

patterns of earnings changes than women. In addition, because this is identified by switchers, it removes the time-invariant person effects captured in the α_w, which might includes differences in human capital.

To quantify the role of sorting in the earnings gap between men and women, I can use two sets of "prices." Considering equation (2.1), the earnings gap is

$$\frac{\sum_{w \in \mathbb{M}} \sum_t y_{wt}}{N_m} - \frac{\sum_{w \in \mathbb{F}} \sum_t y_{wt}}{N_f}, \tag{2.2}$$

where $\mathbb{M}$ is the set of male workers, $\mathbb{F}$ is the set of female workers, N_m is the total number of male person-years and and similarly for N_f. I quantify the role of sorting across firms using either the "male" or "female" prices, denoted $\Psi^m_{\boldsymbol{J}(w,t)}$ and $\Psi^f_{\boldsymbol{J}(w,t)}$ respectively.

2.2.3 Role of sorting in gender earnings gap

Sorting explains a large share of the gender earnings gap, regardless of whether men's and women's prices are used. Panel A of table 2.2 shows that regardless of whether men's or women's firm-specific-pay is used, the sorting of men and women across firms explains about 25% of the gender earnings gap documented in table 2.1. This is quantitatively larger than what Card, Cardoso, and Kline (2015) find for Portugal (they find that the sorting component is about 15%). One plausible explanation is that in the Portuguese data they observe hours and so part of what the sorting component is picking up is differences in hours; that is, there are high- and low-hours firms, and men are women are sorted on this basis. Figure 2.1 depicts the sorting component graphically. It shows the distribution of firm-level earnings at the firms where men and women work for two different sets of "prices:" firm effects estimated using the earnings changes only of men, and only of women. In both panels it is clear that the women's distribution is left-shifted. That is, women work at lower-paying firms.

Sorting across industry explains a large share of the sorting component. In panel B of table 2.2, I aggregate the firm-specific earnings premia to the 4-digit industry level (there are 312 of them). The sorting across industry explains about 75% of the sorting component.

Sorting across locations—i.e. the relative labor supply of men and women differs across high- and low-paying locations—is quantitatively small. Panel C shows that it explains about 3% of the gender earnings gap, or about 10% of the sorting component.

2.3 A model to understand sorting

The previous section established that men and women are sorted in the labor market. A key issue is why. This section writes down a model which contains three explanations for why men and women are at different firms. First, it might be harder for women to climb the job ladder. Formally, the "search" parameters, $\{\{\delta\}, \{\rho\}, \lambda_1\}$, might differ by gender so that women are more likely to lose their jobs involuntarily, or less likely to receive outside offers. Second, preferences might differ. In the model, this is reflected in differences in the gender-specific firm-level values V^e. Third, opportunities might differ. This is reflected in differences in the gender-specific offer distributions, or f^e.

The model is a partial equilibrium search model in the spirit of Burdett and Mortensen (1998) where firms post utility offers. These utility offers might differ between men and women. The utility offers consist of a combination of earnings and amenities. Formally,

$$V^e = \omega(\Psi + a)$$

where V^e is the firm-specific values, the Ψ is the firm-specific earnings from the previous section and a is a bundle of amenities.

Preference differences arise in the model because men and women might have different V^e at a specific firm, which might be due gender specific differences in either earnings or amenities. For example, if there are flexible and inflexible firms and women value flexibility more than men, then at a firm i $a_i, women > a_{i,men}$ and so all else equal women would value firm i more highly. On the other hand, differences in pay would also be included as differences in preferences. If men and women have the same values a at firms, but different levels of earnings at firms, Ψ, then this would also generate differences in valuations and would generate sorting in the labor market. Because preferences are only identified up to a constant factor, it might be women extract less "utility" from the labor market than men, but I will not be able to identify this; for example, an important finding of Card, Cardoso, and Kline (2015) is that in Portugal on average $\Psi_{women} = 0.9\Psi_{men}$, but I cannot make similar statements about the V_e.

Discrimination (or differences in opportunities) arises in the model through the offer distribution, f_i^e. When employed and unemployed, men and women receive offers drawn at random from the offer distribution. Differences in the offer distribution naturally give rise to differences in where men and women work—even if men and women have identical preferences.

A final reason why men and women might be employed in different firms is that the search parameters differ. In the model, an employed worker receives a job offer at rate λ. If this parameter differs between men and women, then they would end up in different firms. Intuitively, a higher draw of λ allows workers to climb the ladder faster. The model also has various exogenous shocks—a shock that forces a worker to make an EE transition (ρ) and one that forces a worker to go to nonemployment (δ)—that might differ by gender. Differences in these parameters can also generate sorting in the labor market for the same reason as differences in λ: it might be harder for women to climb the ladder.

The remainder of this section closely follows Sorkin (2015a). While I allow for differences in preferences between men and women, I allow for a very limited form of preference heterogeneity among men and women. 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 A to B and B to A. This is probably the most controversial assumption in the model. According to this assumption, all worker choices tell us the same thing about firms, whereas in richer models decisions of workers at different points in their careers tell us different things.

A key additional assumption in the model is what Hall and Mueller (2013) term the proportionality-to-productivity hypothesis. The only form of persistent heterogeneity the model accommodates among men and women is a worker-w-specific constant which enters the flow payoff to all employers as well as nonemployment. The search parameters are the same for all workers and so I can use the structure of the search model to infer rejected offers.

The following Bellman equation summarizes this verbal discussion of the model (I omit the gender specific subscripts for notational compactness). A worker at employer

i has the following value function:[10]

$$\underbrace{V^e(v_i)}_{\text{value of being at i}} = \underbrace{v_i}_{\text{flow payoff}} + \underbrace{\beta}_{\text{discounter}} \mathbb{E}\Bigg\{ \delta_i \underbrace{\int_{\iota_1} \{V^n + \iota_1\}dI}_{\text{exogenous job destruction}} + \tag{2.3}$$

$$\rho_i(1-\delta_i) \underbrace{\int_{V^e(v')} \int_{\iota_2} \int_{\iota_3} \underbrace{V^n + \iota_3}_{EN} dIdIdF}_{\text{exogenous employer-to-employer (reallocation)}}$$

$$+ \underbrace{(1-\rho_i)(1-\delta_i)}_{\text{no exogenous shocks}} \times \Bigg[\underbrace{\lambda_1}_{\text{offer}} \underbrace{\int_{V^e(v')} \int_{\iota_4} \int_{\iota_5} \max\{\underbrace{V^e(v'')+\iota_4}_{\text{accept}}, \underbrace{V^e(v_i)+\iota_5}_{\text{reject}}\}dIdIdF}_{\text{endogenous employer-to-employer}}$$

$$+ \underbrace{(1-\lambda_1)}_{\text{no offer}} \underbrace{\int_{\iota_6} \int_{\iota_7} \max\{\underbrace{V^n+\iota_6}_{\text{accept}}, \underbrace{V^e(v_i)+\iota_7}_{\text{reject}}\}dIdI}_{\text{endogenous job destruction}} \Bigg] \Bigg\}. \tag{2.4}$$

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 EE 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 EN transitions. A worker can be hit be 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 in where workers who are hired from nonemployment end up. Formally, a worker who is unemployed has the

_____

[10]The value does not include the current period's idiosyncratic draw. See Arcidiacono and Ellickson (2011, pg. 368) for further discussion of this point.

Bellman equation:

$$\underbrace{V^n}_{\text{value of being nonemployed}} = \underbrace{b}_{\text{flow payoff}} \tag{2.5}$$

$$+ \underbrace{\beta}_{\text{discounter}} \mathbb{E}\{\underbrace{\underbrace{\lambda_0}_{\text{offer } V^e(v')} \int_{\iota_8} \int_{\iota_9} \int \max\{\underbrace{V^e(v') + \iota_8}_{\text{accept}}, \underbrace{V^n + \iota_9}_{\text{reject}}\} dI dI dF}_{\text{endogenous nonemployment-to-employment}}$$

$$+ (1 - \lambda_0) \underbrace{\int_{\iota_{10}} (V^n + \iota_{10}) dI}_{\text{no offer}} \}. \tag{2.6}$$

Reading from left to right, an unemployed worker receives a total value of nonemployment of b, which includes both unemployment benefits as well as the value of non market 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 draw of nonemployment.

2.3.1 Estimation/identification

While Sorkin (2015a) provides complete details on how I estimate the model, this section provides an informal discussion of the features of the data that are used to identify each of the three classes of explanations.

The preference and opportunity explanation are identified by different features of how workers do and do not move across firms. The firm-level values take into account all of the information in how workers move across firms (i.e. the number of workers moving from firm A to B and from B to A), as well as how workers move between firms and nonemployment. In addition, the firm-level values account for the possibility of rejected offers by comparing the size of firms to their prominence in the offer distribution.

The offer distribution is separately identified from the firm-level values because the offer distribution is estimated mainly using the information in flows from nonemployment to each firm, whereas the firm-level values also use the flows between employers as well as the flows from each firm to nonemployment. This does mean that there is a mechanical tendency for the model to find that firms that are more prominent in the estimated offer distribution have lower values.

The search parameters are identified by a variety of features of the data. To

identify the shocks, I rely on the growth rates of firms at the time the worker was separating. Excess probabilities of separating at contracting firms are treated as shocks. To identify the arrival rate of offers, the model matches the level of employer-to-employer transitions.

2.4 Explaining sorting

This section uses the model estimates to show that it is differences in the offer distribution that explain why women work at lower-paying firms than men.

Differences in the search parameters are unlikely to explain sorting because they are very similar by men and women. Table 2.3 shows the parameters in the search model by men and women separately. The parameters are very similar. One notable commonality is that women in general have "better" transitions: a slightly higher share of their transitions are EE (recall that if a person leaves the sample, then this is not counted as a transition—so a woman leaving the labor force would not count as an EN), and women are also slightly less likely to be displaced.

Differences in preferences are unlikely to explain sorting. Table 2.4 displays the correlations between men's and women's values. The overall correlation is 0.89, so this does not leave a lot of room for preferences to explain sorting across firms. Recall from table 2.2, however, that industry explains about two-thirds of the sorting component of earnings gap between men and women. Panel C shows that at the industry level the correlation between men's and women's values is 0.98. So there is not much room for preferences to explain sorting at this level (though there might be within industry).

These similarities in preferences imply that the men and women have similar rankings of amenities at firms. Recall that the values at firms are the sum of amenities and earnings. The table 2.4 shows that what are high-paying firms for men, are also high-paying firms for women with a firm-level correlation of 0.92 and an industry-level correlation of 0.97. Since the values and earnings are similar, it must be that the valuations attached to the amenities are also similar. Note that this result is consistent with the findings of Card, Cardoso, and Kline (2015) in that their results are not about men and women ranking firms differently.

Differences in the offer distribution are likely to explain sorting. The first piece of evidence is by process of elimination: the model contains three explanations and I have just discussed why the non-offer distribution reasons are likely to be quantitatively unimportant in explaining sorting.

More direct evidence comes from directly considering the offer distributions. Table

2.5 shows that the offer distribution from which women search contains lower-paying firms. The difference is about 4 or 5 log points, which is about half of the gap overall. Figure 2.2 shows the same result graphically that women search from a lower paying offer distribution than men.[11] An interesting additional observation is that both offer distributions are left-shifted relative to the employed distributions depicted in 2.1, which is a prediction of search theory and this finding is not mechanical because I do not use the earnings data in estimation of the offer distribution.

2.5   Consequences of sorting

Unlike in earnings where it is unambiguous that women are at worse firms than men, when measured by utility the answer is less clear. Panel A of table 2.6 shows that when using the utilities estimated using choices of both men and women that men are at slightly better firms (the gap is 0.01, where the units are normalized by a type I extreme value distribution with variance 1). Using male valuations, men are at slightly better firms than women, while using female valuations, women are at slightly better firms than men.

The differences in the values of men and women are quantitatively very small relative to the dispersion in the values. Figure 2.3 provides a graphical depiction of the sorting component of the gender earnings gap, but with utilities. It shows the distribution of firm-level utilities at the firms where men and women for two different sets of "prices:" firm-level utilities estimated using the choices of men only, and of women only. In both panels, the left shift of the women's distribution that is apparent in the earnings is not apparent in the utilities.

Panel B of table 2.6 shows that women would be made worse off by switching industry distributions with men. According to the estimates, women would see a slight decline in utility by switching industry locations with men, and men would similarly be made slightly worse off by switching places with women. (Recall from table 2.2 that sorting across industry accounted for about two-thirds of the sorting component of the earnings gap.)

Finally panel C of 2.6 shows that women would prefer the location distribution of where men are employed, but this effect is quantitatively very small.

A natural concern is that the result that men and women are at equal-valued firms would be baked-in to the estimation procedure. The simplest way to demonstrate that

---

[11]Quantifying the contribution of each of these factors would require simulating the model under various counterfactuals. Figuring out how to do so is left for future work.

this is not the case is to observe that I can do the same subgroup analysis by young and old workers (where young means 18-34 and old means 35-61, which approximately divides the set of EE transitions in half). I find a result similar to men and women in the sense that: 1) young and old workers are different firms; and 2) this explains a large share of the earnings gap between young and old workers. But I also find that young workers would be better off if they were employed at the firms that employ older workers. Table 2.7 shows these results. Panel A shows that older workers are at higher paying firms and the magnitude of this gap is similar to that for between men and women.[12] In contrast to the gender earnings gap, panel B shows that young workers would prefer to switch places with older workers.

2.6   Discussion

This chapter uses a random search model to argue that men and women work at different firms because of differences in opportunities, rather than differences in preferences. The chapter interprets differences in opportunities as "search from a different offer distribution."

The key interpretive issue is where the different offer distribution comes from. In a random search model, the exogenous offer distribution is a reduced-form representation of the complicated process by which workers direct their search towards particular firms, and firms direct their search towards particular workers.

Because the offer distribution is a reduced-form representation of worker and firm behavior it might be possible to interpret the differences in the offer distribution as containing revealed preference information, which might change the interpretation of the results in this chapter. The feature of the data that estimates the offer distribution is where men and women are hired from nonemployment. Outside the structure of the model, this data can arise because of either supply or demand factors: it might be that employers do not discriminate at all and men and women simply apply different places. Then the question is whether we would still want to interpret this endogenous decision as reflecting discrimination. To make progress on this question would require a model in which the decision of where to apply is endogenous and arises from an optimizing decision—i.e. a directed search model.

---

[12]Admittedly, the earnings gap between young and old workers is larger than between men and women. It is 51 log points, as opposed to 33 log points for the gender gap.

Table 2.1: Summary statistics and the variance of earnings and sorting

	Men			Women		
	All	S. Connected by EE	S. Connected by EE (restrictions) (men/women)	All	S. Connected by EE	S. Connected by EE (restrictions) (men/women)
	(1)	(2)	(3)	(4)	(5)	(6)
Sample size:						
People-years	259,607,000	233,475,000	198,228,000	245,337,000	223,113,000	197,276,000
People	54,547,000	50,474,000	45,569,000	51,374,000	47,918,000	44,418,000
Employers	4,724,000	1,279,000	519,000	4,556,000	1,237,000	519,000
Summary statistics						
Mean log earnings	10.60	10.62	10.63	10.25	10.27	10.30
Variance of log earnings	0.71	0.69	0.71	0.62	0.61	0.60
Share of variance of earnings explained by each parameter set						
Employers	N/A	0.22	N/A	N/A	0.23	N/A
People	N/A	0.55	N/A	N/A	0.54	N/A
Covariates	N/A	0.13	N/A	N/A	0.11	N/A
Overall fit of AKM decomposition						
R^2	N/A	0.89	N/A	N/A	0.87	N/A
Adj. R^2	N/A	0.86	N/A	N/A	0.83	N/A
Sorting: share of co-workers that are men for:						
Women	0.38		0.40			
Men	0.64		0.60			

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 $3250 ($2011). The table includes person-years in which on December 31 of the year the person was 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 EE transition, and hire a worker from nonemployment.

Table 2.2: Decomposing the earnings gap

	Men	Women	Gap	(pct. of overall gap)
	(1)	(2)	(3)	(4)
Mean log wage	10.63	10.30	0.33	
A. Overall				
Firm earnings overall	0.09	0.01	0.09	(27)
Firm earnings for men	0.10	0.01	0.09	(27)
Firm earnings for women	0.09	0.00	0.09	(27)
B. Industry-level				
Industry earnings overall	0.08	0.02	0.07	(21)
Industry earnings for men	0.10	0.03	0.07	(21)
Industry earnings for women	0.06	0.000	0.06	(18)
C. County-level				
County earnings overall	0.05	0.05	0.01	(3)
County earnings for men	0.10	0.09	0.01	(3)
County earnings for women	0.01	0.000	0.01	(3)

In panel's A - C, the men column weights the prices given in the row name by where men are employed, while the women column weights the prices given in the row name by where women are employed. Column (3) then takes the difference between column (1) and (2).

Table 2.3: Transition probabilities and model parameters

Symbol	Description	Men	Women
	Overall EE Transition Probability	0.10	0.10
	Overall EN Transition Probability	0.15	0.14
	EE Share of Transitions	0.39	0.41
	Pr(displacement | EE)	0.28	0.25
	Pr(displacement | EN)	0.36	0.33
δ	Exogenous EN probability	0.05	0.05
ρ	Exogenous EE probability	0.03	0.03
λ_1	Probability of offer on-the-job	0.25	0.25

All probabilities and parameters are annual. The sample for the transition probabilities is column (1) of Table 2.1. 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 2.1. The ρ is related to the calculated probability of making an exogenous EE transition by $(1 - \delta)\rho$. λ_1 is estimated from the model.

Table 2.4: Relationship between men's and women's values and earnings

A. Overall Values (V^e)

	All	Men	Women
All	1	0.97	0.96
Men		1	0.89
Women			1

B. Overall Earnings (Ψ)

	All	Men	Women
All	1	0.98	0.97
Men		1	0.92
Women			1

C. Industry Values (V^e)

	All	Men	Women
All	1	0.99	0.99
Men		1	0.98
Women			1

D. Industry Earnings (Ψ)

	All	Men	Women
All	1	0.99	0.99
Men		1	0.97
Women			1

E. County Values (V^e)

	All	Men	Women
All	1	0.97	0.96
Men		1	0.86
Women			1

F. County Earnings (Ψ)

	All	Men	Women
All	1	0.97	0.96
Men		1	0.87
Women			1

Industry is 4-digit industry (there are 312 of them).

Table 2.5: Comparing the offer distributions

	Men (1)	Women (2)	Gap (3)	(4)
A. Overall				
Firm earnings overall	-0.07	-0.11	0.05	
Firm earnings for men	-0.07	-0.11	0.04	
Firm earnings for women	-0.06	-0.12	0.05	
B. Industry-level				
Industry earnings overall	-0.02	-0.06	0.04	
Industry earnings for men	-0.01	-0.05	0.04	
Industry earnings for women	-0.03	-0.08	0.04	
C. County-level				
County earnings overall	0.04	0.04	0.00	
County earnings for men	0.09	0.08	0.01	
County earnings for women	0.00	0.00	0.00	

In panel's A - C, the men column weights the prices given in the row name by the offers that men are estimated to receive, while the women column weights the prices given in the row name by the offers that women are estimated to receive. Column (3) then takes the difference between column (1) and (2).

Table 2.6: Welfare consequences of sorting	Men	Women	Gap
	(1)	(2)	(3)
A. Overall			
Firm utility overall	1.25	1.23	0.01
Firm utility for men	1.22	1.15	0.07
Firm utility for women	1.41	1.42	- 0.01
B. Industry-level			
Industry utility overall	1.24	1.25	-0.01
Industry utility for men	1.22	1.20	0.02
Industry utility for women	1.40	1.42	-0.02
C. County-level			
County utility overall	1.24	1.24	0.00
County utility for men	1.22	1.21	0.01
County utility for women	1.42	1.42	0.00

In panel's A - C, the men column weights the prices given in the row name by where men are employed, while the women column weights the prices given in the row name by where women are employed. Column (3) then takes the difference between column (1) and (2).

Table 2.7: Age decomposition results

	Old	Young	Gap
	(1)	(2)	(3)
A. Earnings			
Firm earnings overall	0.09	0.00	0.09
Firm earnings for old	0.06	-0.02	0.08
Firm earnings for young	0.11	0.01	0.09
B. Utility			
Firm utility overall	1.44	0.95	0.49
Firm utility for old	1.08	0.67	0.41
Firm utility for young	1.27	0.89	0.38

In panel's A - B, the old column weights the prices given in the row name by where older workers are employed, while the young column weights the prices given in the row name by where the younger workers are employed. Column (3) then takes the difference between column (1) and (2).

Figure 2.1: Women are at lower-paying firms no matter how measured

(a) Men's earnings changes
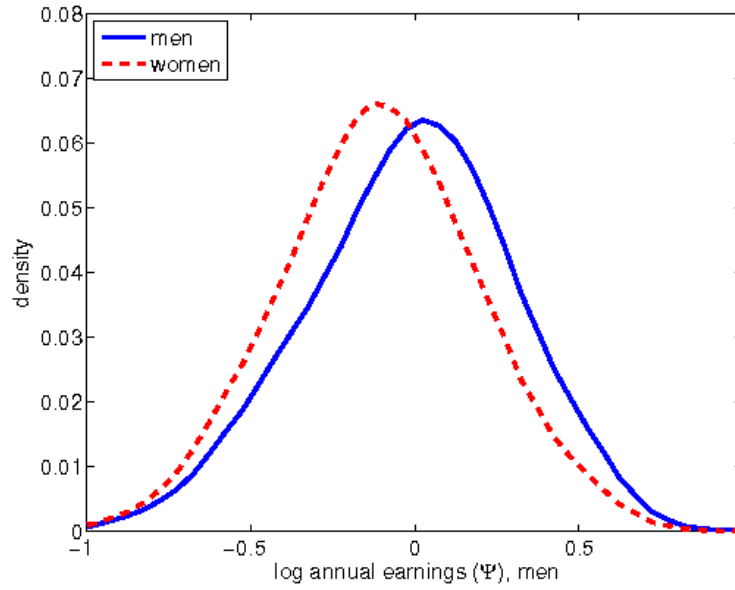
(b) Women's earnings changes
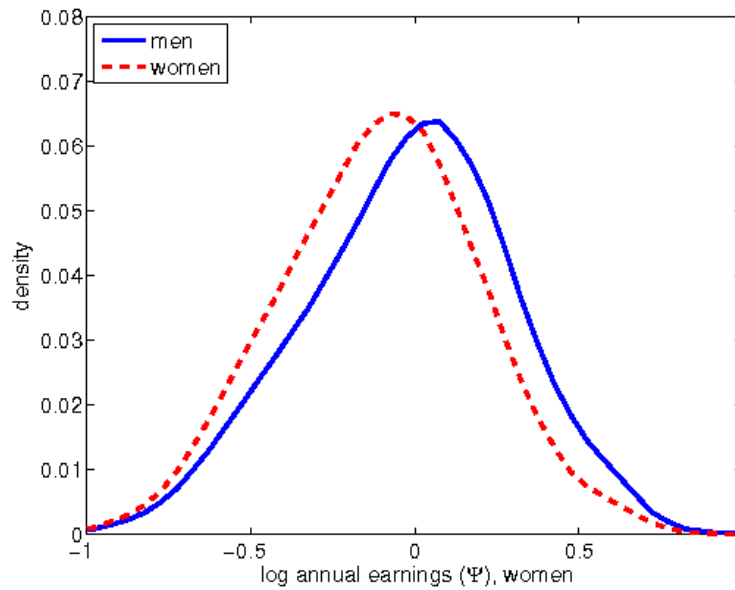
These figures show distribution of firm-level earnings, when these earnings are computed using the earnings changes of different groups of workers. The sample of firms is the same in all three pictures: columns (3) and (6) in table 2.1. The earnings are normalized so that median value in the male distribution is 0.

Figure 2.2: Women search from a lower-paying offer distribution, no matter how measured

(a) Men's earnings changes
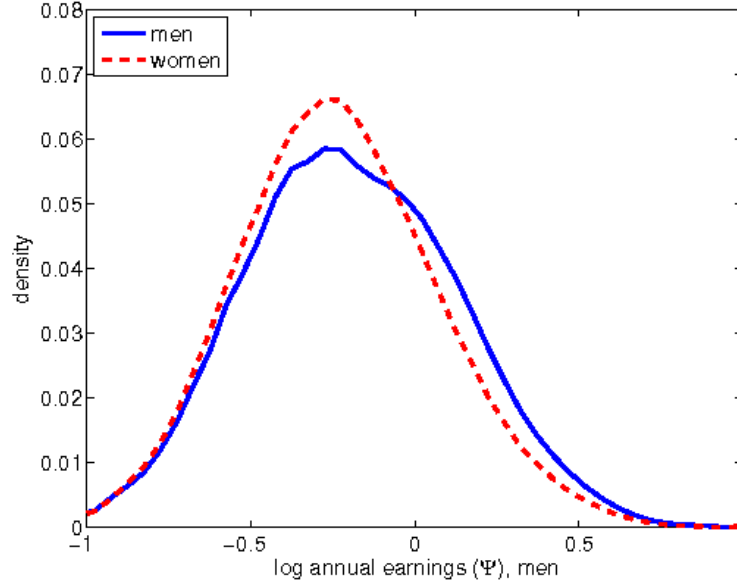
(b) Women's earnings changes
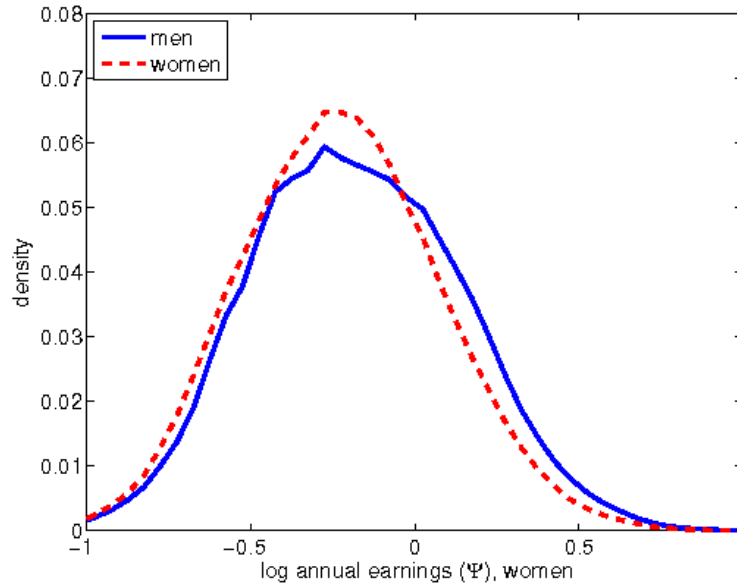
These figures show the offer distribution facing men and women, where the offer distribution is parameterized using the firm-level earnings estimated on men and women separately. The sample of firms is the same in both pictures: columns (3) and (6) in table 2.1. The earnings are normalized so that 0 is the median in the male *employed* distribution (i.e. the male distribution in figure 2.1).

Figure 2.3: Women are at similarly-valued firms as men, no matter how measured

(a) Men's choices
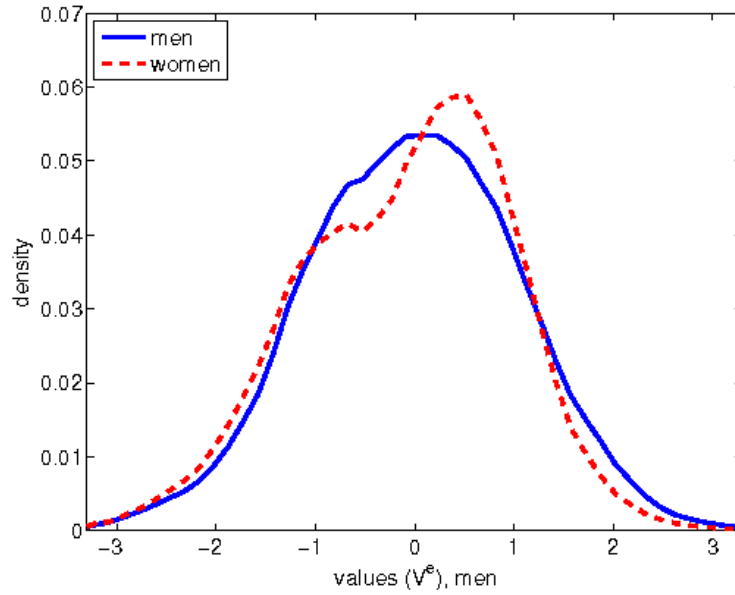
(b) Women's choices
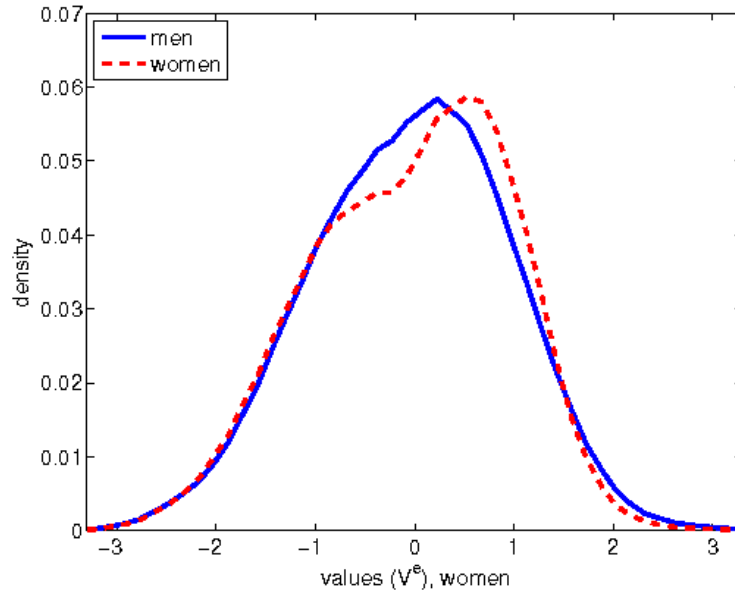
These figures show distribution of firm-level utility, when these utilities are computed using the choices of different groups of workers. The sample of firms is the same in both pictures: columns (3) and (6) in table 2.1. The utilities are first all relative to a common firm, and then normalized so that median value in the male distribution is 0.

CHAPTER III

Are there long-run effects of the minimum-wage?

Inflation and rising real wages make most minimum wage increases temporary.[1] As such, the empirical minimum wage literature has made substantial progress estimating the *short-run* employment effects of minimum wage increases. This effect appears to be small.[2] Despite apparent consensus, the profession remains divided about the employment effects of minimum wage increases.[3]

A reasonable reading of this divide is that there are some questions about the effects of minimum wage increases for which the empirical consensus provides the answer. For other questions, however, economists extrapolate differently depending on whether they think that the relevant short- and long-run employment elasticities differ.[4] To the question: "what is the employment effect of a temporary nominal minimum wage increase likely to be?", the empirical consensus suggests that there are unlikely to be significant employment effects because similar increases have not

---

[1] This observation is not new. Stigler (1946, pg. 358) opens with: "The minimum wage provisions of the Fair Labor Standards act of 1938 have been repealed by inflation." Baker, Benjamin, and Stanger (1999) also emphasize the temporary nature of minimum wage increases in the United States. Of course, there might still be persistent shocks. See the end of this introduction and section 3.6 for more discussion.

[2] Card and Krueger (1995), Brown (1999) and Neumark and Wascher (2008) survey the vast minimum wage literature. Autor, Manning, and Smith (2010, pg. 4, n 6)write "We assume no disemployment effects at the modest minimum wage levels mandated in the US, an assumption that is supported by a large recent literature." Such literature includes Card and Krueger (1994) and Dube, Lester, and Reich (2010). There exist papers that find larger elasticities (e.g. Baskaya and Rubinstein (2012)).

[3] An IGM Expert Panel question on February 26, 2013 asked for responses to the statement: "Raising the federal minimum wage to $9 per hour would make it noticeably harder for low-skilled workers to find employment." In the certainty-weighted responses 40% agreed and 38% disagreed with 22% uncertain. See `http://www.igmchicago.org/igm-economic-experts-panel/poll-results?SurveyID=SV_br0IEq5a9E77NMV` .

[4] Indeed, Brown (1995) and Hamermesh (1995) discuss the possibility of the short- and long-run differing as an important limitation of the results in Card and Krueger (1995). And Neumark and Wascher (2008, pg. 65) emphasize the importance of considering long-run impacts.

resulted in significant employment effects. To the question: "what is the employment effect—after a few years—of a permanent minimum wage increase?," the empirical consensus suggests an answer only if the short- and long-run elasticities of minimum wage increases are the same. In the United States, this latter question is of immediate policy relevance: President Obama's 2013 State of the Union address contained a proposal to index the Federal minimum wage to inflation, which would be a more permanent increase.

To contribute to this important debate, this paper studies the empirical implications of a model that has a distinction between the short- and long-run employment elasticities. The model is based on the putty-clay nature of capital. It was first informally discussed in the minimum wage context by Card and Krueger (1995, pg. 366-8) and I build on the Gourio (2011) version.[5] In the model, when firms pay the entry cost of building a machine, they can freely substitute between capital and labor. Once capital is installed, a firm cannot change its labor demand. The key features of the model are that the labor demand choice of an entering firm is a forward-looking, dynamic, decision that depends on the (expected) stochastic process for minimum wages. And because only some firms adjust each period, the industry-level labor demand response to a minimum wage increase is slow, and also depends on the stochastic process for minimum wages.

The model has two main empirical implications. The first empirical implication is that the reduced-form long-run effects estimated in the literature are essentially uninformative about the true long-run elasticity. I simulate employment data from the model to replicate the dataset used in Dube, Lester, and Reich (2010).[6] They find very small short-run employment effects and, using a common reduced-form long-run regression, no distinction between the short- and long-run employment effects of minimum wages in the United States. They interpret these results as evidence

---

[5]Putty-clay technology was originally developed in Johansen (1959). The main contribution relative to Gourio (2011) is to place the model in industry equilibrium by adding a product demand curve to endogenize product prices and study the dynamics of labor demand. Previously putty-clay technology has been used to study the effect of energy price changes on the economy (e.g. Atkeson and Kehoe (1999)), business cycles (Gilchrist and Williams (2000)) and asset pricing (Gourio (2011)). By deriving dynamic labor demand from dynamic industry equilibrium, the model in this paper is similar to the dynamic labor demand model based on embodied technology in Caballero and Hammour (1994). Aaronson and French (2007) also discuss putty-clay technology in the minimum wage context.

[6]Neumark, Salas, and Wascher (2013) argue that the identification strategy in Dube, Lester, and Reich (2010) is problematic because border counties are not a good control group. In simulated data, the border counties are an appropriate control group for the empirical exercise the literature has focused on.

against the view that short- and long-run elasticities differ.[7] On the simulated data, however, the reduced-form regression recovers a long-run employment effect that is barely different than the short-run employment effect.[8]

The second empirical implication is that the putty-clay model is consistent with the pass-through of minimum wage increases to product prices commonly found in the literature, even though minimum wage increases are relatively temporary. Card and Krueger (1994, pg. 792) emphasize that their finding of product price rises in response to minimum wage increases are inconsistent "with models in which employers face supply constraints (e.g., monopsony or equilibrium search models)." Despite this, the minimum wage literature has focused on models of search frictions to rationalize the small employment effects, without focusing on the price results.[9]

Figure 3.1 suggests why the stochastic process apparently generating US Federal minimum wage variation is unpromising for finding long-run effects: the variation in the real value of the Federal minimum wage follows a "sawtooth" pattern of regular nominal increases that are temporary because they are eroded by inflation and rising real wages. Meer and West (2013, pg. 10) provide evidence that state-level variation is similar.[10] As such, other countries might present promising opportunities for finding long-run effects. Unfortunately, the literature suggests that such opportunities are few and far between (or difficult to exploit). For example, in Dolado et al. (1996)'s comprehensive survey of minimum wages in Europe, they do not distinguish between short- and long-run responses. Pereira (2003) studies an interesting coverage change

---

[7]This interpretation has entered the policy discussion. Greenstone (2013) writes: "The empirical evidence now pretty decisively shows no employment effect, even a few year later. See Dube, Lester and Reich in the REStat."

[8]These panel regressions are a generalization of difference-in-diferences. An important challenge in implementing difference-in-differences is to identify a suitable "control" group for the "treated" location. Difference-in-difference faces an equally important challenge to identify the "after" period that contains the treatment effect of interest. In the minimum wage case, the treatment effect of interest is the long-run employment response, which is confounded both by the presence of dynamics in the response, and by the time-varying treatment.Keane and Wolpin (2002) highlight a similar challenge with measuring the behavioral response to welfare.

[9]Examples include Rebitzer and Taylor (1995),Burdett and Mortensen (1998), Flinn (2006), and Ahn, Arcidiacono, and Wessels (2011). Bhaskar and To (1999) is an exception. They demonstrate in a monopsony model with entry and exit that if a minimum wage increase induces enough exit, there is enough substitution towards capital, and product market power increases rapidly enough in exit, then it is possible to have both product price increases and employment increases. The theoretical contribution of this paper relative to Bhaskar and To (1999) is to offer a way of reconciling price and employment results in a model with perfect competition in both the product and employment markets.

[10]In the 2000s, 10 states started indexing their minimum wages to the CPI. Seven states did so in 2007 and one in 2006. Because of the comparatively small number of changes, the Federal minimum wage increase in mid-2007 and a recession starting in late 2007 , this does not provide a compelling source of variation to study the long-run effects of indexed changes.

117

in Portugal, and finds slightly larger effects at a two year horizon than a one year horizon, but does not study longer-run effects. And Lemos (2007) emphasizes that in Brazil minimum wage increases are similarly temporary due to high inflation. The main exception—also emphasized by Neumark and Wascher (2008)—is Baker, Benjamin, and Stanger (1999). They find larger long-run effects of minimum wage increases in Canada than in the US and suggest that this is due to the variation in the US being less permanent than that in Canada.[11] Their long-run elasticity is around −0.6, which is similar to the long-run elasticity in my model.

My paper complements the work in Baker, Benjamin, and Stanger (1999) by developing an economic model which formalizes when the nature of the variation might matter, develops conditions under which the common reduced-form approach to estimating long-run effects is valid, and shows via simulation how quantitatively important the temporariness of the variation might be in masking the long-run effect of minimum wage increases.[12]

This paper proceeds as follows. Section 3.1 discusses evidence on adjustment costs, and other mechanisms in the model. Section 3.2 develops the putty-clay model of labor demand. Section 3.3 analyzes the implications of the model for the interpretation of empirical work. Section 3.4 discusses the calibration and Section 3.5 develops stylized simulations of the model. Section 3.6 shows that given US variation in minimum wages, the estimated short- and long-run employment effects on data simulated from the model do not differ. Section 3.7 concludes.

3.1 Model assumptions

3.1.1 Adjustment costs on labor

There is a difference between the short- and long-run elasticities if there are adjustment costs for labor. There is a tradition in the minimum wage literature of arguing that high turnover among workers in minimum wage industries means that adjust-

---

[11]Similarly, Meer and West (2013) argue that temporariness of minimum wage increases means that it is easier to find effects in *flows* (hiring and separation rates) rather than *stocks* (employment levels) and find some effects on flows consistent with negative employment effects. Dube, Lester, and Reich (2011), Brochu and Green (2013) and Gittings and Schmutte (2013) present related analyses of labor market flows following minimum wage increases.

[12]Two papers present alternative mechanisms in search models to make "masking" arguments. Pinoli (2010) suggests that anticipation of minimum wage increases might make it difficult to find effects by altering the timing of the treatment. Gorry (2011) shows in a search model that minimum wages can have a nonlinear effect in their level and so the relatively low level of minimum wages in the US might have a relatively smaller effect.

ment costs are likely small.[13] High turnover, however, is evidence against adjustment costs for changing the identity of *workers* rather than for changing the number of *jobs*.[14]

There is substantial evidence that establishments incur significant and lumpy costs to adjust the number of *jobs*. Hamermesh (1989) documents that establishments adjust their number of jobs infrequently, while Hamermesh, Hassink, and van Ours (1996) documents that even without adjusting the number of jobs the identity of workers changes. These findings have been confirmed and amplified in other data.[15] Under the assumption that the underlying shock process is relatively smooth, the infrequency of adjustment of the number of jobs points to the importance of lumpy or fixed adjustment costs at the level of the job. Similarly, empirical models of establishment level employment dynamics study lumpy adjustment costs on the number of jobs (e.g. Cooper and Willis (2009)). Finally, companies think about labor demand in terms of jobs. Lazear (1995, pg. 77) (quoted in Campbell and Fisher (2000)) writes: "Human resource managers think in terms of slots or jobs, and think of these slots or jobs as being fundamental to the organization of the firm."

3.1.2 Entry and exit

The price dynamics in the model derive from the empirical feature of the product market that there is always entry and exit of firms. Aggregate statistics suggest that there is always entry and exit. For example, 9% of fast food restaurants that existed in March 2009 exited by March 2010.[16] And there is evidence, for example Campbell and Lapham (2004), that entry and exit among restaurants is responsive to economic shocks.

3.1.3 Capital heterogeneity

The model formalizes the following idea. If adopting new machines is the way that firms substitute between capital and labor, then such substitution is unlikely to be caused by minimum wage increases when these increases are (perceived as) sufficiently

---

[13]E.g. Brown, Gilroy, and Kohen (1982, pg. 496) and Card and Krueger (1995, pg. 63).

[14]Identically, Hamermesh (1993, pg. 207) distinguishes between adjustment costs on *gross* and *net* employment changes.

[15]The lumpiness of employment adjustment is not confined to manufacturing (e.g. Davis, Faberman, and Haltiwanger (2006, pg. 10)). The high flows of workers at firms that are not adjusting employment is also not unique to manufacturing (Davis, Faberman, and Haltiwanger (2012)).

[16]This statistic comes from the Statistics of U.S. Businesses. See `http://www2.census.gov/econ/susb/data/2010/us_naicssector_small_emplsize_2010.xls`.

temporary. Sufficiently temporary increases have little impact on the relative price of capital and labor over the life of the capital. On the other hand, if the change in labor costs is permanent, or the differences in labor costs across locations is permanent, then these technologies will be adopted, or adopted more quickly.

A few examples highlight the idea that labor demand is embodied in the firm's choice of capital.[17] McDonald's in Europe plans on replacing many of its cashiers and their registers with touchscreen terminals so that customers do not interact with cashiers.[18] To do so it needs to purchase new registers that embody the labor-saving capital. Similarly, in the grocery store industry, self-checkout scanners represent a labor-saving technology that requires a new capital stock. And self-service gas stations required new gas pumps.

A similar kind of evidence comes from Seltzer (1997). He studies the seamless hosiery—sock—industry in the US in the 1930s, which was hit by the implementation of Federal minimum wages. The fundamental technological choice facing the seamless hosiery industry was whether to use machines where the top of the stocking was knit on a different machine than the stocking itself, or machines where the top was knit on the same machine as the stocking. The most labor-intensive process used the hand-transfer machine, where the top of the stocking was knit on a separate machine and then carried by hand to the knitting machine. What is striking in Seltzer (1997)'s data is that while the lower-wage plants adjusted their capital stock towards the labor-saving technology, the speed of adjustment was relatively slow: two years after the change in relative labor costs, the use of the most labor-intensive machines declined by less than a quarter. Similarly, Lewis (2011) shows how an influx of low-skilled immigrants slowed the adoption of more technologically intensive (labor-saving) capital stocks.

These examples are in line with the model because two choices of $\frac{K}{L}$ require differ-

---

[17]Of historical interest, Lester (1946)'s pioneering study cited this mechanism as a reason why the standard competitive model was misleading for studying the employment effects of minimum wage increases. Lester (1946, pg. 72-73) writes:

> Most industrial plants are designed and equipped for a certain output, requiring a certain work force. Often effective operation of the plant involves a work force of a given size...Under such circumstances, management does not and cannot think in terms of adding or subtracting increments of labor except perhaps when it is a question of expanding the plant and equipment, changing the equipment, or redesigning the plant...the decision to shift a manufacturing plant to a method of production requiring less or more labor per unit of output because of a variation in wages is not one that the management would make frequently or lightly.

[18]See Louise Lucas. "McDonalds to shake up food ordering system," *The Financial Times,* May 15, 2011.

ent machines, rather than being able to vary $\frac{K}{L}$ while using the same machines. Over the long-term, these new kinds of capital are adopted because their price *relative to labor* falls. Minimum wages also affect the relative price of capital and labor. In the examples, the change in the relative price of capital and labor is permanent and so the substitution occurs. The idea of the model is that the change in the relative price of capital and labor induced by a minimum wage increase is temporary and so the substitution is unlikely to occur.

3.2 The putty-clay model

This section develops a dynamic labor demand model in industry equilibrium (e.g. Hopenhayn (1992)). The model has three features that make it well-suited to study minimum wages. First, it models industry-level variables such as prices. Industry equilibrium means that employment is industry-wide employment as in Dube, Lester, and Reich (2010), rather than employment within a continuing set of firms as in Card and Krueger (1994). Second, it explicitly parameterizes the transition path between the short- and long-run elasticities. Third, the model is sufficiently tractable that it is possible to derive analytic results both about steady state to steady state comparative statics, as well as about the response of the model to temporary shocks. Because minimum wage increases are mostly temporary, the data are dominated by temporary shocks, so understanding these is essential.

The model is based on the putty-clay model in Gourio (2011), which he uses to study asset pricing. Putty-clay technology is putty and flexible when firms make their investment decisions. After installation, the capital hardens to clay and firms can no longer adjust the labor needs of the capital. Only a small share of firms have their capital stock expire each period, so market-level labor demand takes time to adjust to a wage increase.

The model has an alternate interpretation as a time-dependent adjustment cost model. The Poisson machine expiration process is a stand-in for the endogenous decision of firms to pay an adjustment cost to change their labor demand. Section 3.2.1.2 discusses why this is not a (particularly) restrictive assumption.

3.2.1 The model

Since this paper is concerned with minimum wages, wages are set exogenous to the industry. Firms in the industry face a sequence of expected wages given by $E_t\left[(w_{t+j})_{j=0}^{\infty}\right]$, where time is discrete. Minimum wages are always binding. The un-

121

certainty in the wage path reflects only uncertainty about minimum wage policy. The labor market is frictionless and there is an infinite supply of labor at each minimum wage.

The model is closed with a downward-sloping product demand curve and an entry condition. The entry condition is that the expected profits of newly entering firms (or machines) are always zero, though of course firms have to pay an entry fee by building machines. Firm entry and exit drives the product price responses.

3.2.1.1 Production technology

The continuum of firms produce from a Cobb-Douglas production function, $y = k^{\alpha}l^{1-\alpha}$, where k is the amount of capital in the machine, and l is the amount of labor. Ex-post, production is Leontief so that k and l are chosen when the machine is built, and are then fixed for the life of the machine.[19] Because of constant returns to scale, normalize to one worker per machine. Setting $l = 1$, $y = k^{\alpha}$, where k is the size of the machine (the capital intensity of production). Machines fail with probability $\delta \in (0, 1)$ each time period.

3.2.1.2 Firm maximization

A firm's entry cost is the cost of building a machine. When a firm builds a machine at time t, it chooses the capital intensity of a machine and then is stuck with that choice for the life of the machine. The capital choice takes into account the present discounted value of product prices and wage, where the discount rate is the combination of failure probability of the machine and the market discount rate, β. The price of a unit of capital is normalized to 1.

Let

$$q_t \equiv E_t \sum_{j=0}^{\infty} \beta^j (1-\delta)^j P_{t+j} \tag{3.1}$$

be the expected effective present discounted value of product prices (P_t) and let

$$q_{w,t} \equiv E_t \sum_{j=0}^{\infty} \beta^j (1-\delta)^j w_{t+j} \tag{3.2}$$

be the expected effective present discounted value of wages. These quantities are referred to as present discounted values of product prices and wages. The relative

---

[19]In order to make the Leontief assumption binding, capital is industry-specific so that the resale price of capital reflects the shocks facing that industry.

prices that the firm faces reflect time-varying prices and wages, but capital intensity is fixed and the capital cost is paid once.

A firm chooses k_t to maximize

$$\Pi_t = q_t k_t^\alpha - q_{w,t} - k_t. \tag{3.3}$$

k_t is the capital choice of firms that enter in time t, while k is all active capital in the industry. Writing the maximization problem in terms of present discounted values of product prices and wages reflects a no-shutdown assumption. Once capital is in place, a firm's shutdown decision treats capital costs as sunk. A firm that invested in period t' operates in period $t > t'$ if $P_t k_{t'}^\alpha - w_t \geq 0$. Following Gourio (2011) a machine, once installed, is always operated. This no-shutdown assumption never binds in the simulations below, which means that the Poisson adjustment is not too restrictive.[20]

There is good economic reason for the no-shutdown assumption to never bind. In the model, the only shock facing firms is a minimum wage increase. Following a minimum wage increase, incumbent firms see their flow profits change for two reasons: wages rise and the product price rises. For *local* changes from steady state these two effects exactly cancel, which limits how quantitatively important endogenous exit would be in speeding up adjustment.[21] For larger changes these effects do not exactly cancel and incumbents see their flow profits fall. That incumbents are (partially) protected by equilibrium responses is the "insulation effect" of Caballero and Hammour (1994).

3.2.1.3 Product price determination

Two features of the model together pin down the product price: a free entry condition and a product demand curve. The free entry condition does not mean that entry is free: firms need to construct their capital stock when they enter.

Denote gross entry (investment) by h_t, the number of machines built at time t. Free entry implies that

$$h_t \Pi_t = 0 \tag{3.4}$$

for all t. If there is gross entry, then there are zero expected profits—net of the cost of constructing the capital stock—from entering. Following Gourio (2011), there is positive gross entry for all time t so that there are zero expected profits from entering

---

[20]See 3.8.2 for details.

[21]See 3.8.2 for a formal derivation. This results does not hold for non-local changes and if the elasticity of substitution between capital and labor is less than one.

$(\Pi_t = 0 \forall t)$. As with the no-shutdown assumption, this assumption is never binding in the simulations.[22]

The industry faces an isoelastic product demand curve

$$Q_t = \theta P_t^{-\gamma}, \tag{3.5}$$

where Q is market quantity and P is the product price. This demand curve is consistent with the industry making up a small portion of the economy, and the exogenous product demand and factor prices are standard in models of industry equilibrium (e.g. Hopenhayn (1992)). It does rule out general equilibrium explanations for small employment effects of minimum wages such as the argument caricatured by Kennan (1995, pg. 1961) as "teenagers like cheeseburgers."

3.2.1.4 Aggregation and laws of motion

Employment evolves as machines expire with probability δ each time period and the new investment is implemented. The law of motion for machines (and employment) is

$$N_t = (1 - \delta)N_{t-1} + h_t. \tag{3.6}$$

Aggregate output is given by integrating over the distribution of the capital stock of all ages at time t, G_t (recall that k_t is the capital chosen by entering firms at time t and k represents all active capital in the industry):

$$Q_t = \int_0^\infty k^\alpha dG_t(k).$$

Identical to employment, output evolves as machines expire and new investment is implemented:

$$Q_t = (1 - \delta)Q_{t-1} + h_t k_t^\alpha. \tag{3.7}$$

3.2.1.5 Equilibrium

An *equilibrium* in the model is a sequence of endogenous variables $\{k_t, h_t, P_t, Q_t, N_t\}_{t=-\infty}^{+\infty}$ that takes the sequence of realized $\{w_t\}_{t=-\infty}^{+\infty}$ and expected, $\{E_t\left[(w_{t+j})_{j=0}^\infty\right]\}_{t=-\infty}^{+\infty}$, minimum wages as given such that in every time period:

---

[22]See 3.8.2 for details.

- Firm's choose k_t to maximize profits subject to their production technology (equation (3.3));

- The entry condition is satisfied (equation 3.4));

- The product market clears (equation (3.5)); and

- The laws of motion for employment and output hold (equations (3.6) and (3.7)).

A *steady state equilibrium* is an equilibrium in which the the sequence of realized and expected minimum wages are constant. The resulting endogenous variables are all constant as well.

3.2.1.6 Equilibrium computation

Equilibrium computation is straightforward when there is entry in every period and the no-shutdown assumption holds. The core of the model can be reduced to three equations in three unknowns:

- The firm's first order condition of equation (3.3) implies one equation in three unknowns $(q_{w,t}, k_t, q_t)$.

- The sequence of expectations of minimum wages provides a second equation through the sequences of present discounted value of minimum wages: $q_{w,t}, q_{w,t+1}...$

- The entry condition with entry (3.4) implies that the firm's profit in equation (3.3) is equal to zero, providing the third equation.

The resulting capital intensity is:

$$k_t = \frac{\alpha}{1-\alpha} q_{w,t}. \tag{3.8}$$

A firm considers the present discounted value of future relative prices when making its investment decision, unlike in a static model where a firm only considers time t relative prices.

Solving for the expectation of next time period's present discounted value of product prices is straightforward. The sequence of equilibrium present discounted value of product prices can be rewritten as the current product price:

$$P_t = q_t - \beta(1-\delta)E_t[q_{t+1}]. \tag{3.9}$$

A minor complication arises in computing the equilibrium when there is uncertainty in the wage path because the present discounted value of prices, q_t, is a nonlinear function of the present discounted value of wages, $q_{w,t}$.

3.2.2 Comparison to long-run elasticities

The putty-clay model is a dynamic version of the textbook static labor demand elasticities. In particular, the long-run (steady state) elasticities in the model match the long-run textbook results exactly.[23]

3.8.1 derives expressions for the number of workers and the product price in steady state in the putty-clay model. Differentiating these expressions with respect to the present discounted value of wages gives the long-run elasticities for employment

$$\frac{\partial N}{\partial q_w}\frac{q_w}{N} = -\underbrace{\gamma(1-\alpha)}_{\text{scale}} - \underbrace{\alpha}_{\text{substitution}} \tag{3.10}$$

and for prices

$$\frac{\partial P}{\partial q_w}\frac{q_w}{P} = 1 - \alpha. \tag{3.11}$$

These elasticities are identical to the long-run textbook elasticities with respect to the *wage* and imply complete pass-through of the minimum wage increase to product prices. The employment effect operates through two channels. The *scale* effect measures the reduction in employment because of the contraction in the size of the product market from the product price increase. The *substitution* effect measures the reduction in employment because of substitution between capital and labor.

The difference between the putty-clay model and the textbook framework emerges in the short-run and in the transition to the long-run. In the textbook model, the capital-labor ratio can adjust for *all* firms in the short-run, while in the putty-clay model only some firms adjust the capital-labor ratio. In the textbook model, no firms can acquire capital in the short-run, while in the putty-clay model all firms have the option of doing so. Unlike the textbook model, the putty-clay model explicitly parameterizes the transition from the short- to long-run. As the rate of capital expiration, δ, increases the long-run arrives more quickly.

---

[23]See Hamermesh (1993, pg. 24) for the long-run elasticities and Cahuc and Zylberberg (2004, pg. 172) for the distinction between the short- and long-run used here.

3.3 Interpreting empirical work in light of the model

This section analytically shows how employment and product prices move in response to temporary and permanent minimum wage increases. Surprisingly, product prices move even in response to temporary minimum wage increases.

In this section, a temporary minimum wage increase is an increase that lasts for a single time period. This is a stylized way of capturing what happens if the minimum wage is increased in nominal terms and then eroded by inflation. Section 3.5 numerically studies environments where temporary minimum wage increases are more persistent than one period and where real minimum wages follow a sawtooth pattern. The perfect foresight assumption implicit in this section is relaxed in section 3.6.

3.3.1 Short-run employment response to temporary increase is through the change in output

The minimum wage literature has focused on estimating short-run employment response to minimum wage increases. Card and Krueger (1994) look at a 9 month window around a minimum wage increase, while in the panel data literature, e.g. Neumark and Wascher (1992) and Dube, Lester, and Reich (2010) the focus is on the response within a quarter. These estimates are interpreted as testing "[t]he prediction from conventional economic theory" (Card and Krueger (1994, pg. 772)) about the employment effects of minimum wage increases. In the model, these very short-run employment responses capture the part of the employment response that operates through the change in output (the *scale* effect), and very little of the substitution of labor for capital (the *substitution* effect). This second effect is quantitatively much larger in my simulations.

The following result shows how these two effects are reflected in the contemporaneous employment response to a temporary minimum wage increase. The proof of this and all other results are in 3.8.3.

Result 4. *The elasticity of contemporaneous employment with respect to a temporary minimum wage hike from steady state is:*

$$\frac{\partial N_t}{\partial w_t} \frac{w_t}{N_t} = - \underbrace{\gamma(1-\alpha)}_{scale} - \underbrace{\alpha\delta(1-\beta(1-\delta))}_{substitution}.$$

The first term shows that the full scale effect occurs contemporaneously, which

127

follows from the result in a subsequent subsection that prices move with wages. Unlike the scale effect, the substitution effect is dramatically attenuated relative to the long-run benchmark of α displayed in Equation (3.10). The attenuation occurs because each period only δ share of firms adjust and because the capital decision is forward-looking and the increase is temporary, the firms that adjust engage in very little capital-labor substitution. This second form of attenuation is captured by the $1 - \beta(1 - \delta)$ term, which can be very small.

3.3.2 Long-run observed employment response depends on how permanent the increase is

While the literature has focused on the short-run employment response, as a robustness exercise some papers include lags of the minimum wage to capture any potential long-run effects (e.g. Neumark and Wascher (1992, Table 5), Baker, Benjamin, and Stanger (1999, Table 7) and Dube, Lester, and Reich (2010, Equation 7)). That the coefficients on these lags are often quite small is interpreted as evidence that the long-run employment effects of a minimum wage increase are the same as the short-run employment effects (Dube, Lester, and Reich (2010, pg. 956)).[24]

Formally, the standard robustness exercise builds on the following distributed lag form:

$$\ln N_{i,t} = \alpha_i + \alpha_t + \beta_0 \ln w_{i,t} + \sum_{j=1}^{J} \beta_j \ln w_{i,t-j} + \epsilon_{i,t}, \qquad (3.12)$$

where N is employment in location i at time t, $w_{i,t}$ is the measure of minimum wages at location i at time t and $w_{i,t-j}$ is lags of the minimum wage. The model provides a "structural" interpretation of this robustness exercise. Equation (3.34) shows that employment in time t is a function of past, current and expected minimum wages and parameters of the model. A first-order expansion of equation (3.34) with respect to a minimum wage increase in a given time period provides a justification for estimating equation (3.12) since it contains employment on the left hand side and functions of the minimum wage and parameters of the model on the right hand side.

When is the structural interpretation warranted? When there is a distinction between the short- and long-run elasticities ($\delta \neq 1$), two conditions jointly imply that the coefficients of the regression in equation (3.12) are structural parameters

---

[24]Greenstone (2013) echoes this interpretation: "The empirical evidence now pretty decisively shows no employment effect, even a few years later. See Dube, Lester and Reich in the REStat." Similarly, Card and Krueger (1995, pg. 366-8) cite the absence of evidence on long-run employment effects as an argument against the importance of dynamic models such as putty-clay.

and do not depend on the nature of minimum wage variation. First, the minimum wage increase is one-time and permanent and is perceived as permanent. Second, employment was initially in steady state. The following result provides the explicit time-path that such a regression would return.

Result 5. *If $q_{w,j} = E_j[q_{w,j+1}] = q_w \ \forall j \leq t$ and the change in wages at time t is unexpected and permanent (and perceived as such), then*

$$\beta_0 = \frac{\partial N_t}{\partial q_{w,t}} \frac{q_{w,t}}{N_t} = -\gamma(1-\alpha) - \delta\alpha \tag{3.13}$$

$$\beta_j = \frac{\partial N_{t+j}}{\partial q_{w,t}} \frac{q_{w,t}}{N_{t+j}} = -(1-\delta)^j \delta\alpha, \ \forall j > 0. \tag{3.14}$$

Moreover, $\sum_{j=0}^{\infty} \frac{\partial N_{t+j}}{\partial q_{w,t}} \frac{q_{w,t}}{N_{t+j}} = -\gamma(1-\alpha) - \alpha$, which is the elasticity given in (3.10).[25]

The temporary nature of the mininum wage hike mutes the observed long-run effects of the minimum wage hike for two reasons. First, the contemporaneous response is smaller when firms expect the hike to be temporary because the extent of firms' adjustment depends on the increase in the expected present discounted value of wages. Formally, $\frac{\partial q_{w,t}}{\partial w_t} \frac{w_t}{q_{w,t}} < 1$. Second, the subsequent adjustments are made to smaller and smaller minimum wage increases. Formally, $\frac{\partial q_{w,t+j}}{\partial q_{w,t}} \frac{q_{w,t}}{q_{w,t+j}} < 1$. Hence, if the short- and long-run elasticities to permanent minimum wage hikes differ, then both the short- and long-run reduced-form estimates would be biased down relative to the structure for temporary minimum wage hikes.

Being out of steady state can also induce bias. The bias is ambiguous in sign. If the capital stock was on average installed when the minimum wage was expected to be higher than it is after the increase, then employment can actually rise following a minimum wage increase. Conversely, the employment decline can be larger if the capital stock was installed when the minimum wage was expected to be lower than it is after the increase.

3.3.3 Short- and long-run observed price response: Product prices move with the wage

While much of the minimum wage literature since Card and Krueger (1994) has focused on their employment results, they also document some evidence of an increase in the product price following minimum wage increases. Subsequent work by Aaronson (2001) and Aaronson, French, and Macdonald (2008) has extended and

---

[25]Since this is an elasticity, it is for local changes so that $N_{t+j} = N_t$ for all j.

confirmed this finding. Card and Krueger (1994), Aaronson and French (2007) and Aaronson, French, and Macdonald (2008) emphasize that these price increases are consistent with full pass-through of the minimum wage increase to product prices and inconsistent with employment increases. In most models, if prices go up then output falls so inputs, including labor, have fallen.

The putty-clay model is consistent with full pass-through of the minimum wage increase to product prices in the short-run because there is always entry (and exit). With entry, the market price is set by the marginal entrant and so it is responsive to changes in market conditions. Consider an increase in minimum wages that is known to last only one quarter. To see what has to happen to prices in the quarter of the increase, let us examine what happens after the minimum wage has fallen again. In quarters after the minimum wage has fallen, new entrants face the same expected costs as before the increase. So free entry implies that the product price immediately returns to its pre-increase level. Now let us return to the quarter of the increase. Potential entrants in the quarter of the minimum wage increase face higher costs that quarter, but know that prices will fall immediately. To compensate entrants in the quarter of the minimum wage increase, there must be complete pass-through of the minimum wage increase to the product price. There is an additional channel that generates more than complete pass-through. If entrants adjust their input mix so that it is suboptimal in later periods, then to compensate firms for this distortion the contemporaneous product price response is larger following a temporary minimum wage increase than a permanent increase.

The following result shows that the contemporaneous price response to a temporary minimum wage increase is very similar to the response to a permanent increase documented in Equation (3.11). When firms are "myopic" and assume that the current wage lasts forever, they do not need to be compensated for having a suboptimal input mix in later periods. In this case, the contemporaneous response is the same to temporary and permanent increases.

Result 6. *Under perfect foresight, the elasticity of the product price with respect to the contemporaneous wage is:*

$$\frac{\partial P_t}{\partial w_t}\frac{w_t}{P_t} = \underbrace{(1-\alpha)}_{\text{long-run}}\underbrace{\frac{w_t}{\tilde{w}_t}\frac{1-\beta(1-\delta)}{1-\beta(1-\delta)\left(\frac{\tilde{w}_{t+1}}{\tilde{w}_t}\right)^{1-\alpha}}}_{\text{adjustment term}}$$

where $\tilde{w}_t = (1-\beta(1-\delta))q_{w,t}$ is the flow equivalent of period t's present discounted

value of wages and $\tilde{w}_{t+1} = (1 - \beta(1 - \delta))q_{w,t+1}$ is the flow equivalent of period $t + 1$'s present discounted value of wage.

Under myopic expectations ($\beta = 0$):

$$\frac{\partial P_t}{\partial w_t}\frac{w_t}{P_t} = 1 - \alpha.$$

The additional increase in the product price from a temporary increase due to a suboptimal input mix can be seen by considering the two components of the adjustment term separately. Consider first the $\frac{w_t}{\tilde{w}_t}$ term. This term is the ratio of the contemporaneous wage to the flow equivalent of wages. If wages fall over time, then the contemporaneous wage is bigger than the flow equivalent. Consider second the $\frac{1 - \beta(1-\delta)}{1 - \beta(1-\delta)\left(\frac{\tilde{w}_{t+1}}{\tilde{w}_t}\right)^{1-\alpha}}$ term. Its magnitude relative to one depends on whether $\frac{\tilde{w}_{t+1}}{\tilde{w}_t}$ is bigger or less than one. If wages do not fall too rapidly, then this term will be close to one. Hence, the $\frac{w_t}{\tilde{w}_t}$ term will dominate and the adjustment term can be greater than one.

3.4 Parameter values

Much empirical work in the minimum wage literature, including Dube, Lester, and Reich (2010), has focused on the restaurant industry. Hence, I calibrate the model to the restaurant industry. Table 3.1 displays the parameter values.

A key parameter is the share of *minimum wage* workers in firms' expenses, which is set to 0.1. This number is arrived at as follows. Aaronson and French (2007) report that the labor share is 0.3 and the minimum wage worker share in the wage bill is 0.17. They report that about one-third of workers are minimum wage workers, but about two-thirds of workers are low-skill. They attribute the fact that not all low-skill workers are paid the minimum wage to wage dispersion across labor markets, rather than to skill dispersion among low-skill labor workers (pg. 181). Since in the model the minimum wage is binding in all labor markets (and Dube, Lester, and Reich (2010) treat all labor markets identically), I assume that all low-skill workers earn the minimum wage. Hence, low-skill workers account for $0.3 \times 0.17 \times 2 \approx 0.1$ share of firms' expenses.

The choice of capital pins down materials use as well as high and low-skill labor use. As a result, materials and high-skill labor adjusts with the capital stock and so they can be combined. This assumption means the short-run substitutability between minimum wage labor and intermediate inputs and high-skill labor is zero, which is

the same as that between minimum wage labor and capital.[26] Aaronson and French (2007) report a land, structure and machines share of 0.3, material share of 0.4 and my calculations above leave high-skill workers with a share of 0.2. Hence, α, the non-minimum-wage-worker share of inputs, is 0.9. The long-run price elasticity in the model is thus $1 - \alpha = 0.1$, and Aaronson (2001) finds a price elasticity of 0.07. Setting $\alpha = 0.9$ represents a low estimate of the importance of the non-minimum wage worker inputs.

I report elasticities of the combination of high-skill and low-skill employment, not just employment of workers subject to the minimum wage. This choice aligns with Dube, Lester, and Reich (2010) who use total employment in the restaurant industry as their measure of employment. It is also similar to other studies of minimum wages where researchers have not been able to directly identify minimum wage workers (see for example the discussion of studies of teenagers in Brown (1999, pg. 2114)). Letting a bar over a variable represent steady state, define employment in time t, $E_t = N_t + n_h \frac{\bar{N}}{\bar{K}} K_t$, where K_t is the aggregate capital stock and n_h is a constant. The quantity $n_h \frac{\bar{N}}{\bar{K}} K_t$ represents high-skill employment, which moves with the capital stock. Set $n_h = \frac{1}{2}$ so that in steady state low-skill employment, N_t, is two-thirds of total employment. Set the price elasticity of demand for output to $\gamma = 0.6$, which is within the range reported by Aaronson and French. Set the market discount rate, β, to $(0.95)^{\frac{1}{4}} \approx 0.987$ on a quarterly basis, which is standard.

The final parameter is the machine expiration rate, δ. Unfortunately, there is not detailed evidence as to the value of δ in the restaurant industry, though the main result turns out not to be particularly sensitive to δ. Since I build on his model, I use the value in Gourio (2011), 0.08 on an annual basis (0.0206 on a quarterly basis), which he chooses as a high estimate of capital depreciation rates. Two alternative approaches yield similar values of δ. First, δ could be thought of as the exit probability of a fast food restaurant. In 2009-2010, this rate was 0.09.[27] Second, δ could be thought of as the rate at which existing fast food restaurants are remodeled. Some anecdotal evidence on this point comes from McDonald's. As of 2003, many McDonald's restaurants looked "as they did during the Reagan administration." And a plan launched in 2003 to renovate all McDonald's had resulted in only half being renovated

---

[26]The hours margin of minimum wage labor can be ignored without loss of generality. Redefine L to be total minimum wage labor input and then a firm with no reason to adjust total minimum wage labor input also has no reason to adjust hours.

[27]See the U.S. Census Statistics of U.S. Businesses among limited service eating places for 2009-2010 at
`http://www2.census.gov/econ/susb/data/2010/us_naicssector_small_emplsize_2010.xls`.

by 2010.[28] The first anecdote suggests that McDonald's are updated less frequently than once every 15 years (or δ smaller than 0.08), while the second suggests a half-life of 6 years (or δ of about 0.08). In section 3.6.7 I show that the main result of this paper is quite robust to varying δ, even to using $\delta = 0.33$ as suggested in Card and Krueger (1995).

3.5 Quantitative implications of the model with stylized minimum wage variation

To illustrate the quantitative implications of the calibrated version of model, I perform three minimum wage experiments. The first experiment considers a one-time and permanent increase. This experiment corresponds to the *ceteris paribus* condition implicit in interpreting reduced-form long-run regression coefficients as long-run elasticities. The second experiment considers a one-time increase that is eroded by inflation. This experiment shows how the observed long-run response differs dramatically depending on the nature of the variation. The third experiment considers a stylized version of the sawtooth pattern that characterizes US data of repeated temporary increases. This experiment shows how sawtooth variation is expected to affect estimates.

This section and the next section reports reduced-form elasticities of employment with respect to the magnitude of the minimum wage hike. This practice is in keeping with the convention in the empirical literature. One point that this paper makes is that these elasticities are not structural objects, and in fact depend on the nature of the policy change.

To starkly display the quantitative implications of the model, in this section firms have perfect foresight about the minimum wage process. In section 3.6, firms expect that minimum wages follow a stochastic process that is estimated from the data.

3.5.1 One-time changes

I report results of a one-time and permanent 15% increase in the minimum wage from steady state that is permanent, and such an increase that is eroded by 2.2% annual inflation and so is temporary. Consistent with the industry being small relative to the economy, the product price does not feed back into the inflation rate. 3.9 describes the simulation algorithm.

---

[28]Both anecdotes come from: Daniel Smith. "The Renovation Rush." May 2010. QSR Magazine. `http://www.qsrmagazine.com/competition/renovation-rush`.

Figure 3.2 displays the response of employment, product prices and the aggregate capital-labor ratio to the permanent and temporary minimum wage increase. The product price moves with the minimum wage, while the capital-labor ratio and employment are slower moving variables.

Table 3.2 displays the arc elasticities of employment, defined as the percent change in employment divided by the percent change in wages. Column (1) shows that the long-run elasticities of a permanent minimum wage increase are substantial. The contemporaneous elasticity is about -0.06, which is within the range of elasticities discussed by, for example, Brown (1999) and Dube, Lester, and Reich (2010). After six years the employment elasticity with respect to the initial minimum wage increase is -0.25. Once all adjustment is complete, the elasticity is -0.55, which is larger than elasticities discussed in Brown (1999). Column (2) shows that when the increase is temporary the contemporaneous employment effect is similar, but over time the employment effect fades as the minimum wage erodes.

Table 3.3 displays the arc elasticities of prices. Column (1) shows that all of the price response occurs immediately. Column (2) shows that even when the minimum wage hike is temporary, the contemporaneous price response is not muted. In fact, the price response is slightly larger because the firms that enter have to be compensated for the fact that their capital-labor ratio will be suboptimal in all periods. The price response fades as the minimum wage erodes.

3.5.2   Repeated temporary increases: a sawtooth equilibrium

A message of Figure 3.1 is that all minimum wage increases are not permanent and in fact follow a sawtooth pattern of regular temporary increases that are eroded. I simulate the model in a sawtooth equilibrium. The shape of the sawtooth is chosen as an approximation to Federal minimum wages, where the minimum wage has been raised once every 7 years on average.[29] In the sawtooth equilibrium, minimum wages increase every seven years by 15% and the inflation rate of 2.2% annually is chosen so that the minimum wage does not have a trend.[30]

---

[29]I count phased increases as one increase. Hence, the initial years of the 11 increases in 74 years since minimum wages were implemented in 1938: 1939, 1945, 1950, 1956, 1961, 1967, 1974, 1978, 1990, 1996, 2007.

[30]The model is simulated using the algorithm described in 3.9, where the model is started in the steady state defined by the minimum value of the wage.

Results

Column (3) of Table 3.3 shows that the contemporaneous price response in the sawtooth equilibrium is basically identical to that from even the permanent minimum wage increase. The contemporaneous price response is insensitive to the nature of the variation: it is similar across permanent, temporary and repeated and predictable temporary minimum wage increases.

Column (3) of Table 3.2 displays the cumulative time path of the employment response to a minimum wage increase in the sawtooth equilibrium and makes two points. The first message of the table is that repeated temporary increases are not well-suited to to finding large employment effects of minimum wage increases even in a model that embodies large long-run elasticities. The contemporaneous employment effect is -0.06 and there are no observed long-run effects.

A second message of Table 3.2 is that long-run elasticities explain the employment consequences of a policy of repeated temporary minimum wage increases. On average, employment is about 3% lower than if wages were always at their lowest level in the simulation, and real minimum wages are about 6% above the lowest level. The implied "policy elasticity" of the average employment level with respect to the average minimum wage is -0.58, which is identical to the long-run elasticity of -0.55. Why is this? Firms choose a capital-intensity well-suited to the long-run average level of the minimum wage. Employment fluctuates around this lower steady state level, but these fluctuations are small relative to the lower steady state. These level difference in employment would not be easy to measure in empirical work. In standard empirical approaches, such level differences would be absorbed in the location fixed effects. The location fixed effects contain many other differences across places that are probably not solely due to the average level of minimum wages.

3.6 Quantitative implications of the model with actual minimum wage variation

So far I have studied a dynamic model embodying an important distinction between short- and long-run elasticities. If minimum wage increases are sufficiently temporary, I have shown that there would be little difference in the observed short- and long-run employment responses.

The model emphasizes two channels through which temporariness matters. First, temporariness matters because of how many firms adjust to the realized minimum

135

wage increase: if in realization the minimum wage increase is temporary, then few firms adjust and the long-run effects are small. Second, temporariness matters because of firms' expectations: if firms expect minimum wage increases to be temporary, then the firms that adjust do so by less and the long-run effects are small.

To address whether minimum wages used in empirical work are sufficiently temporary to make inference from standard empirical approaches misleading, I simulate the model to replicate the dataset used by Dube, Lester, and Reich (2010). In the simulation, expectations are pinned down by estimating a stochastic process for minimum wages in-sample.

Replicating a dataset that contains actual minimum wages directly addresses the first channel through which temporariness matters: the simulation uses realized variation, which might contain a mix of temporary and permanent changes. Estimating a stochastic process on the data partially addresses the second channel: in the model, the extent to which firms expect increases to be temporary is tied to the data. Of course, the estimated stochastic process may be misspecified.

Simulating the model using realized minimum wages directly addresses another concern: the use of counties also subject to inflation as the control group renders the main mechanism for temporariness moot. In the simulations, the control counties are also subject to inflation.

3.6.1 Data description

Dube, Lester, and Reich (2010) use county-level variation in the minimum wage in the US from 1990-2006. Their research design generalizes the Card and Krueger (1994) case study approach by pairing bordering counties with different minimum wages to control for local economic shocks. They study employment in the restaurant industry using data from the Quarterly Census of Employment and Wages.

Figure 3.3 shows both the real value of the Federal minimum wage and the 95th percentile of the effective county-level minimum for the sample used by Dube, Lester, and Reich (2010). The 95th percentile of the effective minimum wage is often above the Federal minimum wage because many states set a minimum wage above the Federal minimum.[31]

---

[31]My sample ends up slightly smaller than Dube, Lester and Reich's, with very little effect on point estimates, because some of their estimation sample is sometimes missing the measure of total employment and so they do not have a balanced panel of counties, whereas the text of their paper suggests that this is their goal.

3.6.2 Expectations: the stochastic process of wages

An important input to the simulation is firms' expectations about the path of minimum wages. A minimum wage process with two features attempts to capture the sawtooth pattern. First, a minimum wage increase is more likely the lower is the minimum wage relative to its long-run average. Second, conditional on an increase occurring, there is still some uncertainty about the size.

Formally, the minimum wage evolves according to the following process:

$$
w_{i,t+1}|w_{i,t} = \begin{cases} (1-inf)w_{i,t} \text{ with probability } (1-h(w_{i,t}-\bar{w}_i)) \\ w_{i,t}(1+g) \text{ with probability } h(w_{i,t}-\bar{w}_i), \end{cases}
$$

where I now define notation. The location-time invariant inflation rate is inf. The long-run location-specific average minimum wage is $\bar{w}_i$. The current minimum wage is $w_{i,t}$. The realization of the size of a minimum wage increase is g, which is drawn from a location-time invariant distribution G. Finally, $h : \mathbb{R} \to [0,1]$ is the function that captures the idea that minimum wage increases are more likely the lower is the minimum wage relative to its long-run average. In particular, this is true when $h' < 0$.

I estimate all parameters of this process, except for inf, in-sample on the pooled contiguous border county sample of Dube, Lester, and Reich (2010). The inf parameter is chosen so that the estimated stochastic process does not have a trend. 3.9.2 provides details. Figure 3.4a shows that the estimated quarterly probabilities of a minimum wage hike are consistent with sawtooth dynamics. The probability of an increase becomes very small (below 5%) when the current minimum wage is above the long-run average and rises rapidly when the current minimum wage falls below the long-run average—when the current minimum wage is 30% below the long-run average the per period probability of an increase exceeds 50%. Figure 3.4b shows that the empirical distribution of minimum wage hikes places the most mass between 15% and 25% hikes.

An important assumption in estimating the stochastic process that governs firms' expectations on the pooled sample is that the process—which reflects a particular form of temporariness of minimum wage increases—is common across locations. To explore the validity of this assumption, 3.9.2.1 specifies an auxiliary statistical model of the persistence of minimum wages and shows that the dispersion across locations in the data is similar to that in data simulated from the stochastic process. Using only in-sample information, firms would be unlikely to reject the null that observed minimum wages are consistent with the estimated stochastic process, rather than

there being place specific minimum wage processes. Because I do not horserace the fit of auxiliary statistical models with and without permanent shocks, this exercise does not shed light on whether or not there are permanent shocks in the minimum wage process.

3.6.3   Simulated data

3.9 describes both the simulation algorithm and construction of the minimum wage series in detail. Firms' expectations about the minimum wage are given by the stochastic process described in subsection 3.6.2. The county-level employment series is generated using the *realized* minimum wage histories at the county-level. I use 9 runs through the realized minimum wage histories as a burn-in period and store the 10*th* run. Visual inspection reveals that 9 runs are more than adequate to ensure convergence. The resulting dataset is exactly the same size as the Dube, Lester, and Reich (2010) dataset and includes the measures of employment and population from the Dube, Lester, and Reich (2010) dataset.

3.6.4   Specification

I estimate a distributed lag specification on the actual dataset and the simulated dataset. Following Dube, Lester, and Reich (2010) I use logs and their preferred specification (equation (7)).[32] The unit of observation is county-level restaurant employment (*restemp*) in county i, border-pair p, and at time t. A county enters the dataset if it has a bordering county with a different minimum wage at any point between 1990:I-2006:II. A county can be in the dataset multiple times if it borders multiple counties with different minimum wages. There are two additional data series total employment (*totemp*) and population (*pop*). To control for local economic shocks the regression includes a border-pair time period specific fixed effect, τ_{pt}.[33]

The estimating equation is:

$$\ln(\text{restemp})_{ipt} = a_1 + \sum_{j=0}^{23} \eta_{-j}\Delta\ln(MW_{i,t-j}) + \eta_{-24}\ln(MW_{i,t-24})$$

$$+a_2\ln(\text{totemp})_{it} + a_3\ln(\text{pop})_{it} + \phi_i + \tau_{pt} + \epsilon_{ipt} \tag{3.15}$$

---

[32]Neumark and Wascher (1992) and Baker, Benjamin, and Stanger (1999) use specifications in levels rather than logs. I have estimated the below specification in levels, and converted the level specification to elasticities at the mean of the sample. On the simulated data the specifications in levels yields elasticities slightly smaller in magnitude than the specifications in logs, so in "clean" data this specification choice is not important.

[33]See Dube, Lester, and Reich (2010, pg. 951-2) for details on how this is estimated.

where ϕ_i is a county fixed effect and ϵ_{ipt} is an error term. Equation 3.15 follows Baker, Benjamin, and Stanger (1999) and examines six year (24 quarters) of lagged responses, rather than two years of leads and four years of lags as in Dube, Lester, and Reich (2010). This choice allows comparison to other Baker, Benjamin, and Stanger (1999) specifications discussed in Section 3.6.6.

Equation 3.15 uses differenced minimum wages, which means that η_{-j} is the average employment effect of the jth lag of minimum wages. The η's are identified based on periods when *within* a given county-border pair the two counties have different minimum wages; in periods when the county-border pair have the same minimum wages the fixed effect absorbs the common component of restaurant employment. Following Dube, Lester, and Reich I report Cameron, Gelbach, and Miller (2011) standard errors clustered at the state and border segment separately, and corrected for certain forms of heteroskedasticity.[34]

3.6.5 Employment results

Figure 3.5 shows that because minimum wage increases are temporary there are not large estimated employment effects of minimum wage changes. The figure plots the elasticities estimated on the actual data and the simulated data.[35] The estimates on simulated data are very small negative elasiticities in the first quarters—about -0.05—and in subsequent periods the employment response barely grows, reaching -0.09 after 24 quarters. At all horizons the simulated data lies within the confidence bands of the actual data, though these bands are very wide.

3.6.6 Additional specifications

3.10 reports the results of the two additional long-run specifications used by Baker, Benjamin, and Stanger (1999). These results are consistent with small employment effects of minimum wage changes. Neither of these specifications yield an elasticity after six years that exceeds -0.07 on the simulated data.

3.6.7 Robustness

It might seem like the results should be very sensitive to the rate at which firms adjust, δ. To provide intuition as to why this is not the case, Table 3.4 reproduces

---

[34]There are standard errors on the simulated data because the simulated dataset is the same size as the actual dataset, and so in principle there is sampling variability. In practice, the standard errors from regressions estimated on simulated data are very small.

[35]The coefficients are in Table 3.7.

the calculations of employment elasticities with respect to permanent, temporary and sawtooth minimum wage increases presented in Table 3.2 and discussed in Section 3.5 for many of values of δ. For sawtooth increases, the interaction of the erosion of the minimum wage and the expectation of future increases limits the sensitivity to δ. Indeed, as in the more complex simulations described immediately below, the elasticities with $\delta = 0.08$ and $\delta = 0.16$ are basically the same. On the other hand, the Table shows that the response to permanent increases is very sensitive to δ.[36] Consistent with the main themes of this paper, results for permanent increases provide a misleading guide to increases in a sawtooth environment.

For the robustness exercise, I replicate the Dube, Lester, and Reich (2010) data using $\delta \in \{0.16, 0.33, 0.5\}$. Table 3.5 reports the point estimates from the distributed lag specification for the contemporaneous elasticity and the "long-run" (at 24 quarters) elasticity. The contemporaneous and long-run elasticities are basically identical for 0.08 and 0.16. This indicates that raising δ in this range increases the extent of attenuation, since Table 3.4 shows the employment effect of a permanent increase at 24 quarters is 50% larger with the higher value of δ. With $\delta = 0.33$, the long-run elasticity grows to -0.16. Similar to $\delta = 0.08$, this estimate is about one-third of the long-run elasticity implied by the model, because the response to a permanent increase grows with δ.[37] With $\delta = 0.5$, the contemporaneous elasticity is -0.08 and the long-run elasticity is -0.24. Again, this exhibits substantial attenuation relative to a permanent minimum wage increase (in Table 3.4 it is -0.54).

3.7   Conclusion

If there were differences between short- and long-run employment elasticities would it be possible to tell? This paper has suggested that because of the nature of variation in minimum wages the answer is no. In particular, because minimum wage increases are mostly temporary, the *ceteris paribus* condition implied in the long-run elasticities—that of a permanent minimum wage increase—is unlikely to obtain. Even without such a *ceteris paribus* condition, it might be the case that standard regressions could overcome this challenge. In the presence of a distinction between the short- and long-run elasticities, standard regressions in the literature do not overcome this challenge and so have potentially been misinterpreted.

---

[36]For the smaller values of δ, the main departure from linearity is because the contemporaneous response is dominated by the scale effect—or market quantity change—which is the same regardless of the value of δ. Subtracting off the year 0 effect gets much closer to a linear relationship.

[37]See Table 3.4: $0.08/0.252 \approx 0.16/0.505 \approx 1/3$.

As such, the paper suggests that it would be a mistake to infer from existing empirical work on the employment effects of minimum wage increases that the President's 2013 proposal to index minimum wages to inflation would have minimal effects on employment. Taking the model at face value shows how misleading such an inference might be: the results in Table 3.2 show that a contemporaneous elasticity of -0.002 in response to a temporary increase is consistent with an elasticity after 6 years of -0.252 for a permanent increase.

The putty-clay model also offers a rationalization of two divergent strands of empirical work on effects of minimum wage increases. On the one hand, the evidence of small short-run employment effects is inconsistent with standard static models of labor demand; on the other hand, product price increases are inconsistent with models of "supply side constraints" (e.g. search models). The putty-clay model is consistent with both. In the model, observed long-run effects of temporary minimum wage increases are also small, though the long-run effect on employment of a permanent change in the real value of the minimum wage is large and equilibrium employment is below what it would have been were it not for the minimum wage.

This paper suggests several avenues for future research. The putty-clay model could be extended along a couple dimensions. First, the labor market is assumed to be perfectly competitive, which omits potentially important mechanisms that models of search emphasize. Second, the firm's problem is simplified by using an exogenous adjustment probability, rather than an endogenous adjustment process. More broadly, it would be interesting to bring detailed evidence to bear on several mechanisms that are emphasized by the putty-clay model: the speed with which employers of low-skill labor adjust labor demand in response to shocks, how firms form expectations about minimum wages, and how these expectations affect decisionmaking.

Table 3.1: Annual parameter values

Parameter	Value	Description
	0.3	Land, structures and machines
	0.4	Materials share
	0.2	High skill labor share
α	0.9	Non-minimum wage-labor share
δ	0.08	Expiration chance
γ	0.6	Price elasticity of demand
β	0.95	Discount rate
n_h	0.5	$\frac{\text{high-skill labor}}{\text{low-skill labor}}$ in steady state

Source: See the discussion in Section 3.4.

Table 3.2: Employment response to minimum wage increases

Year	Permanent (1)	Temporary (2)	Sawtooth (3)
0	−0.062	−0.059	−0.056
1	−0.100	−0.063	−0.049
2	−0.136	−0.061	−0.040
3	−0.169	−0.057	−0.030
4	−0.199	−0.049	−0.020
5	−0.227	−0.040	−0.010
6	−0.252	−0.029	−0.002
∞	−0.546	0	NA
			Policy Elasticity
			−0.584

Note: An elasticity is defined as the percent change in employment divided by the percent change in wages. The permanent columns shows the total employment elasticity after X years from a one-time and permanent 15% minimum wage increase. The temporary column reports the total employment elasticity after X years from a one-time 15% minimum wage increase that is eroded at 2.2% per year. The sawtooth equilibrium is minimum wages increasing by 15% every seven years and then inflating away at 2.2% per year. The sawtooth column shows the total employment elasticity after X years following the minimum wage increase in the sawtooth equilibrium. The policy elasticity is defined using the average level of employment in the sawtooth equilibrium relative to never having a minimum wage and the wage is the average wage in the sawtooth equilibrium relative to having wages always at the lowest level.

Table 3.3: Price response to minimum wage increases

Year	Permanent (1)	Temporary (2)	Sawtooth (3)
0	+0.094	+0.096	+0.093
1	+0.094	+0.081	+0.078
2	+0.094	+0.066	+0.063
3	+0.094	+0.051	+0.049
4	+0.094	+0.036	+0.034
5	+0.094	+0.021	+0.020
6	+0.094	+0.006	+0.007
∞	+0.094	0	NA
			Policy Elasticity
			+0.097

Note: An elasticity is defined as the percent change in price divided by the percent change in wages. The permanent columns shows the total price elasticity after X years from a one-time and permanent 15% minimum wage increase. The temporary column reports the total employment elasticity after X years from a one-time 15% minimum wage increase that is eroded at 2.2% per year. The sawtooth equilibrium is minimum wages increasing by 15% every seven years and then inflating away at 2.2% per year. The sawtooth column shows the total price elasticity after X years following the minimum wage increase in the sawtooth equilibrium. The policy elasticity is defined using the average level of prices in the sawtooth equilibrium relative to never having a minimum wage and the wage is the average wage in the sawtooth equilibrium relative to having wages always at the lowest level.

Table 3.4: Employment response to minimum wage increases: varying δ

Year	$\delta = 0.08$	$\delta = 0.16$	$\delta = 0.33$	$\delta = 0.5$
	A. Permanent increase			
0	-0.062	-0.073	-0.099	-0.130
1	-0.100	-0.148	-0.246	-0.338
2	-0.136	-0.212	-0.345	-0.442
3	-0.169	-0.265	-0.411	-0.494
4	-0.199	-0.310	-0.456	-0.520
5	-0.227	-0.348	-0.485	-0.533
6	-0.252	-0.379	-0.505	-0.539
∞	-0.546	-0.546	-0.546	-0.546
	B. Temporary increase			
0	-0.059	-0.065	-0.087	-0.117
1	-0.063	-0.090	-0.168	-0.256
2	-0.061	-0.100	-0.197	-0.284
3	-0.057	-0.099	-0.191	-0.257
4	-0.049	-0.089	-0.164	-0.203
5	-0.040	-0.075	-0.125	-0.139
6	-0.029	-0.057	-0.084	-0.077
	C. Sawtooth increase			
0	-0.056	-0.057	-0.066	-0.081
1	-0.049	-0.055	-0.082	-0.125
2	-0.040	-0.046	-0.071	-0.107
3	-0.030	-0.033	-0.044	-0.061
4	-0.020	-0.019	-0.013	-0.007
5	-0.010	-0.006	+0.012	+0.037
6	-0.002	+0.002	+0.018	+0.044
Policy elast.	-0.584	-0.584	-0.583	-0.580

Note: see footnotes to table 3.2.

Table 3.5: Distributed lag specification: varying δ

δ	Contemporaneous	Long-run
0.08	-0.051	-0.094
0.16	-0.050	-0.089
0.33	-0.062	-0.159
0.50	-0.084	-0.238

Note: This table reports the contemporaneous (quarter 0) and long-run elasticities (24 quarters) of using the preferred Dube, Lester, and Reich (2010) specification to estimate the employment effects of minimum wages. I estimate equation (3.15) having simulated the model as described in section 3.6.3.

Figure 3.1: Federal minimum wage relative to average hourly earnings in the private sector
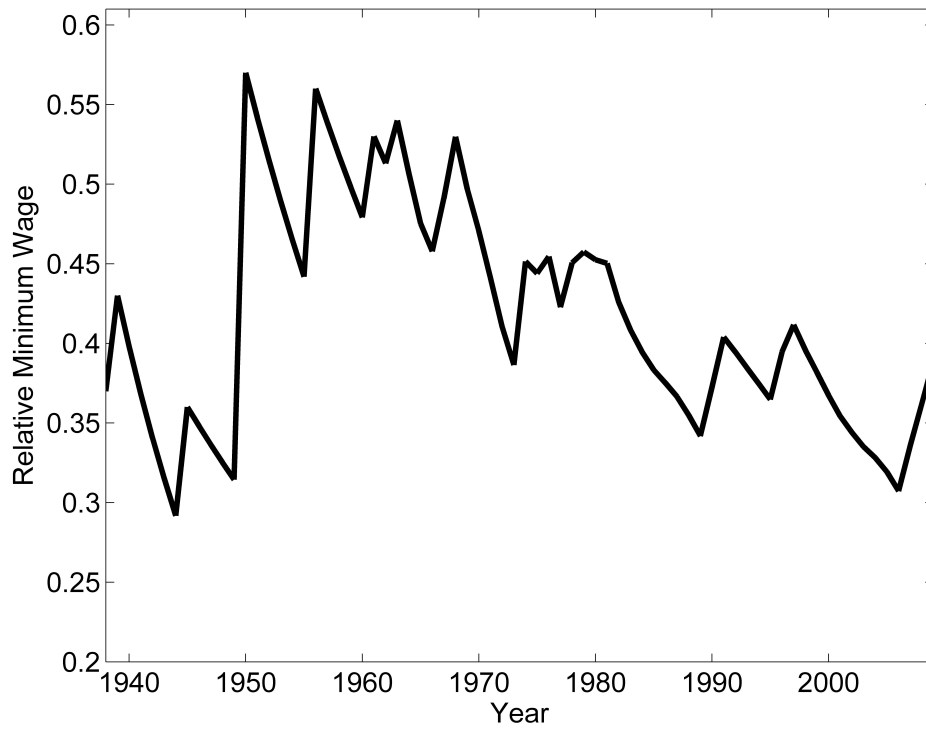
Source: Brown (1999) table 1 updated with the Bureau of Labor Statistics for the nominal federal minimum wage (http://www.dol.gov/whd/minwage/chart.htm), and then deflated using the Current Employment Statistics for average hourly wage in the private sector among production and non-supervisory employees (historical hours and earnings, table B-2). This figure uses annual data and plots the largest minimum wage in each year.

Figure 3.2: Impulse responses to temporary and permanent minimum wage increases
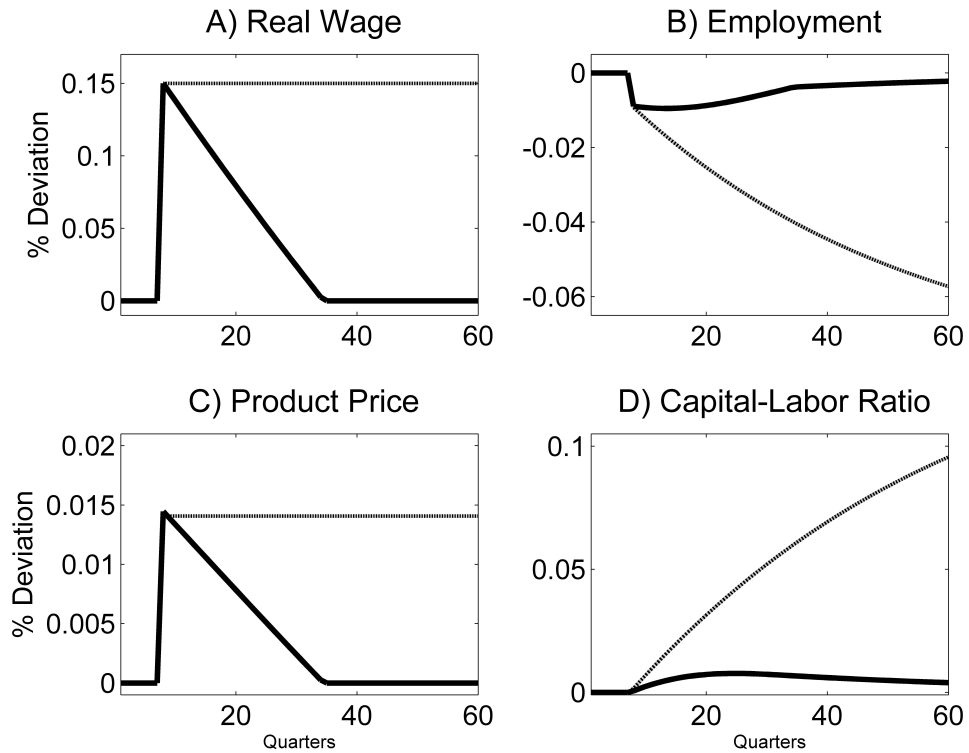
Note: based on simulations of the putty-clay model using the parameterization in Table 3.1. The thick line shows the impulse response to a temporary increase: one-time unanticipated increase of 15% that is eroded by annual inflation of 2.2% a year. The thin line shows the impulse-response to a permanent increase of 15%.

Figure 3.3: County-level variation in minimum wages in the US: 1984-2006
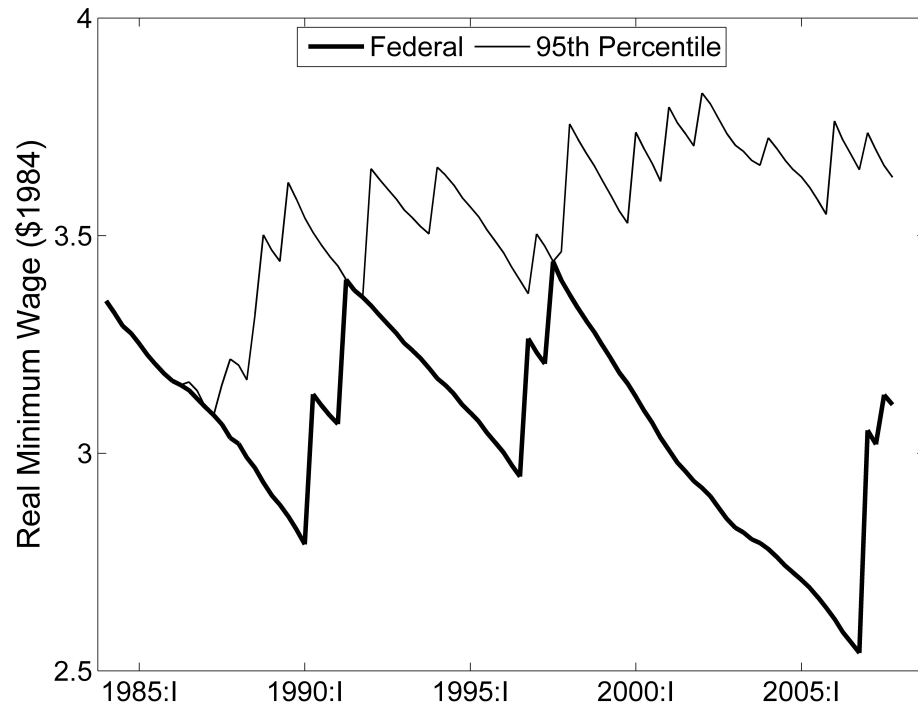
Note: Units are minimum wages deflated by average private sector hourly wage to 1984:I. The figure plots the minimum wages in the contiguous border county sample in Dube, Lester, and Reich (2010).

Figure 3.4: The stochastic process for minimum wages

(a) Estimated Probability of a Minimum Wage Hike   (b) Cumulative Distribution of Minimum Wage Hikes
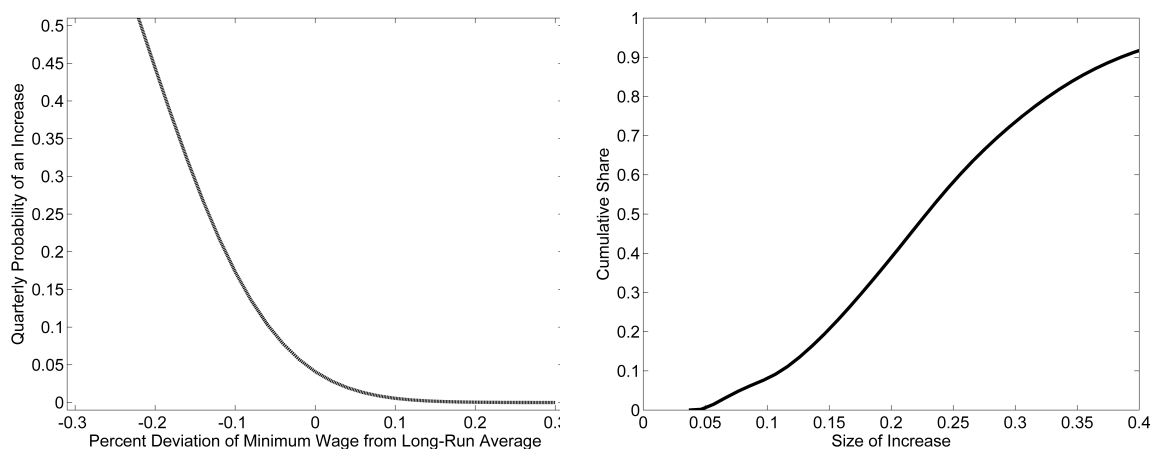
Note: The left panel plots implied probabilities of minimum wage hikes at different values relative to the long-run average of the minimum wage estimated on the contiguous border county sample (dropping duplicate counties) from 1984:I-2007:IV. The regression coefficients are in Table 3.6. The right panel plots the distribution of minimum wage increases in the contiguous border county sample (dropping duplicate counties) from 1984:I-2007:IV.

Figure 3.5: Distributed lag specification
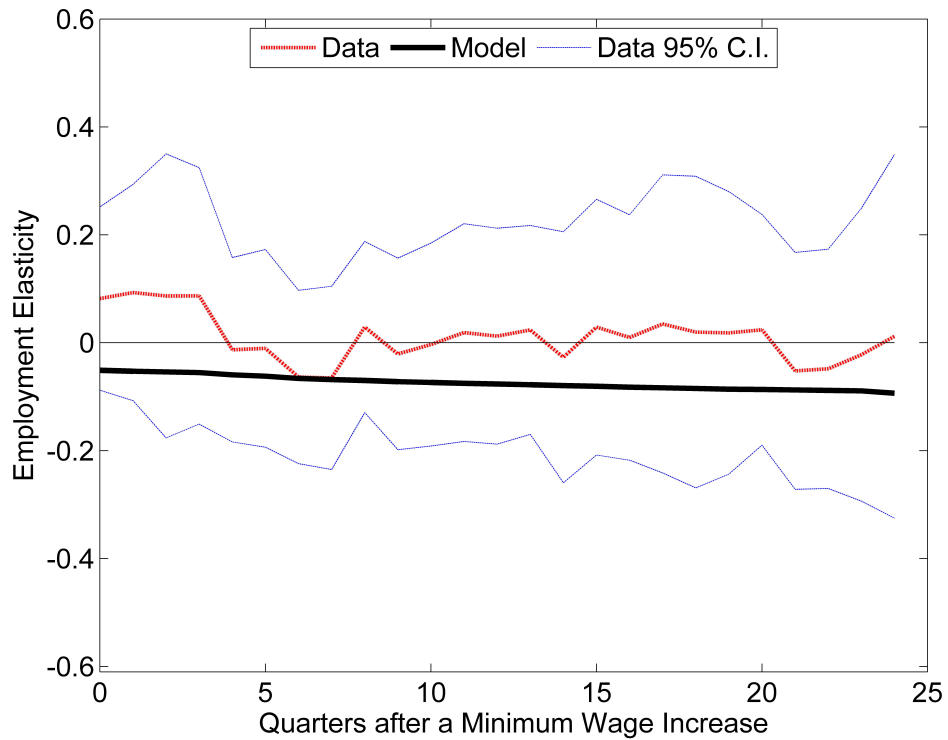
Note: This figure plots the coefficients from the distributed lag specification. The regressions also include a border pair times period interaction term. The outcome is total employment in the restaurant industry. The regressions use the contiguous border sample. The thin line shows the pointwise 95% confidence interval from the data, where the standard errors are clustered at the state and border segment separately. The data plot shows estimates based on actual data. The model line plots coefficients based on data simulated from the putty-clay model. The coefficients are in table 3.7.

3.8 Appendix: Omitted calculations

This appendix derives the steady state relationships in the model, demonstrates that firms never want to shutdown, that there is entry in every period, details the insulation effect, and provides complete proofs of all results stated in the text.

3.8.1 Steady state values

Given a constant wage,

$$q_w = \sum_{j=0}^{\infty} \beta^j (1-\delta)^j w = \frac{w}{1-\beta(1-\delta)}. \tag{3.16}$$

From (3.8),

$$k = q_w \frac{\alpha}{1-\alpha}. \tag{3.17}$$

Combining the optimal size of a machine, equation (3.17), with the zero profit condition, equation (3.3) set to zero in steady state, gives

$$q = \alpha^{-\alpha} \left(\frac{q_w}{1-\alpha} \right)^{1-\alpha}. \tag{3.18}$$

From (3.9) with a constant q,

$$P = (1-\beta(1-\delta)) \frac{q_w^{1-\alpha}}{\alpha^\alpha (1-\alpha)^{1-\alpha}}. \tag{3.19}$$

From (3.5),

$$Q = \theta P^{-\gamma} = \theta \left[\frac{\alpha^{\gamma\alpha}(1-\alpha)^{\gamma(1-\alpha)}}{(1-\beta(1-\delta))^\gamma q_w^{\gamma(1-\alpha)}} \right]. \tag{3.20}$$

The number of workers in steady state is

$$N = \frac{Q}{k^\alpha} = \theta \left[\frac{\alpha^{\alpha(\gamma-1)}(1-\alpha)^{\gamma(1-\alpha)+\alpha}}{(1-\beta(1-\delta))^\gamma q_w^{\gamma(1-\alpha)+\alpha}} \right]. \tag{3.21}$$

3.8.2 No shutdown, entry in every period and the insulation effect

No shutdown

A machine built at time t has capital $k_t = \frac{\alpha}{1-\alpha} q_{w,t}$. Assume that after time $t' > t$ the wage will be constant and equal to $w_{t'}$. Considering permanent movements allow

152

me to derive results in closed-form. The price in period t' is $P_{t'} = (1 - \beta(1 - \delta))\frac{q_{w,t'}^{1-\alpha}}{\alpha^\alpha(1-\alpha)^{1-\alpha}}$. If in period $t' > t$ a firm wants to shut down a machine that was installed in period t this implies that $P_{t'}k_t^\alpha - w_{t'} < 0$ or using the expressions for k_t and $P_{t'}\frac{q_{w,t}}{q_{w,t'}} < (1-\alpha)^{\frac{1}{\alpha}}$. Then shutdown requires (using $\alpha = 0.9$) $q_{w,t} < q_{w,t'}(1-\alpha)^{\frac{1}{\alpha}} = q_{w,t'}0.077$. Shutdown requires that the present discounted value of wages to rise by more than a factor of 10 from when the investment was made. In the dataset the real value of wages varies by less than a factor of two, and the present discounted value of wages is less volatile than the wage. So firms never want to shutdown.

Entry in every period

Entry in every period requires: $h_t > 0 \Leftrightarrow Q_t - (1 - \delta)(Q_{t-1}) > 0 \Leftrightarrow P_t^{-\gamma} - (1 - \delta)P_{t-1}^{-\gamma} > 0 \Leftrightarrow \left(\frac{1}{1-\delta}\right)^{\frac{1}{\gamma}} > \frac{P_t}{P_{t-1}}$. Using the baseline calibration, having no entry in a period requires one quarter increases in the present discounted value of wages of more than 40%, while in the simulation the largest one quarter increase in the minimum wage is 34% (and the present discounted value of wages moves by even less). So there is always entry.

Insulation effect

This section demonstrates the insulation effect. The comparative static is an incumbent firm's revenue from steady state with respect to the wage, where the wage and the price is allowed to respond, but the capital stock is not.

$$R = Pk^\alpha - w \tag{3.22}$$

$$\frac{\partial R}{\partial w}\frac{w}{P} = \frac{\partial P}{\partial w}\frac{w}{P}k^\alpha - \frac{w}{P}. \tag{3.23}$$

Combining equations (3.16) and (3.19), in steady state

$$\frac{w}{P} = q_w^\alpha \alpha^\alpha (1-\alpha)^{1-\alpha}. \tag{3.24}$$

Result 6 has as a special case with perfect foresight from steady state that

$$\frac{\partial P}{\partial w}\frac{w}{P} = 1 - \alpha. \tag{3.25}$$

153

Combining the previous expressions with equation (3.17) gives

$$\frac{\partial R}{\partial w}\frac{w}{P} = (1-\alpha)^{1-\alpha}\alpha^\alpha q_w^\alpha - (1-\alpha)^{1-\alpha}\alpha^\alpha q_w^\alpha \tag{3.26}$$

$$= 0. \tag{3.27}$$

3.8.3 Proofs

First write labor demand in terms of model parameters and the $q_{w,t}$.

Writing labor demand in terms of model parameters.

Iterating equation (3.6) gives

$$N_t = \sum_{j=0}^{\infty}(1-\delta)^j h_{t-j}. \tag{3.28}$$

The goal is to write equation (3.28) in terms of $q_{w,t-j}$. Rearrange equation (3.7):

$$h_t = \frac{Q_t - (1-\delta)Q_{t-1}}{k_t^\alpha}. \tag{3.29}$$

Combine equations (3.3) set to zero, (3.5), and (3.9):

$$Q_t = \theta P_t^{-\gamma} \tag{3.30}$$

$$= \theta[q_t - (\beta(1-\delta))\,E_t[q_{t+1}]]^{-\gamma} \tag{3.31}$$

$$= \theta\left(\frac{\alpha^{-\alpha}}{(1-\alpha)^{1-\alpha}}\right)^{-\gamma}\left[q_{w,t}^{1-\alpha} - (\beta(1-\delta))\,E_t[q_{w,t+1}^{1-\alpha}]\right]^{-\gamma}. \tag{3.32}$$

Combine equations (3.8), (3.29) and (3.32):

$$h_t = \frac{\theta\left\{\left[q_{w,t}^{1-\alpha} - (\beta(1-\delta))\,E_t[q_{w,t+1}^{1-\alpha}]\right]^{-\gamma} - (1-\delta)\left[q_{w,t-1}^{1-\alpha} - (\beta(1-\delta))\,E_{t-1}[q_{w,t}^{1-\alpha}]\right]^{-\gamma}\right\}}{\frac{\alpha^{1-\alpha\gamma}}{(1-\alpha)^{1+\gamma(1-\alpha)}}q_{w,t}}. \tag{3.33}$$

Substitute suitably modified versions of equation (3.33) into (3.28):

$$N_t$$

$$= \sum_{j=0}^{\infty} (1 - \delta)^j \left(\frac{\alpha^{\alpha - \alpha\gamma}}{(1 - \alpha)^{\alpha - \gamma(1-\alpha)}} q_{w,t-j}^{\alpha} \right)^{-1} \tag{3.34}$$

$$\times \theta \left\{ \left[q_{w,t-j}^{1-\alpha} - (\beta(1 - \delta)) E_{t-j}[q_{w,t+1-j}^{1-\alpha}] \right]^{-\gamma} \right. \tag{3.35}$$

$$\left. - (1 - \delta) \left[q_{w,t-1-j}^{1-\alpha} - (\beta(1 - \delta)) E_{t-1-j}[q_{w,t-j}^{1-\alpha}] \right]^{-\gamma} \right\}. \tag{3.36}$$

Proof of result 4.

Proof. Equation (3.6) gives

$$N_t = h_t + (1 - \delta)N_{t-1}. \tag{3.37}$$

Using equation (3.37) and converting to an elasticity at the end,

$$\frac{\partial N_t}{\partial w_t} = \frac{\partial h_t}{\partial w_t} \tag{3.38}$$

$$= \frac{\partial \left\{ \frac{Q_t - (1-\delta)Q_{t-1}}{k_t^{\alpha}} \right\}}{\partial w_t} \tag{3.39}$$

$$= \frac{\partial Q_t}{\partial w_t} \frac{1}{k_t^{\alpha}} - \alpha \frac{\partial k_t}{\partial w_t} \frac{1}{k_t} \frac{Q_t - (1 - \delta)Q_{t-1}}{k_t^{\alpha}} \tag{3.40}$$

$$= \frac{\partial Q_t}{\partial w_t} \frac{1}{k_t^{\alpha}} - \alpha \frac{\partial k_t}{\partial w_t} \frac{h_t}{k_t} \tag{3.41}$$

$$\frac{\partial N_t}{\partial w_t} \frac{w_t}{N_t} = \frac{\partial Q_t}{\partial w_t} \frac{1}{k_t^{\alpha}} \frac{w_t}{N_t} - \alpha \frac{\partial k_t}{\partial w_t} \frac{w_t}{k_t} \frac{h_t}{N_t}. \tag{3.42}$$

Solve for the RHS of (3.42) in pieces. I am interested in an increase of w_t from steady state.

First, collect some useful facts about steady state. Start with equation (3.2)

$$q_{w,t} = w_t + \beta(1 - \delta)\mathbb{E}_t[q_{w,t+1}] \tag{3.43}$$

$$\frac{\partial q_{w,t}}{\partial w_t} = 1 \tag{3.44}$$

$$\frac{\partial q_{w,t}}{\partial w_t} \frac{w_t}{q_{w,t}} = \frac{w_t}{q_{w,t}}. \tag{3.45}$$

155

In steady state (equation (3.16)):

$$\frac{w}{q_w} = 1 - \beta(1 - \delta). \tag{3.46}$$

Rewrite equation (3.28) in steady state.

$$N = \sum_{j=0}^{\infty} (1-\delta)^j h = \frac{h}{\delta} \Leftrightarrow \delta = \frac{h}{N}. \tag{3.47}$$

Rewrite equation (3.29) in steady state.

$$h = \frac{Q - (1-\delta)Q}{k^\alpha} = \frac{\delta Q}{k^\alpha}. \tag{3.48}$$

Combine (3.47) and (3.48):

$$\delta N = \frac{\delta Q}{k^\alpha} \Leftrightarrow Q = N k^\alpha. \tag{3.49}$$

This completes the preliminary facts.

Now solve for terms on the RHS of equation (3.42). Substitute equation (3.49) into the first term on the RHS of equation (3.42) and expand:

$$\frac{\partial Q_t}{\partial w_t} \frac{w_t}{Q_t} = \frac{\partial Q_t}{\partial P_t} \frac{\partial P_t}{\partial w_t} \frac{w_t}{Q_t} \frac{P_t}{P_t} \tag{3.50}$$

$$= \frac{\partial Q_t}{\partial P_t} \frac{P_t}{Q_t} \frac{\partial w_t}{\partial P_t} \frac{P_t}{w_t} \tag{3.51}$$

$$= -\gamma(1 - \alpha), \tag{3.52}$$

where the last step computes an elasticity from equation (3.5) and uses the special case of Result 6 from steady state.

Now consider part of the second term on the RHS of equation (3.42) combined with equation (3.47).

$$\frac{\partial k_t}{\partial w_t} \frac{w_t}{k_t} = \frac{\partial k_t}{\partial q_{w,t}} \frac{\partial q_{w,t}}{\partial w_t} \frac{w_t}{k_t} \frac{q_{w,t}}{q_{w,t}} \tag{3.53}$$

$$= \frac{\partial k_t}{\partial q_{w,t}} \frac{q_{w,t}}{k_t} \frac{\partial q_{w,t}}{\partial w_t} \frac{w_t}{q_{w,t}} \tag{3.54}$$

$$= \frac{w_t}{q_{w,t}} \tag{3.55}$$

$$= 1 - \beta(1 - \delta), \tag{3.56}$$

156

where the elasticities are computed from (3.8) and (3.45) and equation (3.56) uses (3.46).

Finally, substituting equation (3.56), (3.52), and (3.47) into the RHS of (3.42) in steady state gives the result:

$$\frac{\partial N_t}{\partial w_t}\frac{w_t}{N_t} = \frac{\partial Q_t}{\partial w_t}\frac{1}{k_t^\alpha}\frac{w_t}{N_t} - \alpha\frac{\partial k_t}{\partial w_t}\frac{w_t}{k_t}\frac{h_t}{N_t} \tag{3.57}$$

$$= -\gamma(1 - \alpha) - \alpha\delta(1 - \beta(1 - \delta)). \tag{3.58}$$

$\square$

Proof of result 5.

Proof. I first compute the contemporaneous elasticity. I then compute the lagged effects. Throughout, I make use of steady state facts derived in result 4.

Equation (3.6) gives

$$N_t = h_t + (1 - \delta)N_{t-1}. \tag{3.59}$$

I am interested in the effect of a change in $q_{w,t}$ on N_t (contemporaneous effect of a permanent minimum wage increase on employment). Taking the relevant derivative of (3.59)

$$\frac{\partial N_t}{\partial q_{w,t}} = \frac{\partial h_t}{\partial q_{w,t}}, \tag{3.60}$$

since N_{t-1} is predetermined. Substitute in (3.29):

$$\frac{\partial N_t}{\partial q_{w,t}} = \frac{\partial\left[\frac{Q_t - (1-\delta)Q_{t-1}}{k_t^\alpha}\right]}{\partial q_{w,t}} \tag{3.61}$$

$$= \frac{\partial Q_t}{\partial q_{w,t}}\frac{1}{k_t^\alpha} - \alpha\frac{\partial k_t}{\partial q_{w,t}}\frac{h_t}{k_t}. \tag{3.62}$$

Convert to an elasticity:

$$\frac{\partial N_t}{\partial q_{w,t}}\frac{q_{w,t}}{N_t} = \frac{\partial Q_t}{\partial q_{w,t}}\frac{1}{k_t^\alpha}\frac{q_{w,t}}{N_t} - \alpha\frac{\partial k_t}{\partial q_{w,t}}\frac{h_t}{k_t}\frac{q_{w,t}}{N_t} \tag{3.63}$$

$$= \frac{\partial Q_t}{\partial q_{w,t}}\frac{q_{w,t}}{Q_t} - \alpha\delta\frac{\partial k_t}{\partial q_{w,t}}\frac{q_{w,t}}{k_t}, \tag{3.64}$$

157

where this last equation makes use of (3.49) and (3.47). From equation (3.8),

$$\frac{\partial k_t}{\partial q_{w,t}} \frac{q_{w,t}}{k_t} = 1. \tag{3.65}$$

The last term to solve for is $\frac{\partial Q_t}{\partial q_{w,t}} \frac{q_{w,t}}{Q_t}$. The assumption of permanence—and expectations thereof—is that $E_t[q_{w,t+1}^{1-\alpha}] = q_{w,t}^{1-\alpha}$ and $\frac{\partial E_t[q_{w,t+1}^{1-\alpha}]}{\partial q_{w,t}} = (1-\alpha)q_{w,t}^{-\alpha}$. Using equation (3.32) and then substituting in for the expectations:

$$\frac{\partial Q_t}{\partial q_{w,t}} \frac{q_{w,t}}{Q_t} = -\gamma \frac{(1-\alpha)q_{w,t}^{1-\alpha} - \beta(1-\delta)\frac{\partial E_t[q_{w,t+1}^{1-\alpha}]}{\partial q_{w,t}}q_{w,t}}{\left[q_{w,t}^{1-\alpha} - (\beta(1-\delta))E_t[q_{w,t+1}^{1-\alpha}]\right]} \tag{3.66}$$

$$= -\gamma \frac{(1-\alpha)q_{w,t}^{1-\alpha} - \beta(1-\delta)(1-\alpha)q_{w,t}^{1-\alpha}}{\left[q_{w,t}^{1-\alpha} - (\beta(1-\delta))q_{w,t}^{1-\alpha}\right]} \tag{3.67}$$

$$= -\gamma(1-\alpha). \tag{3.68}$$

Substituting (3.65) and (3.68) into (3.64) gives the result for the contemporaneous elasticity:

$$\frac{\partial N_t}{\partial q_{w,t}} \frac{q_{w,t}}{N_t} = -\gamma(1-\alpha) - \alpha\delta. \tag{3.69}$$

Now consider the lagged response. I am interested in the effect of a change in $q_{w,t}$ on N_{t+n} (the lagged effects of a permanent minimum wage increase on employment). From equation (3.34), $q_{w,t}$ appears in two terms making up N_{t+n} (since we want partial and not total derivatives). Using the representation in equation (3.28) gives

$$N_{t+n} = h_{t+n} + ... + (1-\delta)^{n-1}h_{t+1} + (1-\delta)^n h_t + ... \tag{3.70}$$

Hence the direct effect of a change in $q_{w,t}$ on N_{t+n} is (substituting in (3.29)):

$$\frac{\partial N_{t+n}}{\partial q_{w,t}} = (1-\delta)^{n-1}\frac{\partial h_{t+1}}{\partial q_{w,t}} + (1-\delta)^n \frac{\partial h_t}{\partial q_{w,t}} \tag{3.71}$$

$$= (1-\delta)^{n-1}\frac{\partial \left[\frac{Q_{t+1}-(1-\delta)Q_t}{k_{t+1}^\alpha}\right]}{\partial q_{w,t}} + (1-\delta)^n \frac{\partial \left[\frac{Q_t-(1-\delta)Q_{t-1}}{k_t^\alpha}\right]}{\partial q_{w,t}} \tag{3.72}$$

$$= -(1-\delta)^n \frac{1}{k_{t+1}^\alpha}\frac{\partial Q_t}{\partial q_{w,t}} + (1-\delta)^n \left[\frac{\partial Q_t}{\partial q_{w,t}}\frac{1}{k_t^\alpha} - \alpha \frac{\partial k_t}{\partial q_{w,t}}\frac{h_t}{k_t}\right]. \tag{3.73}$$

Since the increase is permanent, $q_{w,t} = q_{w,t+1}$ so that (using equation (3.8)) $k_{t+1}^\alpha = k_t^\alpha$. Simplifying, converting to elasticities, applying equation (3.65), using the fact that

158

the change is local so that $N_t = N_{t+n}$, and finally that the economy starts in steady state so that equation (3.47) holds gives the result:

$$\frac{\partial N_{t+n}}{\partial q_{w,t}} = -\alpha(1-\delta)^n \frac{\partial k_t}{\partial q_{w,t}} \frac{h_t}{k_t} \tag{3.74}$$

$$\frac{\partial N_{t+n}}{\partial q_{w,t}} \frac{q_{w,t}}{N_{t+n}} = -\alpha(1-\delta)^n \frac{\partial k_t}{\partial q_{w,t}} \frac{h_t}{k_t} \frac{q_{w,t}}{N_{t+n}} \tag{3.75}$$

$$= -\alpha(1-\delta)^n \frac{h_t}{N_{t+n}} \tag{3.76}$$

$$= -\alpha\delta(1-\delta)^n. \tag{3.77}$$

$\square$

Proof of result 6.

Proof. Start with equation (3.9) and substitute in the combination of equation (3.8) and the zero profit condition, equation (3.3), set to zero. Then expand $q_{w,t}$ using equation (3.2):

$$P_t = q_t - \beta(1-\delta)E_t[q_{t+1}] \tag{3.78}$$

$$= \alpha^{-\alpha} \left[\frac{q_{w,t}}{1-\alpha}\right]^{1-\alpha} - \beta(1-\delta)\alpha^{-\alpha}E_t\left\{\left[\frac{q_{w,t+1}}{1-\alpha}\right]^{1-\alpha}\right\} \tag{3.79}$$

$$= \alpha^{-\alpha} \left[\frac{w_t + \beta(1-\delta)E_t[q_{w,t+1}]}{1-\alpha}\right]^{1-\alpha} - \beta(1-\delta)\alpha^{-\alpha}E_t\left\{\left[\frac{q_{w,t+1}}{1-\alpha}\right]^{1-\alpha}\right\}. \tag{3.80}$$

Imposing perfect foresight provides a convenient alternative representation of P_t:

$$P_t = \alpha^{-\alpha} \left[\frac{q_{w,t}}{1-\alpha}\right]^{1-\alpha} - \beta(1-\delta)\alpha^{-\alpha}\left\{\left[\frac{q_{w,t+1}}{1-\alpha}\right]^{1-\alpha}\right\} \tag{3.81}$$

$$= \alpha^{-\alpha} \left[\frac{q_{w,t}}{1-\alpha}\right]^{1-\alpha} \left[1 - \beta(1-\delta)\left\{\left[\frac{q_{w,t+1}}{q_{w,t}}\right]^{1-\alpha}\right\}\right]. \tag{3.82}$$

Take a derivative of equation (3.80) with respect to w_t, convert to elasticity form

and substitute in for P_t while imposing perfect foresight (equation (3.82)):

$$\frac{\partial P_t}{\partial w_t} = \alpha^{-\alpha} \left[\frac{w_t + \beta(1-\delta)E_t[q_{w,t+1}]}{1-\alpha} \right]^{-\alpha} \tag{3.83}$$

$$\frac{\partial P_t}{\partial w_t} \frac{w_t}{P_t} = \alpha^{-\alpha} \left[\frac{w_t + \beta(1-\delta)E_t[q_{w,t+1}]}{1-\alpha} \right]^{-\alpha} \frac{w_t}{P_t} \tag{3.84}$$

$$= \alpha^{-\alpha} \left[\frac{q_{w,t}}{1-\alpha} \right]^{-\alpha} \frac{w_t}{\alpha^{-\alpha} \left[\frac{q_{w,t}}{1-\alpha} \right]^{1-\alpha} \left[1 - \beta(1-\delta) \left(\frac{q_{w,t+1}}{q_{w,t}} \right)^{1-\alpha} \right]} \tag{3.85}$$

$$= (1-\alpha) \frac{1}{q_{w,t}} \frac{w_t}{\left[1 - \beta(1-\delta) \left(\frac{q_{w,t+1}}{q_{w,t}} \right)^{1-\alpha} \right]}. \tag{3.86}$$

Let $\tilde{w}_t = (1 - \beta(1-\delta))q_{w,t}$ and $\tilde{w}_{t+1} = (1 - \beta(1-\delta))q_{w,t+1}$. Then:

$$\frac{\partial P_t}{\partial w_t} \frac{w_t}{P_t} = (1-\alpha) \frac{w_t}{\tilde{w}_t} \frac{1 - \beta(1-\delta)}{1 - \beta(1-\delta) \left(\frac{\tilde{w}_{t+1}}{\tilde{w}_t} \right)^{1-\alpha}}. \tag{3.87}$$

For local changes from steady state, $w_t = \tilde{w}_t$ and $\tilde{w}_{t+1} = \tilde{w}_t$ so that

$$\frac{\partial P}{\partial w_t} \frac{w_t}{P_t} = (1-\alpha). \tag{3.88}$$

$\square$

3.9   Appendix: Simulation details

3.9.1   Simulation algorithm for perfect foresight

This appendix describes how to simulate the model in section 3.5. The model is solved under perfect foresight. There are two steps: first, initializing the model, and second, solving the dynamics given a varying minimum wage.

Initializing the model to steady state

Suppose that the economy is in the steady state defined by w. Then the steady state values are computed in the following order using the formulas in 3.8.1:

$$
\begin{aligned}
q_w &= \frac{w}{(1 - \beta(1 - \delta))}, \\
k &= q_{w,t}\frac{\alpha}{1 - \alpha}, \\
q &= \alpha^{(-\alpha)}\left(\frac{q_w}{1 - \alpha}\right)^{(1-\alpha)}, \\
P &= q_t(1 - \beta(1 - \delta)), \\
Q &= \theta P^{-\gamma}, \\
N &= \frac{Q}{k^\alpha}, \\
K &= Nk,
\end{aligned}
$$

where K_t is the aggregate capital stock. The employment numbers include the high-skill workers:

$$
E = (1 + n_h)N.
$$

Dynamics to changing minimum wages

Suppose that after time t the wage will vary. At the beginning of time $t + 1$ firms know $\{w_{t+1+j}\}_{j=0}^{\infty}$. Define

$$
q_{w,t+1+j} = \sum_{i=0}^{\infty}(\beta(1 - \delta))^i w_{t+1+j+i} \forall j \geq 0.
$$

Given the sequence of $q_{w,t+j}$, the following loop shows how to simulate the model starting from the time t steady state.

For $j = 1, 2, 3 \ldots$:

$$q_{t+j} = \alpha^{(-\alpha)} \left(\frac{q_{w,t+j}}{1 - \alpha} \right)^{(1-\alpha)}$$

$$q_{t+j+1} = \alpha^{(-\alpha)} \left(\frac{q_{w,t+j+1}}{1 - \alpha} \right)^{(1-\alpha)}$$

$$P_{t+j} = q_{t+j} - \beta(1 - \delta)q_{t+j+1}$$

$$Q_{t+j} = \theta P_{t+j}^{-\gamma}$$

$$k_{t+j} = q_{w,t+j}\frac{\alpha}{1 - \alpha}$$

$$h_{t+j} = \frac{Q_{t+j} - Q_{t+j-1}(1 - \delta)}{k_{t+j}^{\alpha}}$$

$$N_{t+j} = N_{t+j-1}(1 - \delta) + h_{t+j}$$

$$K_{t+j} = K_{t+j-1}(1 - \delta) + h_{t+j}k_{t+j}$$

$$E_{t+j} = N_{t+j} + n_h\frac{N_t}{K_t}K_{t+j}.$$

3.9.2   The stochastic process for wages

Estimating the stochastic process requires estimating the following objects: $\bar{w}_i$, inf, $h(w_{i,t} - \bar{w}_i)$, and G (the distribution of increases). I proceed in the following steps:

- Recover $\bar{w}_i$ as the location-specific fixed effect by running a regression on 1984:I - 2007:IV data:

$$\ln rmw_{it} = \bar{w}_i + \epsilon_{it},$$

  where rmw_{it} is the real minimum wage in location i in period t. Store the residuals, $\hat{\epsilon}_{it}$, which are the deviations from long-run averages.

- Assume that the h function takes a probit form. Let $Increase_{it}$ be an indicator for a minimum wage hike in period i in location t. To account for the fact that there are staggered increases, only count the first increase in a five quarter window as an increase, and to not count changes due to indexation an increase has to be least 5%. To measure the size, sum together the two increases. Estimate h based on the following regression:

$$Pr(Increase_{it} = 1|\hat{\epsilon}_{it}) = \Phi(\alpha_0 + \alpha_1\hat{\epsilon}_{it}),$$

  where Φ is the normal CDF.

- Estimate the distribution of increases, G, nonparametrically by taking the distribution of increases observed in the data.

- Compute the inflation parameter, inf, so that in expectation minimum wages do not have a trend. In practice, I use simulation to calibrate the inflation parameter by picking the inflation parameter such that after $100,000$ runs of the series the average wage is zero. This yields an annual inflation rate of 3.3%.

3.9.2.1 Does the minimum wage process vary across locations?

Operationalize *persistence* as the AR(1) coefficient of a basic time series regression on a single location i:

$$\hat{\epsilon}_{it} = \alpha_i + \rho_i \hat{\epsilon}_{it-1},$$

Does persistence, the estimated ρ_i, vary across locations? It does. The appropriate null hypothesis, however, is that the dispersion in ρ_i is equal to that from the estimated stochastic process. To construct this null distribution, simulate 50,000 minimum wage series of length 96 (the same length as in the data) based on the estimated stochastic process and estimate the same AR(1) regression.

Figure 3.6 plots the PDFs of the two distributions of ρ_i, where both distributions are centered at their respective medians (the median from regressions on data simulated from the stochastic process is 0.8895, and on the data is 0.9105.). Visual inspection of the Figure suggests that the dispersion in the estimated ρ_i in the data is very similar to what the estimated stochastic process would imply.

3.9.3 Simulation algorithm for uncertainty

This appendix describes how to simulate the model in section 3.6. This simulation takes into account the uncertainty in the wage path. Firms expect the minimum wage process to be that described in section 3.6.2. This modifies the algorithm described in the previous subsection in that the the present discounted values of prices and wages are computed via simulation.

The following describes how to simulate the present discounted value of wages and prices, and the product price, in a period given the log deviation.

- Given $\hat{\epsilon}_{i,t}$, simulate T periods of log deviations of real minimum wages using the stochastic process in appendix 3.9.2: $\{\hat{\epsilon}_{i,t+j}\}_{j=1}^{T}$.

- Convert to real levels using the long-run average in the real value of the minimum wage: $w_{i,t+j} = \bar{w}_i(1 + \hat{\epsilon}_{i,t+j})$, where $w_{i,t}$ is the real minimum wage and $\bar{w}_i$

is the long-run average of the real value of the minimum wage.

- Compute and store $q^s_{w,t} = \sum_{x=1}^{T}(\beta(1-\delta))^{x-1} w_{i,t+x}$.

- Repeat the previous three steps S times.

- For the present discounted value of wages: $q_{w,t} = \frac{1}{S}\sum_{s=1}^{S} q^s_{w,t}$.

- For the contemporaneous present discounted value of prices: $q_t = \alpha^{(-\alpha)}\left(\frac{q_{w,t}}{1-\alpha}\right)^{(1-\alpha)}$.

- For next period's expected present discounted value of prices:

$$E_t[q_{t+1}] = \frac{1}{S}\sum_{s=1}^{S}\left\{\alpha^{(-\alpha)}\left(\frac{q^s_{w,t+1}}{1-\alpha}\right)^{(1-\alpha)}\right\},$$

  where the $q^s_{w,t+1}$ are based on simulations using the period t wage.

- Compute the price: $P_t = q_t - \beta(1-\delta)E_t[q_{t+1}]$.

Set $T = 960$ and $S = 100$.

The remaining aspects of the simulation are identical to the simulation with perfect foresight.

3.9.4 Dube, Lester, and Reich (2010) dataset

Construct a county-level series of real quarterly minimum wages. Take the quarterly nominal minimum wages at the county-level from the Dube, Lester, and Reich (2010) dataset from 1984:I - 2007:IV (96 entries). Use as a measure of inflation the average hourly wage of production and non-supervisory workers from the Bureau of Labor Statistics (series code CES0500000008). This series is monthly. Convert it to quarterly using the geometric average of the three months. Create an index with 1984:I as the base period and equal to 1. Use this index to convert the nominal minimum wages to real minimum wages.[38]

Then simulate a county-level employment series using *realized* variation in minimum wages. For each county, start with the real minimum wage series from 1984:I - 2007:IV. In particular, the regression in appendix 3.9.2 implies for each county-period a real minimum log deviation pair: $(rmw_{it}, \hat{\epsilon}_{it})$. Feed through the sequence of $\hat{\epsilon}_{it}$ to

---

[38]Using a measure of wages rather than inflation to deflate nominal minimum wages to real wages is standard in the literature: Brown (1999) table 2 uses this deflator, and Baker, Benjamin, and Stanger (1999) deflate by the average manufacturing wage. Results are indistinguishable using the CPI to deflate the wage series, since inflation-adjusted average wages barely moved in this period.

the algorithm described in the previous subsection. Start with employment in the steady state defined by average minimum wages from 1984:I-2007:IV. Use this steady state to define the adjustment ratio for high-skill labor.

To deal with initial conditions, copy this minimum wage series and expectations 10 times. This means that the first 96 entries are 1984:I-2007:IV, and the 97th entry is 1984:I. Run the resulting 960 periods through the algorithm and store the final 96 observations.

3.10   Appendix: Additional specifications

Baker, Benjamin, and Stanger (1999) present three long-run specifications. Two specifications filter the minimum wage series into their high and low frequency components. The last specification is the distributed lag specification already presented in the text.

First, the "informal" filter splits minimum wage movements into low and high-frequency components, where these components are defined as the two-period moving average and the first difference of minimum wages. Baker, Benjamin, and Stanger use annual data and so to be consistent compute the moving average four periods apart and use fourth differences. This specification is:

$$\ln(\text{restemp})_{ipt} = a_1 + \eta_1 \frac{\Delta_4 \ln(MW)_{it}}{2} + \eta_2 \frac{(\ln MW_{it} + \ln(MW)_{i,t-4})}{2}$$
$$+ a_2 \ln(\text{totemp})_{ipt} + a_3 \ln(\text{pop})_{it} + \phi_i + \tau_t + \tau_{pt},$$

where η_1 is the elasticity with respect to high-frequency movements and η_2 is the elasticity with respect to low frequency movements. Table 3.8 shows that the low-frequency filter returns an elasticity on data simulated from the model of -0.07.

Second, the formal filter is based on a formal frequency decomposition of the natural log of the real minimum wage series. See Baker, Benjamin, and Stanger (1999, pg. 334) for details, or Hamilton (1994, pg. 159) for a statement of the relevant theorem. Baker, Benjamin, and Stanger decompose the real minimum wage series directly, I decompose the natural log. Doing the analysis in levels yields smaller coefficients on the simulated data. Following Baker, Benjamin, and Stanger enter the nine lowest frequencies as five sums, where adjacent frequencies are summed starting from the lowest, except for the highest frequency. They report that they enter five frequencies, which are each the sum of two frequencies. The 10th frequency is a

constant. The results are identical if the frequencies are entered separately. Table 3.9 shows that the lowest frequency movements in the minimum wage estimated on data simulated from the model yield elasticities of -0.07.

Table 3.6:   Probit model of the probability of a minimum wage increase

	Increase
Deviation	-6.673
	(0.182)
Constant	-1.991
	(0.016)
N	36480

Note: The outcome variable is an indicator for a minimum wage increase. The deviation is defined as the residual from a regression of log minimum wages on county fixed effects using 1984:I-2007:IV data. The sample is a modified form of the contiguous border county sample from Dube, Lester, and Reich (2010), where duplicate counties are dropped.

Table 3.7: Distributed lag specification

	Data	Model
$\Delta lnMW\_(t)$	0.082	-0.051
	(0.087)	(0.001)
$\Delta lnMW\_(t+1)$	0.093	-0.053
	(0.102)	(0.001)
$\Delta lnMW\_(t+2)$	0.087	-0.054
	(0.134)	(0.001)
$\Delta lnMW\_(t+3)$	0.087	-0.055
	(0.121)	(0.001)
$\Delta lnMW\_(t+4)$	-0.013	-0.060
	(0.087)	(0.001)
$\Delta lnMW\_(t+5)$	-0.010	-0.062
	(0.093)	(0.001)
$\Delta lnMW\_(t+6)$	-0.063	-0.066
	(0.082)	(0.001)
$\Delta lnMW\_(t+7)$	-0.065	-0.068
	(0.087)	(0.001)
$\Delta lnMW\_(t+8)$	0.029	-0.070
	(0.081)	(0.001)
$\Delta lnMW\_(t+9)$	-0.021	-0.072
	(0.091)	(0.001)
$\Delta lnMW\_(t+10)$	-0.003	-0.074
	(0.096)	(0.001)
$\Delta lnMW\_(t+11)$	0.019	-0.075
	(0.103)	(0.001)
$\Delta lnMW\_(t+12)$	0.012	-0.076

Table 3.7 – continued from previous page

	(0.102)	(0.001)
$\Delta lnMW\_(t+13)$	0.024	-0.078
	(0.099)	(0.001)
$\Delta lnMW\_(t+14)$	-0.027	-0.079
	(0.119)	(0.001)
$\Delta lnMW\_(t+15)$	0.029	-0.081
	(0.121)	(0.001)
$\Delta lnMW\_(t+16)$	0.010	-0.082
	(0.116)	(0.001)
$\Delta lnMW\_(t+17)$	0.035	-0.084
	(0.141)	(0.001)
$\Delta lnMW\_(t+18)$	0.020	-0.085
	(0.147)	(0.001)
$\Delta lnMW\_(t+19)$	0.018	-0.086
	(0.134)	(0.001)
$\Delta lnMW\_(t+20)$	0.024	-0.087
	(0.109)	(0.001)
$\Delta lnMW\_(t+21)$	-0.052	-0.087
	(0.112)	(0.001)
$\Delta lnMW\_(t+22)$	-0.048	-0.088
	(0.113)	(0.001)
$\Delta lnMW\_(t+23)$	-0.022	-0.089
	(0.138)	(0.001)
$lnMW\_(t+24)$	0.012	-0.094
	(0.172)	(0.002)
N	41448	41448

Note: The regressions also include a border pair times period interaction term, a measure of total employment and a measure of population. The outcome is total employment in the restaurant industry. The regressions use the contiguous border sample. The standard errors are clustered at the state and border segment separately. The data column reports estimates based on the data. The model column reports estimates based on data simulated from the putty-clay model.

Table 3.8: "Lag operator" filtered effects of minimum wages.

	Data	Model
High Frequency Minimum Wage	0.166	-0.029
	(0.158)	(0.004)
Low Frequency Minimum Wage	-0.008	-0.072
	(0.076)	(0.002)
N	41448	41448

Note: The high frequency minimum wage is an annual log difference. The low frequency minimum wage is an annual log moving average. The other variables are in logs. The regressions also include a border pair times period interaction term, a measure of total employment and a measure of population. The outcome is total employment in the restaurant industry. The regressions use the contiguous border sample. The standard errors are clustered at the state and border segment separately. The data column reports estimates based on actual data. The model column reports estimates based on data simulated from the putty-clay model.

Table 3.9:   Frequency decomposition-based dynamic lagged specification.

	Data	Model
Lowest Frequency	-0.001	-0.067
	(0.122)	(0.002)
Lower Frequency	0.017	-0.064
	(0.122)	(0.003)
Medium Frequency	0.183	-0.064
	(0.135)	(0.003)
Higher Frequency	0.116	-0.059
	(0.203)	(0.006)
Highest Frequency	-0.015	-0.051
	(0.206)	(0.008)
N	41448	41448

Note: The frequency decomposition is based on the logged real minimum wage series deflated using average hourly wage of private sector production and non-supervisory workers from 1987:2-2006:2. Lowest frequency are movements at a frequency at 19.2 year and 9.6 years; lower frequency are movements at 6.4 years and 4.8 years; medium frequency are movements at 3.85 and 3.2 years; higher is at 2.75 and 2.4; and highest is at 2.1 years. The regressions also include a border pair times period interaction term, a measure of total employment and a measure of population. The outcome is total employment in the restaurant industry. The regressions use the contiguous border sample. The standard errors are clustered at the state and border segment separately. The data column reports estimates based on actual data. The model column reports estimates based on data simulated from the putty-clay model.
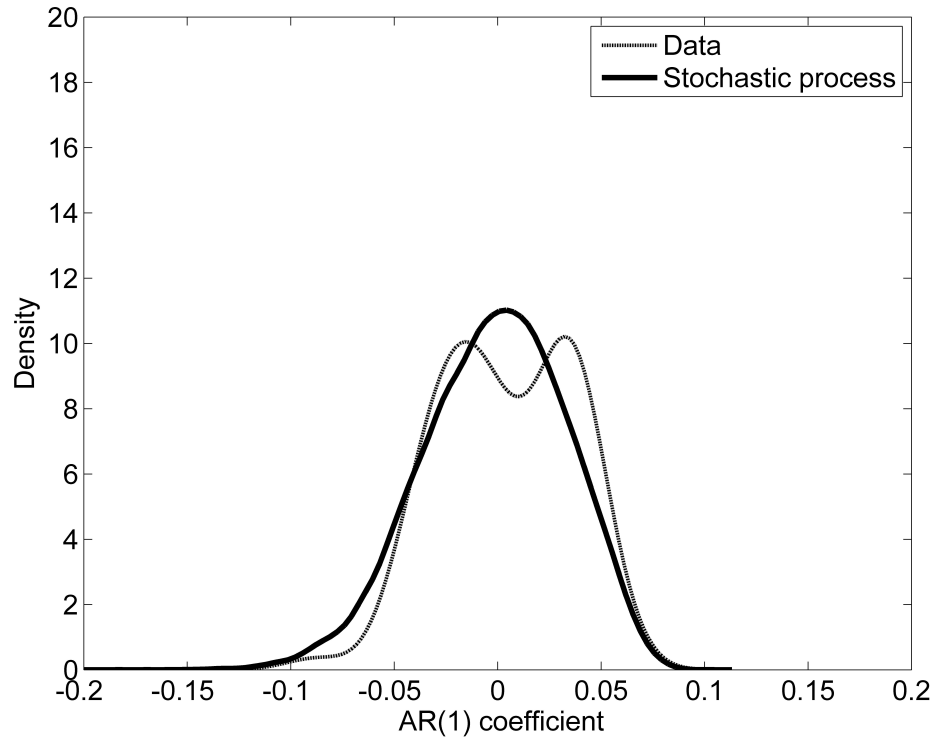
Figure 3.6: Persistence in the data and the estimated stochastic process

Note: This figure plots the PDF of the persistence coefficients of an AR(1) regression of minimum wages on lags of itself within location, with their median removed. The median from regressions on data simulated from the stochastic process is 0.8895, and on the data is 0.9105. The minimum wage is defined as the residual from a regression of log minimum wages on county fixed effects using 1984:I-2007:IV data. The sample is a modified form of the contiguous border county sample from Dube, Lester, and Reich (2010), where duplicate counties are dropped. The procedure to generate the distribution from the stochastic process is given in 3.9.2.1.

BIBLIOGRAPHY

BIBLIOGRAPHY

Aaronson, Daniel. 2001. "Price Pass-Through and the Minimum Wage." *Review of Economics and Statistics* 83 (1):158–169.

Aaronson, Daniel and Eric French. 2007. "Product Market Evidence on the Employment Effects of the Minimum Wage." *Journal of Labor Economics* 25 (1):167–200.

Aaronson, Daniel, Eric French, and James Macdonald. 2008. "The Minimum Wage, Restaurant Prices, and Labor Market Structure." *Journal of Human Resources* 43 (3):688–720.

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.

Ahn, Tom, Peter Arcidiacono, and Walter Wessels. 2011. "The Distributional Impacts of Minimum Wage Increases When Both Labor Supply and Labor Demand Are Endogenous." *Journal of Business and Economic Statistics* 29 (1):12–23.

Aizawa, Naoki and Hanming Fang. 2013. "Equilibrium Labor Market Search and Health Insurance Reform." Working paper.

175

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.

Atkeson, Andrew and Patrick J. Kehoe. 1999. "Models of Energy Use: Putty-Putty versus Putty-Clay." *American Economic Review* 89 (4):1028–1043.

Autor, David H., Alan Manning, and Christopher Smith. 2010. "The Contribution of the Minimum Wage to U.S. Wage Inequality over Three Decades: A Reassessment." Working Paper 16533, National Bureau of Economic Research.

Bagger, Jesper, Francois Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin. Forthcoming. "Tenure, Experience, Human Capital and Wages: A Tractable Equilibrium Search Model of Wage Dynamics." *American Economic Review* .

Bagger, Jesper and Rasmus Lentz. 2014. "An Empirical Model of Wage Dispersion with Sorting." Working Paper 20031, NBER.

Baker, Michael, Dwayne Benjamin, and Shuchita Stanger. 1999. "The High and Lows of the Minimum Wage Effect: A time-series cross-section study of the Canadian Law." *Journal of Labor Economics* 17 (2):2318–2350.

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.

Baskaya, Yusuf Soner and Yona Rubinstein. 2012. "Using Federal Minimum Wages to Identify the Impact of Minimum Wages on Employment and Earnings across the U.S. States." *Working Paper* .

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.

Bhaskar, V. and Ted To. 1999. "Minimum Wages for Ronald McDonald Monopsonies: A Theory of Monopsonistic Competition." *The Economic Journal* 109:109–123.

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. *BLS Handbook of Methods.*

Bonacich, Phillip. 1987. "Power and Centrality: A Family of Measures." *Journal of Public Economics* 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.

Brochu, Pierre and David A. Green. 2013. "The Impact of Minimum Wages on Labour Market Transitions." *The Economic Journal* 123 (573):1203–1235.

Brown, Charles. 1995. "Comment." *Industrial and Labor Relations Review* 48 (4):828–830.

———. 1999. "Minimum Wages, Employment, and the Distribution of Income." In *Handbook of Labor Economics*, edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier, 2101–2163.

Brown, Charles, Curtis Gilroy, and Andrew Kohen. 1982. "The Effect of the Minimum Wage on Employment and Unemployment." *Journal of Economic Literature* 20 (2):487–528.

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.

Caballero, Ricardo J. and Mohamad L. Hammour. 1994. "The Cleansing Effect of Recessions." *American Economic Review* 84 (5):1350–1368.

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.

Cahuc, Pierre and Andre Zylberberg. 2004. *Labor Economics.* Cambridge, MA: MIT Press.

Cameron, A.Colin, Jonah B. Gelbach, and Doug L. Miller. 2011. "Robust inference with multiway clustering." *Journal of Business and Economic Statistics* 29 (2):238–249.

Campbell, Jeffrey R. and Jonas Fisher. 2000. "Aggregate Employment Fluctuations with Microeconomci Asymmetries." *American Economic Review* 90 (5):1323–1345.

Campbell, Jeffrey R. and Beverly Lapham. 2004. "Real Exchange Rate Fluctuations and the Dynamics of Retail Trade Industries on the U.S.-Canada Border." *The American Economic Review* 94 (4):1194–1206.

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.

———. 2015. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." *Quarterly Journal of Economics* .

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.

Card, David and Alan B. Krueger. 1994. "Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania." *American Economic Review* 84 (4):772–784.

———. 1995. *Myth and Measurement: The New Economics of the Minimum Wage.* Princeton, NJ: Princeton University Press.

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.

Cooper, Russell and Jonathan Willis. 2009. "The Cost of Labor Adjustment: Inferences from the Gap." *Review of Economic Dynamics* 12 (4):632–647.

Davis, Steven J., R. Jason Faberman, and John Haltiwanger. 2006. "The Flow Approach to Labor Markets: New Data Sources and Micro-Macro Links." *Journal of Economic Perspectives* 20:3–26.

———. 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.

Dolado, Juan, Francis Kramarz, Stephen Machin, Alan Manning, David Margolis, and Coen Teulings. 1996. "The economic impact of minimum wages in Europe." *Economic Policy* 11 (23):317–372.

Dube, Arindrajit, T. William Lester, and Michael Reich. 2010. "Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties." *Review of Economics and Statistics* 92 (4):945–964.

———. 2011. "Do Frictions Matter in the Labor Markets? Accessions, Separations and Minimum Wage Effects." Working Paper 5811, IZA.

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.

Flinn, Christopher. 2006. "Minimum Wage Effects on Labor Market Outcomes under Search, Matching, and Endogenous Contact Rates." *Econometrica* 74 (4):1013–1062.

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

Gilchrist, Simon and John C. Williams. 2000. "Putty-Clay and Investment: A Business Cycle Analysis." *Journal of Political Economy* 108 (5):928–960.

Gittings, R. Kaj and Ian M. Schmutte. 2013. "Getting Handcuffs on an Octopus: Minimum Wages, Employment and Turnover." Working paper, University of Georgia.

Gorry, Aspen. 2011. "Minimum Wages and Youth Employment." Working paper, UC Santa Cruz.

Gourio, Francois. 2011. "Putty-Clay Technology and Stock Market Volatility." *Journal of Monetary Economics* 58 (2):117–131.

Greenstone, Michael. 2013. "IGM Expert Panel Question about Minimum Wages." Poll question of 2/26. URL `http://www.igmchicago.org/igm-economic-experts-panel/poll-results?SurveyID=SV_br0IEq5a9E77NMV`.

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

Haltiwanger, John, Henry R. Hyatt, and Erika M. McEntarfer. 2014. "Cyclical Reallocation of Workers Across Employers by Firm Size and Firm Wage." Working paper.

Hamermesh, Daniel, Wolter Hassink, and Jan van Ours. 1996. "Job Turnover and Labor Turnover: A Taxonomy of Employment Dynamics." *Annales d'Economie et de Statistique* (41/42):21–40.

Hamermesh, Daniel S. 1989. "Labor Demand and the Structure of Adjustment Costs." *American Economic Review* 79:674–689.

———. 1993. *Labor Demand.* Princeton, NJ: Princeton University Press.

———. 1995. "Comment." *Industrial and Labor Relations Review* 48 (4):835–838.

Hamilton, James D. 1994. *Time Series Analysis.* Princeton, NJ: Princeton University Press.

Hellerstein, Judith, David Neumark, and Melissa Mcinerney. 2008. "Changes in Workplace Segregation in the United States between 1990 and 2000: Evidence from Matched Employer-Employee Data." In *The Analysis of Firms and Employees: Quantitative and Qualitative Approaches*, edited by Stefan Bender, Julia Lane, Kathryn Shaw, Fredrik Andersson, and Till von Wachter. University of Chicago Press, 163–195.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and firm Dynamics in Long Run Equilibrium." *Econometrica* 60 (5):1127–1150.

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 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.

Johansen, Leif. 1959. "Substitution versus Fixed Production Coefficients in the Theory of Economic Growth: A Synthesis." *Econometrica* 27 (2):157–176.

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.

Keane, Michael and Kenneth Wolpin. 2002. "Estimating Welfare Effects Consistent with Forward-Looking Behavior, Part I: Lessons From a Simulation Exercise." *Journal of Human Resources* 37 (3):570–599.

Kennan, John. 1995. "The Elusive Effects of the Minimum Wage." *Journal of Economic Literature* 33 (4):1950–1965.

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.

Lazear, Edward. 1995. *Personnel Economics.* MIT Press.

Lemos, Sara. 2007. "Minimum Wage Effects on Wages, Employment and Prices in Brazil." *Resarch in Labor Economics* 26:397–413.

Lester, Richard A. 1946. "Shortcomings of Marginal Analysis for Wage-Employment Problems." *American Economic Review* 36 (1):63–82.

Lewis, Ethan G. 2011. "Immigration, Skill Mix, and Capital-Skill Complementarity." *Quarterly Journal of Economics* 126 (2):1029–1069.

Meer, Jonathan and Jeremy West. 2013. "Effects of the Minimum Wage on Employment Dynamics." Working Paper 19262, National Bureau of Economic Research.

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.

Neumark, David, J.M. Ian Salas, and William L. Wascher. 2013. "Revisiting the Minimum Wage-Employment Debate: Throwing Out the Baby with the Bathwater?" Working Paper 18681, National Bureau of Economic Research.

Neumark, David and William L. Wascher. 1992. "Employment Effects of Minimum Wages and Subminimum Wages: Panel Data and State Minimum Wages Laws." *Industrial and Labor Relations Review* 46 (1):55–81.

———. 2008. *Minimum Wages.* Cambridge, MA: MIT Press.

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.

Pereira, Sonia C. 2003. "The Impact of Minimum Wages on Youth Employment in Portugal." *European Economic Review* 47 (2):229–244.

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.

Pinoli, Sara. 2010. "Rational Expectation and the Puzzling No-Effect of the Minimum Wage." Discussion Paper 4933, IZA.

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.

Rebitzer, James B. and Lowell J. Taylor. 1995. "The Consequences of Minimum Wage Laws: Some New Theoretical Ideas." *Journal of Public Economics* 56 (2):245–255.

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.

Seltzer, Andrew J. 1997. "The Effects of the Fair Labor Standards Act of 1938 on the Southern Seamless Hosiery and Lumber Industries." *Journal of Economic History* 57 (2):396–415.

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

Sorkin, Isaac. 2015a. "Ranking firms using revealed preference." Work in progress.

———. 2015b. "Why do men and women work at different firms?" Work in progress.

Stigler, George. 1946. "The Economics of Minimum Wage Legislation." *American Economic Review* 36 (3):358–365.

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.