I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11966

Monopsony in Labor Markets: A Meta-Analysis

Anna Sokolova
Todd Sorensen

NOVEMBER 2018

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11966

Monopsony in Labor Markets: A Meta-Analysis

Anna Sokolova
University of Nevada, Reno

Todd Sorensen
University of Nevada, Reno and IZA

NOVEMBER 2018

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Monopsony in Labor Markets: A Meta-Analysis*

When jobs offered by different employers are not perfect substitutes in the minds of workers, employers gain wage-setting power; the extent of this power can be captured by the elasticity of labor supply that each employer faces. Estimates of this parameter reported by the literature vary broadly. We collect 801 estimates from published studies, record 20 aspects of each study's design and perform Bayesian Model Averaging to show that this observed variation is systematic and can be attributed to four groups of factors. First, estimates depend on methodologies used by the researchers: different specifications produce systematically different results that are also affected by whether the study employs an identification strategy; the choice between linear and non-linear estimation techniques also matters. Second, estimates vary with the underlying data: labor markets seem to be more competitive in Europe, and less competitive in developing countries - compared to the US, Canada and Australia. The market for medical workers appears to be more monopsonistic compared to others. Third, there is evidence of publication bias in parts of the literature, which results in negative estimates of supply elasticities receiving lower probability of being reported, and a (slightly) exaggerated mean. Fourth, estimates seem to vary with study quality, with top journals publishing higher estimates and studies using larger data sets producing more evidence of competitive behavior.

JEL Classification: J42, C83

Keywords: monopsony, labor supply, meta-analysis, Bayesian Model Averaging

Corresponding author:
Todd Sorensen
Department of Economics
University of Nevada, Reno
1664 N. Virginia Street
Reno, NV 89557
USA

E-mail: tsorensen@unr.edu

* We thank Laura Giuliano, Gautam Gowrisankaran, Tomas Havranek, Boris Hirsch, Alan Manning and Suresh Naidu for their feedback. We are also grateful for feedback received at the University of Leuvan, Université Libre de Bruxelles, Leibniz RWI and the University of Nevada, Reno. Anna Sokolova acknowledges support from the Basic Research Program at the National Research University Higher School of Economics (HSE) and from the Russian Academic Excellence Project `5-100'.

1 Introduction

Manning (2003) opens his book by questioning the standard assumption of perfect competition in labor markets with a simple thought experiment: "What happens if an employer cuts the wage it pays its workers by one cent?" If we are to believe in perfectly competitive labor markets characterized by individual firms facing perfectly elastic labor supply curves, then we also must believe that such a small wage cut would lead to a wholesale exodus of workers from the firm, as workers instantly and costlessly pick up work with competing employers. This narrative does not appear very plausible; it also contradicts the empirical evidence provided by a large body of literature on job search, that has clearly established the existence of frictions to mobility between jobs.

An alternative way of thinking about labor markets is to picture a model of imperfect competition in which firms have some wage-setting power—a model of monopsony, or monopsonistic competition. Monopsony as a labor market structure was first introduced by Robinson (1933) in her book *The Economics of Imperfect Competition*. Robinson argued that firms possessing wage-setting power can explain features that the model of perfect competition fails to capture, such as wage discrimination against marginalized groups. She also pointed out that employment in such a market would be less sensitive to changes in the minimum wage. The latter argument proved to be very powerful in light of the work of Card & Krueger (1994) and Card & Krueger (1995a) who empirically document the lack of an employment response to new minimum wage regulations, challenging the predictions of the conventional labor market model.

Over the last several decades, employers have grown in size and labor markets have become more concentrated (Azar *et al.* 2017). Some employers now use non-compete agreements that can bind workers to a specific firm, while others enter into no-poaching agreements with competing employers, further dampening labor market competition (Krueger & Posner 2018). These trends explain the present increase in interest toward the monopsony literature. Recent studies have argued that the monopsony model can shed light on a variety of observed phenomena, for example, firms paying for general training (Acemoglu & Pischke 1999), pro-cyclical real wages (Depew & Sørensen 2013, Hirsch *et al.* 2018b and Webber 2018), the urban wage premium (Hirsch *et al.*, 2018a), wage inequality (Card *et al.* 2018a and Krueger & Posner 2018), pay differences between groups of workers (Hirsch & Jahn, 2015) and bunching in wages (Dube *et al.*, 2018b). Several works also call for policy changes to limit employer wage-setting power, such as revising antitrust policy to address the impact of mergers on input markets, and labor markets in particular (Naidu *et al.*, 2018) and imposing more restrictions on non-compete and no-poaching agreements (Krueger & Posner 2018).[1]

Whether a monopsonistic or a perfectly competitive model is the best approximation for the labor market is ultimately an empirical question. The extent of firms' monopsony power, and thus the ability of the monopsony model to explain the phenomena mentioned above, can be

[1] However, Naidu & Posner (2018) argue that anti-trust policy is unlikely to be an effective tool for addressing monopsony power if market power exists even when employer concentration is low and absent factors such as non-compete or no-poaching agreements.

summarized by the level of the elasticity of labor supply to the firm. In a perfectly competitive labor market, labor supply should have infinite elasticity. While any finite positive value is consistent with a monopsonistic model, firm wage-setting power increases exponentially as the parameter decreases. Following the publication of Manning (2003), a new empirical literature estimating this parameter has blossomed. Although empirical methods have greatly evolved, the findings reported by different strands of literature on monopsony remain very diverse. Studies document different values of supply elasticity and, as a consequence, firm wage-setting power.

In this paper we attempt to pin down and examine the sources of this variation. Estimates may vary because of differences in estimation strategy used by researchers, or because of variation in data that causes the 'true' value of the supply elasticity parameter to vary across studies. It is also possible that the estimates are affected by variation in quality of research papers produced by the authors. Finally, published estimates may be affected by the preferences of the profession, with certain values receiving higher probability of being reported. Our goal is to disentangle these different sources of variation. To this end, we employ the framework of a meta-analysis to perform a quantitative synthesis of estimates produced by the literature.

In economics, meta-analysis as a research tool has been used by Card & Krueger (1995b) and Doucouliagos & Stanley (2009) who revisit the effect of increases in minimum wage on employment; Card *et al.* (2018b) who examine the effects of active labor market programs; Havranek (2015) who looks at elasticities of substitution reported in the consumption literature; Havranek *et al.* (2017) who study habit formation in consumption; Chetty *et al.* (2013) who present a meta-analysis of Hicksian and Frisch elasticities of labor supply, and others. To our knowledge, we conduct the first meta-analysis that synthesizes evidence on the elasticity of labor supply *to the firm* and quantifies the degree of firm wage-setting power.

We collect 801 estimates of the elasticity of labor supply to the firm reported in 38 published studies. The mean estimate reported in the literature is 3.75, implying that the last worker hired is paid about 79% of their worth. However, the standard deviation of 36.9 in the reported supply elasticity estimates raises concerns as to whether the literature as a whole produces reliable evidence for monopsony. We argue that most of this variation can be explained through differences in data and estimation methodology, publication characteristics and some selective reporting, and conclude that, in fact, the evidence for monopsonistic labor markets is quite strong.

For each of the 801 estimates, we also collect information on 20 aspects that govern study design, which we believe could contribute to observed variation in the estimates. The main challenge in evaluating the effects of these 20 variables is that we do not know which combination of these controls belongs to the 'true' data generating process for supply elasticity estimates, and therefore risk misspecifying the empirical model. We therefore employ Bayesian Model Averaging, a technique designed to address model uncertainty. We show that the most (economically) important distinctions stem from differences in estimation strategies employed in the studies. Papers that focus on the inverse elasticity of labor supply produce much larger estimates when converted to 'direct' elasticity, compared to studies that use other estimation strategies—consistent with predictions of Manning (2003) who argues that this stock-based approach may result in estimates that exhibit upward bias. We find that this bias is less pronounced in studies that employ an identification strategy.

For studies that use other approaches, the underlying supply elasticity parameter appears to be small; on random data, these papers would occasionally obtain estimates that lie in negative territory. However, negative estimates would imply a downward-sloping supply curve, a result that is hard to interpret. We show that such estimates receive lower probability of being reported, which results in the mean reported estimate being exaggerated and biased away from where the most precise estimates cluster. We note, however, that, although the effect of this "selective reporting" is statistically significant, it has a relatively low magnitude in economic terms.

Differences in data used by researchers are important, both statistically and economically. To our knowledge, we are the first study to produce evidence of systematic differences in monopsony power across countries. We show that estimates of the supply elasticity that come from data from developing countries display more evidence of monopsonistic labor markets when compared with advanced economies. For the advanced economies, estimates of the supply elasticity obtained on European data are higher than those for the US, Canada and Australia, suggesting that European labor markets might be more competitive.

Estimates of the elasticity of labor supply to the firm also appear to differ across some industries. Our results indicate that the market for medical workers appears relatively more monopsonistic, perhaps due to higher employer concentration. We also show that linear estimation methods produce results that are systematically different from those obtained via non-linear estimation techniques. Studies published in top journals and authors that have access to larger data sets report larger estimates of the supply elasticity. At the same time, the most cited results are those showing evidence of monopsony power, which could reflect the overall interest of the profession in this topic.

We provide fitted estimates of the elasticity of labor supply to the firm for different estimation strategies, conditional on 'best practice' (e.g. the study having a large and fairly fresh data set, being published in a high-ranking journal, correcting for selective reporting, etc.). Our approximation for the supply elasticity estimate obtained using data on worker separations and a hazard model is around 2.75 for American data with the conventional 95% confidence interval of $[1.21; 4.29]$ and a wider confidence interval of $[0.46; 5.88]$—with wild bootstrap clusters. This fitted estimate implies a wage markdown of 26.7%, and strong evidence for monopsonistic markets. Even at the upper bound of our 95% confidence interval, 5.88, the implied markdown is 14.5%.

2 Estimating the Elasticity of Labor Supply to the Firm

During the 20th century, the labor literature largely focused on the pure monopsony model in which a single firm comprised the entirety of demand for labor in a market (e.g. in a company town).[2] As a consequence, relatively little attention was paid to the more general case of imperfect competition, where several competing firms exercise wage-setting power. The foundation for this broader way of thinking about imperfect competition, however, was laid 85 years ago. Robinson (1933) described three specific reasons why the perfectly competitive model of the labor market may fail, even when there are many firms in the market competing for labor. She argued that a firm may end up facing an upward-sloping labor supply curve because of geographical isolation and differences in commuting distances to a work cite, because workers may prefer their employer for reasons other than compensation, or because workers may not be fully aware of opportunities existing at other firms. Such labor markets, in which a firm faces upward sloping supply despite the presence of many competitors, are termed monopsonistic (or oligopsonistic).

The Manning (2003) book *Monopsony in Motion* inspired a conceptual shift in the literature by applying the Burdett & Mortensen (1998) model to formalize the notion of a monopsonistic labor market, in which firms possess wage-setting power due to labor market frictions. His work also provided a relatively straightforward estimation framework, which paved the way for a new empirical literature on monopsony. In addition to papers estimating the elasticity of labor supply to the firm, recent work has begun to revisit possible causes of market power, focusing on issues such as input market concentration (Brummund 2011, Webber 2015, Azar *et al.* 2017, Benmelech *et al.* 2018 and Rinz *et al.* 2018), legal restrictions to mobility (Naidu 2010, Naidu & Yuchtman 2013, Balasubramanian *et al.* 2018 and Krueger & Ashenfelter 2018) and moving costs (Ransom 2018).

Here, we provide some background on the monopsony market structure and the key way to quantify firms' wage-setting power. Consider a firm that faces an upward-sloping labor supply curve and chooses the number of workers to solve the maximization problem

$$\Pi = \max_{L} \left[p \times f(L) - w(L) \times L \right], \tag{1}$$

where p is the price, L is the labor input, $f(.)$ is the production function, and $w(L)$ is the wage that the firm pays its workers, depending on how many workers are hired. This problem yields a solution that links the wage paid by the firm, the marginal revenue product of labor and the elasticity of labor supply:

$$w = MRP_L \frac{\epsilon}{1+\epsilon}, \tag{2}$$

where MRP_L is the marginal revenue product of labor and ϵ is the elasticity of labor supply to an individual firm with respect to the wage, $\epsilon \equiv \frac{\partial L}{\partial w}\frac{w}{L}$. If supply is perfectly elastic (and $\epsilon = \infty$),

[2] Manning (2003) demonstrates this by examining the contents of contemporary labor economics textbooks. In the meantime, other fields moved on to adopt models in which markets failed to yield perfectly competitive outcomes despite the presence of many firms in the market, on account of factors such as differentiated products (e.g. Berry *et al.* 1995, Krugman 1980 and Melitz 2003).

then the last worker hired is paid her worth to the firm: equation (2) implies $w = MRP_L$. By contrast, the worker is payed 90% of her worth to the firm if $\epsilon = 9$, and half of her worth if $\epsilon = 1$. It is, however, unclear that firms are able to exercise all of their monopsony power, as factors such as minimum wages, union contracts, social norms or worker responses to perceptions of fairness (see Dube *et al.* Forthcoming) may also affect wage outcomes. Nevertheless, this simple model does provide important insight into how monopsony power may affect wages, and ϵ, the elasticity of labor supply to the firm, provides important insights into the degree of wage-setting power that firms possess.

In this section we will discuss different ways in which the estimates of ϵ can be obtained. Perhaps the most straightforward approach for estimating the elasticity of labor supply involves a direct regression of the number of workers employed at a given firm on the wage paid to those workers:

$$\ln(L_i) = \epsilon \cdot \ln(w_i) + \xi_i \tag{3}$$

where L_i is labor employed by the firm, and w_i denotes wages payed. This approach is used by Bodah *et al.* (2003), Staiger *et al.* (2010), Falch (2010) and others. Authors that employ this method typically come up with estimates of elasticity $\hat{\epsilon}$ that do not exceed two, implying that workers are paid less than two thirds of their value to the firm.

An alternative approach that also uses the stock of workers employed by a firm at a given time reverses the left- and right-hand sides of the regression in equation (3) to estimate:

$$\ln(w_i) = \chi \cdot \ln(L_i) + \xi_i, \tag{4}$$

where $\hat{\chi}$ is the inverse elasticity of labor supply. This approach is employed in Fakhfakh & FitzRoy (2006), Sulis (2011), Matsudaira (2014) and others. A reader may expect that the estimates $\hat{\epsilon}$ and $\hat{\chi}$ would be linked through an inverse relationship, $\hat{\epsilon} = \frac{1}{\hat{\chi}}$, and therefore the estimates of $\hat{\chi}$ should cluster somewhere above 1/2. This, however, is not what this literature typically reports: the most common estimates $\hat{\chi}$ lie below 1/2, with only a small fraction exceeding this mark. This suggest some inconsistency and possible structural differences between the two estimation methods, a pattern previously pointed out by Manning (2003).

Manning (2003) provides an alternative framework that is not a stock-based, but a turnover-based approach. Motivated by the idea that perfect competition in labor markets fails due to several sources of frictions, this approach stems from the results of a simplified Burdett & Mortensen (1998) search model, in which firms face search costs, and frictions inhibit the mobility of workers between jobs. Workers choose to separate from jobs that pay lower wages, and the overall job separation rate is a function of the wage.[3] Card & Krueger (1995a) point out the relationship between the elasticity of separation with respect to wage and the labor supply elasticity

$$\epsilon = \epsilon_R - \epsilon_S. \tag{5}$$

In (5), $\epsilon_S \equiv \frac{\partial s(w)}{\partial w} \frac{w}{s(w)}$ is the elasticity of separations where $s(w)$ is the separation rate, and

[3]Readers interested in a more detailed discussion of this model may refer to Manning (2003) Chapter 4.4.

$\epsilon_R \equiv \frac{\partial R(w)}{\partial w}\frac{w}{R(w)}$ is the elasticity of new recruitment where $R(w)$ is the recruitment function. Equation (5) states that the elasticity of labor supply to the firm can be characterized by how the wage affects worker inflows (through the recruitment elasticity) and how it affects worker outflows (through the separation elasticity). It is rare that a researcher would have reliable data to competently estimate both ϵ_R and ϵ_S. A useful practical solution was suggested by Manning (2003): in a steady-state, the elasticities of separation and recruitment should be linked through $\epsilon_S = -\epsilon_R$. Under this assumption, two additional ways of estimating ϵ naturally arise:

$$\epsilon = -2\epsilon_S, \tag{6}$$

$$\epsilon = 2\epsilon_R. \tag{7}$$

Estimating the recruitment elasticity requires not only information about the employees of a firm, but also on how many qualified applicants a position received. This kind of data is hard to come by, so very few papers have estimated ϵ_R. Using high quality administrative data on Norwegian teachers, a field experiment in Mexico, and field data from Amazon Turk, Falch (2017), Dal Bó *et al.* (2013) and Dube *et al.* (2018a), respectively, provide estimates of the elasticity of recruitments with respect to the wage.

Estimating the separation elasticity requires the use of payroll data which contains information on the length of an employee's tenure at a firm and their wage. Measuring how tenure and wage covary identifies the separation elasticity. This approach is much more common, it was adopted, for example, in Ransom & Sims (2010), Booth & Katic (2011), Depew & Sørensen (2013) and others. Econometric models employed to estimate this relationship include linear probability models, probits, logits and hazard models. Studies estimating separation elasticities typically come up with numbers that imply supply elasticities less then two; at the same time, there are some studies that estimate it to be higher. Estimates obtained using recruitment elasticities appear to be slightly higher. An important research question is whether the assumption of $\epsilon_S = -\epsilon_R$ is in fact justified—this will be one of the questions we will attempt to address in Section 5 of this paper.

Finally, some researchers employ techniques that impose more structural assumptions than the papers estimating either the straightforward correlation between wages and labor supply or wages and turnover, i.e. Fleisher & Wang (2004); Naidu *et al.* (2016); Dobbelaere & Mairesse (2013); Ogloblin & Brock (2005).

An important caveat is the potential endogeneity problem that exists when modeling the relationship between wages and employment; understanding the effect of employing an identification strategy is therefore of crucial importance. Studies estimating ϵ via the regression model in (3) can use firm-specific shocks to the wage to identify the supply slope. This approach is taken by Falch (2010) who uses wage premiums paid to teachers in schools facing teacher shortages in Norway. On the other hand, studies that estimate χ with the regression model in (4) require labor demand shifters to identify the supply slope. For example, Matsudaira (2014) exploits increases in demand for nurses at the hospital level on account of a new staffing regulation. Studies that use data on separations to estimate ϵ_S can instrument for worker wages to purge unobserved individual heterogeneity, as is done in Ransom & Sims (2010) who use wages based upon union contracts as an instrument. For the estimate of ϵ_R based on recruitment rates, Dal Bó *et al.* (2013) run a field experiment to generate exogenous variation in wages.

3 Data

We first employ Google Scholar to search for published studies in the field; we prefer Google Scholar over other search engines because of its ability to search through the full text versions of the papers rather than only the abstract and keywords. We selected search parameters based on the following criteria: 1) the search would return papers related to monopsony and 2) it would return papers that *estimate* parameters of monopsony power. We ran the search on November 12th 2017, saved the .html files for the first 100 pages of the search and downloaded the .pdf files, when available, for the first 50 pages covering 500 papers. To verify that this list was indeed comprehensive, we also studied the references of the returned papers to include any potential candidates that we missed and added other published papers we were aware of.[4]

We adopted the following inclusion criteria. First, the study needed to present estimates that allow for computing the elasticity of labor supply to the firm. We therefore eliminated papers that examine the relationship between measures of labor market concentration and wages. Even though these studies can provide useful evidence of monopsony power on labor markets, they do not allow for a straightforward computation of the value of the supply elasticity. We also exclude papers estimating the firm size wage effect, unless such an effect was claimed by the authors to be an estimate of the elasticity of labor supply to the firm. Finally, we excluded papers that report estimates of the elasticity of labor supply to an entire labor market, rather than to an individual firm.

Our second inclusion criterion is that the study must report a standard error or present information from which the standard error can be computed, as we would like to investigate whether this literature is prone to publication bias.[5] Third, we limit our search to published studies, as they are likely to have gone through a peer review process and are not subject to further revisions and changes. Published studies are also likely to be better typeset, a feature that facilitates the data collection process.

We found 38 studies that comply with these criteria that together provide 801 estimates complete with standard errors. The search query and the list of studies are available in Appendix C. The oldest study in our data set was published in 1977, the newest—in 2018. Typically, each paper reports several estimates, and the authors do not explicitly state their preference over the reported results. We therefore do not discriminate between reported estimates and collect all results presented in each study.

We would like to investigate how different aspects of study design affect the reported estimate

[4] Specifically, we checked references in both Boal & Ransom (1997) and Manning (2011), which survey the monopsony literature.

[5] We use the delta method to approximate standard errors when the exact estimate is not available, assuming independence of parameters, as is standard in the literature.

of the supply elasticity. To this end, for each of the 801 estimates we also collect information on 20 features related to data, methodology and publication characteristics. The description and summary statistics of these variables are available in Table A1; we also discuss them in detail in Section 5. The final data set is available upon request from the authors.

As discussed in Section 2, estimates of the supply elasticity seem to vary depending on specifications used by researchers. On the one hand, many papers estimate effects that can, through linear transformations (and under assumptions discussed in Section 2), be converted to measures of the supply elasticity (e.g. studies that estimate ϵ with the model in (3), or report ϵ_S or ϵ_R). For convenience, we will refer to these estimates as 'direct'.[6] These estimates comprise 700 out of 801 estimates in our sample. They are depicted in Figure 1(a), with the median estimate around 1.1 implying that workers are payed 52% of their marginal product—strong evidence for monopsony. The distribution of these estimates appears to be relatively close to a bell-shaped curve, but, importantly, it is skewed: the right tail seems more prominent than the left tail, with many estimates clustering below the median and close to zero, signaling even more monopsony power than the median estimate suggests.

The remaining 101 estimates in our data set come from studies estimating the inverse elasticity of labor supply (parameter χ in the model in 4); we depict their distribution in Figure 1(b). The median inverse elasticity is around 0.25, corresponding to a supply elasticity around 4 and a wage markdown of only 20%. This immediately points to an inconsistency between two sets of results, suggesting that there may be deep structural differences between the two approaches. However, there may be other explanations as well. For example, papers estimating inverse elasticities could, by chance, be studying less monopsonistic markets or using techniques that yield larger estimates.

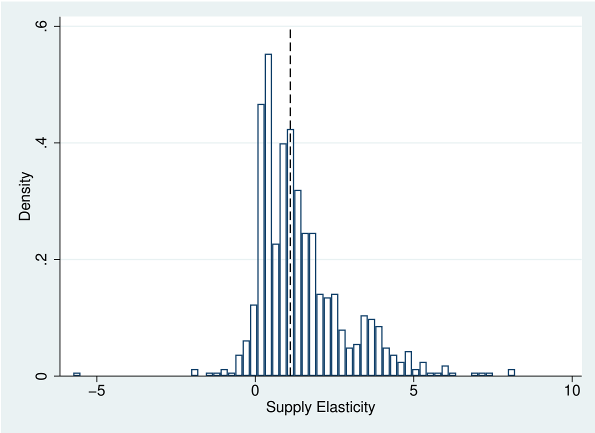
Figure 1(c) plots all estimates of ϵ together, combining those obtained using 'direct' approaches and the inverse regression (i.e. model 4). Again, we note striking differences between these sets of results as they do not appear to come from the same distribution. Table 1 reports sample statistics for the full sample, as well as the subsamples of estimates obtained through 'direct' and inverse methods. For the overall sample, the mean estimate of the supply elasticity is at 3.75, while the median is much lower—only 1.27; we also observe similar patterns when we weigh estimates by the inverse of the number of estimates per study, thereby giving equal weight to each study, regardless of how many estimates it reports. The sample means for 'direct' estimates appear to be somewhat lower (1.46), while the means for inverse estimates are substantially higher (19.66), and very different from the median of 3.77.

Elasticity estimates vary across other dimensions as well. First, we document some variance across geographic regions. The means and medians for estimates coming from Europe are larger than those from other advanced economies and developing countries. This could potentially imply that European labor markets are more competitive. Alternatively, this result could also arise from the fact that a portion of the estimates of the inverse elasticity were obtained using European data—if structural differences between inverse and direct estimations are in fact

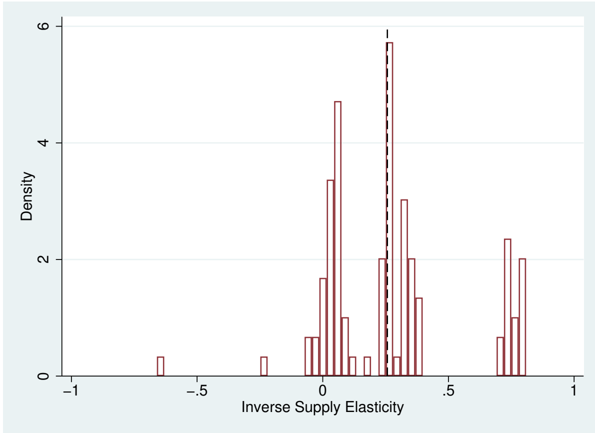
[6] Importantly, this notation is different from the terminology of Manning (2003), who uses the term 'direct regression' to exclusively refer to 'stock'-based regressions of the wage on the stock of labor (see the model in 4).

Figure 1: Estimates of supply elasticity: ‘direct’ vs. ‘inverse’

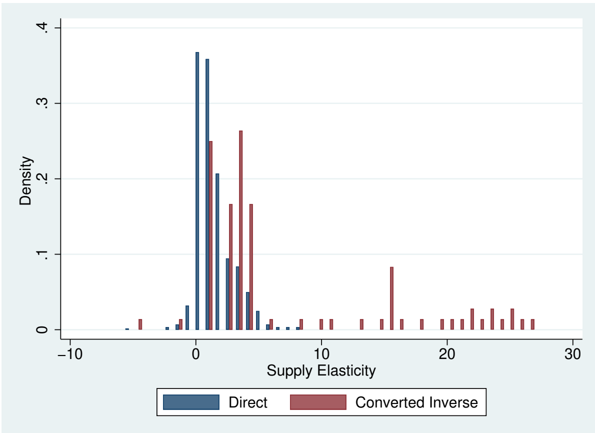
(a) Direct elasticity estimates



(b) Inverse elasticity estimates



(c) All estimates



Notes: The figure displays the distribution of estimates of the elasticity. Figure 1(a) shows estimates of ϵ that were obtained via the ‘direct’ methods; that is, methods that allow for a calculation of ϵ via a linear transformation (i.e. from model (3), using separation or recruitment rates, or performing a structural estimation). Figure 1(b) shows estimates of χ obtained from regression (4). We then convert these estimates to the elasticity of labor supply to the firm using $\epsilon = \frac{1}{\chi}$ and plot the pooled data set in Figure 1(c) (here, we show only estimates between -10 and 30).

important. We also observe that estimates for developing countries appear to be smaller than those corresponding to advanced economies. It is, however, too early to conclude that the labor markets of developing countries are less competitive, as we do not know what other features of the study designs are contributing to this result. We will attempt to disentangle the potential explanations in Section 5.

Aside from geography, we also observe some differences across occupations. A large portion of the literature exclusively focuses on markets for medical workers and teachers, on the grounds of higher potential for monopsony in these markets due to higher employer concentration. There are 203 estimates in our sample exclusively related to either of these markets. From the sample statistics, it would appear that the market for nurses is less competitive compared to the market

Table 1: Supply elasticity estimates by data and methods

	Unweighted				Weighted				
	Mean	Median	5%	95%	Mean	Median	5%	95%	N
All	3.75	1.27	-0.07	5.91	4.70	1.56	-0.35	10.87	801
Europe	7.01	1.43	0.24	19.96	11.28	1.84	0.32	23.31	336
Other advanced	1.43	1.17	-0.26	4.60	0.83	1.47	-4.38	5.12	406
Developing	1.16	1.16	-0.30	2.21	0.96	1.03	-0.35	2.55	59
Nurses	0.95	1.38	-4.38	4.10	-2.65	0.77	-27.36	3.79	78
Teachers	3.08	2.95	1.04	5.44	5.07	3.65	1.06	17.06	102
Inverse	19.66	3.77	-4.38	38.46	21.12	3.77	-27.55	83.33	101
Direct	1.46	1.10	-0.04	4.21	1.67	1.38	-0.07	4.18	700
Separations	1.53	1.22	-0.13	4.36	1.96	1.67	0.23	5.84	496
Recruitments	3.04	3.41	0.15	4.73	2.14	2.13	0.07	4.35	62
L on w	0.87	1.03	0.02	1.69	0.72	0.77	0.02	1.56	47
Structural	0.33	0.31	0.13	0.43	0.44	0.30	-0.35	2.55	95
Top Journal	1.96	1.51	0.08	3.76	2.49	1.74	0.13	5.62	93

Notes: 5% and 95% denote corresponding percentiles. 'Weighted' refers to summary statistics based upon weighting of observations by the inverse of the number of estimates reported in the study, thereby giving each study equal weight.

for teachers and the results coming from other occupations.

Out of the 700 'direct' estimates in our sample, the majority of about 500 estimates comes from studies that use separation rates. The remaining approximately 200 estimates are derived from studies using recruitment rates, regressing labor supply on wage, or using some type of structural estimation. There seems to be some, albeit much smaller, variation across these dimensions as well. Finally, 93 of the estimates in our data set come from papers published in either one of the top five general interest journals, or the top field Journal of Labor Economics (labeled 'Top Journal' in Table 1). These estimates appear quite close to the sample mean of the 'direct' estimates. Overall, there is relatively low variation in 'direct' estimates of the supply elasticity. At the same time, the skewed distribution of those estimates appearing in Figure 1(a) may indicate publication bias in the literature, with negative estimates receiving lower probability of being reported. We investigate these concerns in the next section.

Before proceeding with the estimations, we need to make provisions to improve comparability between inverse and non-inverse estimates. All estimates of supply elasticity obtained via 'direct' methods lie between -6 and 8.5. At the same time, some of the studies estimating the inverse elasticity come up with estimates of $\hat{\chi}$ that lie very close to zero; these estimates become enormous when converted to $\hat{\epsilon}$. Our full sample of 801 estimates includes several estimates converted from inverse elasticity estimates that do not compare with the rest, such as 999.9 with a standard error of 6666.6; 249.9 with a standard error of 520.8, -76.9 with a standard error of 120.2, etc. In order to ensure that we are working with comparable data, we cut the outliers by 1% from each tail. This leaves us with a sample of 787 estimates among which 88 are converted from the inverse elasticity, enough to estimate the contribution of this methodology to the magnitude of supply elasticity estimates. Table A1 compares sample statistics of our control variables for the full sample and the subsample of the 98% of estimates without outliers;

it shows no notable difference between the two samples in terms of the sample properties of key controls. In the next two sections we will focus on this subsample; however, we will also test robustness of our results using the full data set in which outliers from each tail are winsorized at the 1% level.

4 Publication Bias

Estimates of the supply elasticity that are based on 'direct' methods seem to cluster relatively close to zero, implying that the underlying parameter is close to zero as well. When estimated on random data using standard techniques, a model with a small positive underlying parameter would sometimes yield estimates that lie quite far from the true value and are associated with large standard errors. Some of these estimates would be large and positive, while others, given the small 'true' value, would end up in the negative territory. If all estimates of the supply elasticity are reported, then averaging across different results should nevertheless yield a mean close to the underlying effect. If, however, some (e.g. negative) estimates are under-reported, then the mean of this truncated distribution would likely be far from the 'true' effect. What we will investigate here is whether the literature is prone to such 'selective reporting' of the results.[7]

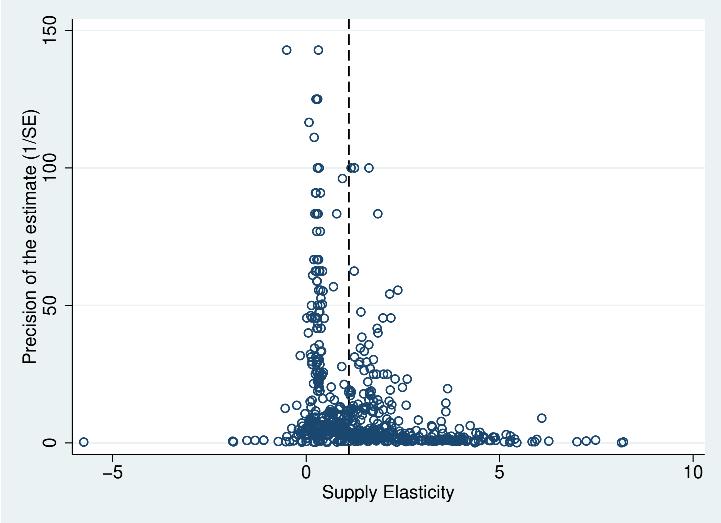
Selective reporting seems to be present in many fields of economics. Ashenfelter *et al.* (1999) find publication bias in the literature estimating returns to schooling; Card & Krueger (1995b) and Doucouliagos & Stanley (2009) document this for studies of the effect of minimum wage regulation on employment. Rose & Stanley (2005) and Havranek (2010) examine literature on the effects of currency unions on trade and find that negative estimates have lower probability of being reported. Similarly, Havranek & Sokolova (2018) find evidence of 'selection for the right sign' in the literature estimating the degree of excess sensitivity in consumption to predictable changes in income.

Positive values of the elasticity of labor supply to the firm, however large, can easily be interpreted by researchers: a large elasticity indicates that the labor market is close to perfect competition, while an estimate close to zero implies high firm wage-setting power. The same cannot be said for negative values of the supply elasticity, as they imply a downward-sloping supply curve and are therefore much harder to make sense of. It is possible that researchers obtaining negative results would see them as an indication of something being wrong with their model, and would possibly engage in further specification searches. These patterns, albeit unintentional, would lead to a lower probability of reporting for negative estimates which in turn implies that, when averaging results across studies, the mean estimate produced by the literature would exaggerate the 'true' underlying effect.

Figure 3 presents a scatter plot of estimates reported by studies of the 'direct' elasticity.

[7] 'Selective reporting' might be a better, more general description compared to 'publication bias', as the observed under-reporting of the results may not actually be related to the publication process. Nevertheless, the literature has converged on the term 'publication bias' (e.g. Card & Krueger 1995b, Ashenfelter *et al.* 1999, Stanley 2001, Efendic *et al.* 2011, Havranek 2015, Rusnak *et al.* 2013). Here, we also use it for consistency.

Figure 2: Funnel plot of 'direct' elasticity estimates



Notes: The figure plots estimates that were obtained via the 'direct' methods, i.e. estimates shown in Figure 1(a). In the absence of 'selection for the right sign', the funnel would exhibit symmetry around the most precise results.

The values of estimates obtained are plotted against their precision. We observe that the most precise estimates seem to cluster close to zero; this seems to imply that the underlying 'true' elasticity parameters should be rather small. In the absence of selection for the 'right sign', the funnel should appear symmetrical, with less precise estimates being distributed around the 'true' effect (see Egger *et al.* 1997). The funnel on Figure 3 is skewed: the right tail is much more prominent compared to the left tail. It appears that a substantial portion of negative estimates is missing from the funnel plot, which seems to point towards publication bias in the form of selection for a positive sign.

To further investigate possible publication bias, we conduct a formal funnel asymmetry test originally proposed by Card & Krueger (1995b). Common estimation methods rely on the assumption that the ratio of the estimate to its standard error is t-distributed. Under this assumption (or assuming any other symmetrical distribution), the estimate and the standard error should not be correlated. Therefore, in a regression of the estimate on its standard error, the coefficient λ on the standard error should be zero:

$$\hat{\epsilon}_{ij} = \epsilon_0 + \lambda \cdot SE(\hat{\epsilon}_{ij}) + u_{ij}, \tag{8}$$

where $\hat{\epsilon}_{ij}$ is the i-th estimate from the j-th study, $SE(\hat{\epsilon}_{ij})$ is its standard error, and u_{ij} is the disturbance term. By contrast, systematic under-reporting of negative estimates would result in a positive relationship between the estimate and the standard error, and a positive coefficient λ in the regression (8)—see Stanley (2005) for a detailed discussion. The coefficient λ can thus be viewed as a measure of the severity of publication bias, while the constant term ϵ_0 gives an approximate value of the unbiased effect.[8]

We estimate model (8) and report the results in Table 2. It is likely that estimates are correlated within studies; we therefore cluster the standard errors at the study level. As our number of clusters is relatively small (33), standard errors from clustered inference may exhibit downward bias. We therefore additionally compute wild bootstrapped clustered p-values, as recommended by Cameron *et al.* (2008). The first column of Table 2 shows the results of OLS estimation of model (8). The coefficient λ appears to be positive and significant, albeit relatively small in magnitude. In the second column we control for study-level fixed effects, accounting for unobserved study-level characteristics. The estimate of the effect of publication bias here is again positive, but less significant compared to OLS. In the third column we only use variation between studies and again find evidence for publication bias, although the number of observations used drops more than tenfold.

Table 2: Testing for publication bias

	OLS	FE	BE	Precision	Study
SE	0.423	0.118	4.016	1.996	0.971
	(0.001)	(0.000)	(0.000)	(0.000)	(0.048)
	[0.123]	[0.313]	[0.000]	[0.016]	[0.000]
Constant	1.255	1.409	0.481	0.459	1.246
	(0.000)	(0.000)	(0.468)	(0.002)	(0.000)
	[0.000]	.	[0.290]	[0.091]	[0.000]
Studies	33	33	33	33	33
Observations	699	699	33	699	699

Notes: The table presents results from the following regression: $\hat{\epsilon}_{ij} = \epsilon_0 + \lambda \cdot SE(\hat{\epsilon}_{ij}) + u_{ij}$, where $\hat{\epsilon}_{ij}$ is the i-th estimate from the j-th study, $SE(\hat{\epsilon}_{ij})$ is the standard error of the estimate, and u_{ij} captures the unobservables in the regression. Standard errors from the regression are clustered at the study level and p-values are shown in parenthesis. We also report p-values from wild bootstrap clustering in square brackets. This is implemented via the `boottest` command in `Stata` (see Roodman 2018). We use Rademacher weights and 9999 replications. The package does not allow for computation of a bootstrapped p-value for the constant term in the fixed effects specification. 'OLS' denotes ordinary least squares, 'FE' is study-level fixed effects, 'BE' is study-level between effects, 'Precision' is a specification with precision weights, and 'Study' is a specification with weights based on the inverse of the number of estimates reported in the study. Here, we use only 699 observations, rather than 700, as our data trimming procedure eliminates one of the 'direct' estimates. Results are similar when all 700 observations are included.

We also apply two alternative weighting strategies to further check robustness of these results. We first weight all estimates by their precision, effectively multiplying equation (8) by the inverse of the standard error. This approach remedies the apparent heteroskedasticity, while at the same time giving more weight to the more precise estimates (see Stanley & Doucouliagos 2015 for a discussion). For our data, precision weighting yields strong evidence for publication bias that is more pronounced when compared to the OLS results. It is worth noting that this technique is not without some caveats. First, it is possible that some estimation methods would produce standard errors that are systematically smaller in magnitude: for example, we expect studies that do not use instrumental variable techniques to report lower standard errors than studies with instruments, other things equal. Weighting by precision would then assign lower

[8]The interpretation of ϵ_0 should be done with caution as the estimate is unbiased only when publication selection is proportional to the standard error. Nevertheless, this linear approximation was documented to work reasonably well in Monte Carlo simulations (e.g. Stanley 2008).

importance to studies that use IV. Furthermore, Lewis & Linzer (2005) show that for models with an estimated dependent variable, a simple OLS would often outperform the weighted estimation. For the literature that we study, the distinction between identified and unidentified estimates is important. We therefore will rely more heavily on the unweighted specification.

Studies in our sample typically report several estimates of the supply elasticity, and we collect all of the estimates reported in each study and explore both within- and between-study variation. However, some studies report many more estimates than others—those studies would then effectively have greater weight in the estimation strategies discussed above. To correct for this potential bias, we weight our data by the inverse of the number of estimates per study and report the results in column five. This strategy also produces results that favor publication bias, though the effect is less pronounced compared to precision weights.

One potential problem with these strategies is that there may be aspects of study design that influence the estimate and the standard error in the same direction. For example, estimates converted from inverse elasticities shown in Figure 1 are systematically larger than the direct estimates, and so are their standard errors. Therefore, when we do publication bias tests of Table 2 on the full sample of 801 estimates, without excluding the inverses, the evidence for publication bias becomes stronger due to this spurious correlation. We therefore limited our focus here to studies that produce 'direct' estimates that seem to be much more homogenous.[9] It is possible, however, that there are other aspects of methodology that could create similar biases within this group. One possible solution to this problem would be to find an instrument for the standard error that is uncorrelated with other aspects of study design. One such instrument could be the number of observations used to produce the results. For this data we find the number of observations to perform poorly in predicting the standard error, which undermines the credibility of the results based on that approach.[10] We nevertheless report them in the appendix in Table B1.

The next section presents an alternative solution to this endogeneity problem. In an effort to explain variation across estimates, we will attempt to control for all aspects of study design that we deem most likely to influence the estimation results. We find evidence for publication bias in this context as well; however, the magnitude of the bias seems to be relatively low. We conclude that, even though economists tend to under-report negative estimates of the supply elasticity, this does not lead to a drastic exaggeration of the mean. Unlike Stanley (2005), who finds that publication bias results in the mean estimate of the elasticity of water demand being four times higher than the unbiased value, we document a very modest effect of the bias on the supply elasticity.

[9] We also performed the publication bias test on the sample of remaining estimates of the inverse supply elasticity depicted in Figure 1. We did not find convincing evidence for publication bias. It is, however, worth noting that this subsample consisted of estimates coming from only six studies.

[10] The first stage coefficient on the instrument is not significant at conventional levels and fails weak identification tests.

5 Why do Estimates of Supply Elasticity Vary?

5.1 Explanatory Variables

So far we have noted a few methodological aspects that are likely to have systematic effects on the estimates of the elasticity of labor supply to the firm. Most importantly, it seems that the estimates are much higher for studies that measure the inverse supply elasticity, and lower for those employing 'direct' methods. There are, however, other aspects of study design that could be affecting the estimates. We will now attempt to control for a subset of these features that a) we believe are important and b) vary sufficiently across studies, while at the same time are not collinear. Our goal is to understand the effects that the researcher's data and method choices have on their inference about firms' monopsony power. To this end, we come up with a set of 20 controls that, we believe, capture the most crucial features of the studies, such as data used and overall study quality, and the most common decisions that researchers make, such as choosing specification and estimation technique. We group these controls into five categories and discuss them below. We also present a full list of controls, their definitions and summary statistics in Table A1.

Data characteristics. It is likely that monopsony power of the firms has changed over the years; we therefore control for the age of the data set by including the midpoint of the data. Next, we include the logarithm of the number of observations used to obtain each estimate, as we believe that results obtained from large data sets may be more reliable. Ashenfelter *et al.* (1999) shows that failing to control for publication bias in the context of meta-regression can result in exaggerated effects attributed to different estimation methods. We therefore include an interaction between the standard error of the estimate and an indicator variable that equals one for estimates obtained through 'direct' methods—to capture publication bias discussed in Section 4.

We also expect that markdown could differ across different demographic groups. For example, Ransom & Oaxaca (2010) find the labor supply of female employees more elastic compared to males. Nine papers in our sample examine gender differences in the supply elasticity, e.g. Galizzi (2001) and Hirsch *et al.* (2010). Several studies in our sample report estimates for males and females separately, and many studies report the female share in the sample they use. We capture this information in a control $female\,share$. Unfortunately, for a substantial portion of the estimates, the information on the demographic structure is not reported. We set $female\,share = 0.5$ for these cases.[11]

Country & Occupation. Strength of institutions varies across countries; it is reasonable to expect that monopsony power of firms would vary as well. To our knowledge, no cross-country studies exist to examine these differences in labor supply elasticity—we are the first study attempting to gather systematic evidence on this topic. Our sample spans data coming from fourteen

[11] We also re-run the estimations assuming $female\,share = 0.3$ for studies missing demographic data; our results are qualitatively robust to this adjustment.

countries, and the papers on gender alone cover six (Russia, Norway, Brazil, Italy, Germany, US). We group the country data into three categories: *Developing* (5 countries), *Europe* (6 countries) and *Other Advanced* (3 countries). The last category is our reference group, it describes the data coming from the US, Canada and Australia and covers 51% of our data set.

Prior to the work of Manning (2003), research on monopsony mostly focused on labor market concentration rather than labor market frictions. Accordingly, much of the literature turned its attention to studying firm market power over nurses and teachers, as these workers are often employed in firms that are large relative to their labor market. In our sample, 21% of estimates are from studies of nurses or teachers. We construct controls that reflect whether the estimate relates exclusively to one of these occupations.

Method & Identification. As discussed in Section 2 and Section 3, one major distinction among the results produced by the literature is between estimates obtained from inverse supply elasticities and those obtained via other methods (that we term 'direct'). The former is a 'stock-based' estimation approach that uses correlation between the wage and the overall number of workers employed by the firm (see model 4 in Section 3). Manning (2003) argues that estimates obtained with this method may be biased on account of unobserved labor supply shocks 'making the slope of the supply curve seem less positive than it really is'. This argument implies that estimates converted from inverse elasticities would exhibit upward bias, a conclusion that so far seems to be in line with sample statistics presented in Section 3. Biases could also arise due to unobserved worker quality, rent sharing, and compensating wage differentials. A firm-specific labor demand shifter could identify the (inverse) elasticity of labor supply in such contexts, reducing the bias. There could therefore arise a systematic difference between estimates obtained with an identification strategy in place and those produced without one. We create controls for identified and unidentified estimates converted from inverse elasticities.

In a stock-based regression of employment on wages (i.e. model 3), a bias in the opposite direction may arise—see Manning (2003). Again, firm-specific shocks (this time to wages) would provide clean identification. In our sample, all estimates obtained via this method are identified through either an IV or a randomized wage strategy. We therefore cannot distinguish between identified and unidentified estimates, and only include a control for the method itself.

Compared to the stock-based methods, turnover-based methodologies that use either separation or recruitment rates employ individual-level data and therefore are subject to much less simultaneity. Nevertheless, threats to identification may still exist. For example, workers with unobservable characteristics which increase their productivity may be rewarded with higher wages as well as more outside job offers, resulting in higher separation rates. We distinguish between separation-based estimates obtained with and without an identification strategy. We do not make the same distinction for recruitment-based estimates, as only seven of them are obtained without an identification strategy in place. Finally, we control for estimates obtained in models that impose additional structural assumptions (e.g. models with production), with and without an identification strategy; we investigate how they compare with the rest of the literature.

Estimation technique. As noted above, there is more than one way to estimate the elasticity of labor supply to the firm, even for a given method such as the separation-based approach. For example, a researcher may run a linear probability model, where a binary outcome of separation from employment depends on the wage, and calculate the separation elasticity (e.g. Depew & Sørensen 2013). Alternatively, a researcher may choose a binary non-linear model, such as probit or logit (e.g. Ransom & Oaxaca 2010). Finally, a number of recent studies have used survival analysis in order to estimate the hazard of separation from employment as a function of the wage (e.g. Hirsch *et al.* 2018b). We assess the impact of these estimation techniques by including corresponding controls. Other estimation technique choices (i.e. OLS versus IV) are largely dictated by whether the study employs an identification strategy and are partially captured by our method-identification controls.

Publication characteristics. Supply elasticity estimates could also vary with unobserved features of the papers related to quality. We control for publication characteristics in an effort to capture some of this variation. We have 93 estimates coming from papers published in either one of the top five general interest journals, or the top ranked field *Journal of Labor Economics*, and we include a corresponding control to account for outlet quality.[12] Next, we constructed a control that records the number of citations listed on Google Scholar, divided by the number of years since the paper first appeared on Google Scholar, to see whether the profession tends to favor certain results over others. We also record publication years for the studies, accounting for advances in statistical methods that occurred over time that perhaps are not fully captured by our controls for methods and techniques.

5.2 Estimation and Results

We would like to pin down the sources of observed variation in supply elasticity estimates—in the previous section we presented our best guess as to what the key sources might be. The effects of some of these factors can be explored within a framework of a single study dealing with labor market data. For example, a researcher could estimate the supply elasticity via different methods using a single data set, and compare results. This is the approach taken by Manning (2003) who draws comparisons between different methods of estimating labor supply elasticity. A researcher could also examine differences in pay of male and female workers within a single firm, as is the case in Ransom & Oaxaca (2010). This within-study comparison approach can shed light on the importance of some features of methodology and data, and the previous section of the current paper builds on the insights coming from the respective studies. At the same time, if the task set by a researcher is to explain overall variation in estimates reported by the literature, this approach would have serious shortcomings.

Estimates of the supply elasticity can differ for a variety of reasons: there could be variation in the 'true' underlying parameter across markets and regions that affects estimation results;

[12] We also considered including the impact factors of the journals, but were forced to exclude this control because of multicollinearity.

there could be certain combinations of methodology, identification strategies and data features that produce evidence of very strong or very weak monopsony power. Finally, there could be variation in the quality of published research papers. All of these features could contribute to observed variation in estimates, and drawing full and consistent comparisons within a framework of a single study of labor market data may be an impossible task. We therefore resort to a more feasible method that, rather than estimating elasticities for each plausible choice of study design and data, utilizes *estimates* obtained by previous studies to perform a meta-regression analysis. We consider the following regression model:

$$\hat{\epsilon}_{ij} = \alpha_0 + \sum_{l=1}^{20} \beta_l X_{l,ij} + u_{ij}, \tag{9}$$

where $\hat{\epsilon}_{ij}$ is estimate i of the supply elasticity reported in study j, $X_{l,ij}$ are values of controls reflecting study design and quality discussed in subsection 5.1 and summarized in Table A1, and u_{ij} is the disturbance term. The model in equation (9) attempts to capture key features of the process governing how researchers obtain estimates of the supply elasticity. We do not claim that this methodology can explain the evolution of the 'true' supply elasticity parameter—but we argue that we employ the second best, feasible option that, nevertheless, allows us to explain much of the variation in estimates reported by the existing literature.

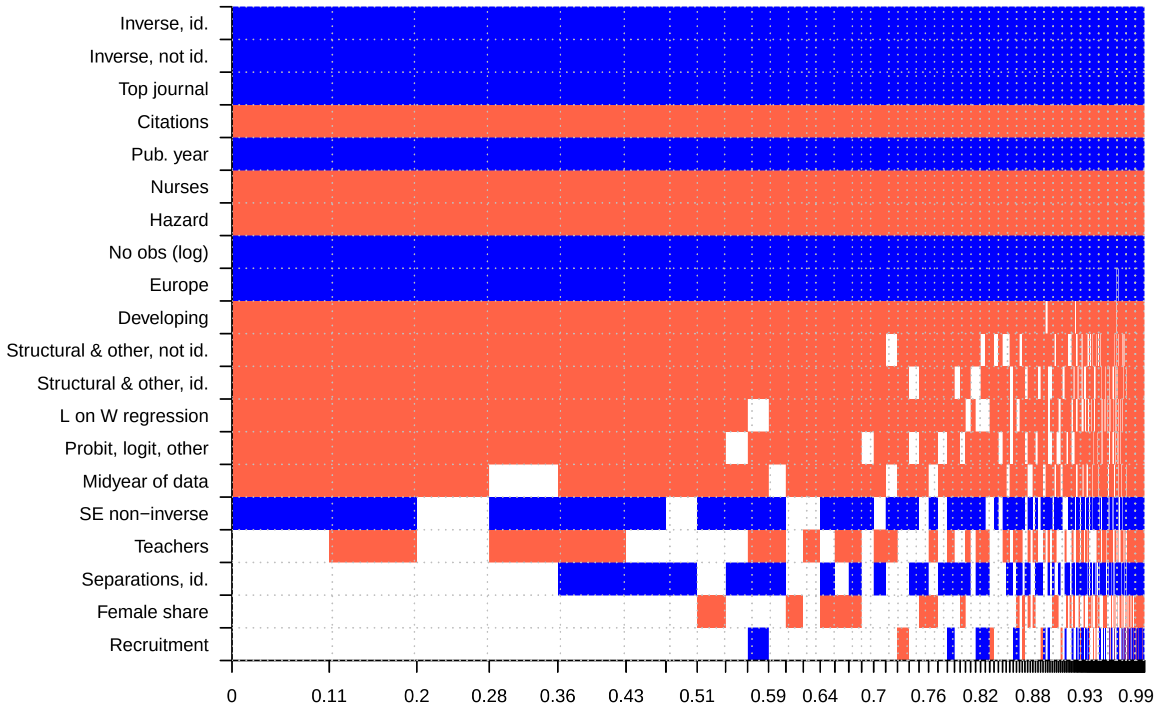
The task we put forward presents one serious challenge: it is unlikely that all of our 20 explanatory variables contribute to the data generating process of supply elasticity estimates in a meaningful way, so model (9) that contains all of these variables is likely misspecified. At the same time, we do not have any theoretical guidelines to help us eliminate the redundant variables *ex ante*. Although sequential t-testing is a popular choice in this context, we find it unsatisfactory: sequential elimination of insignificant regressors may lead us to accidentally exclude some of the variables that belong to the data generating process. To address this model uncertainty, we turn to Bayesian methods and perform Bayesian Model Averaging (BMA).

BMA presents an alternative and, arguably, better solution to the model uncertainty problem. Unlike sequential t-testing, BMA does not discard the information coming from controls that appear insignificant in the broad specification of model (9)—after all, the model with all 20 controls included is only one out of 2^{20} possible combinations of our chosen explanatory variables. Instead, BMA traverses through the space of all of the possible regression models and assigns each a metric called Posterior Model Probability (PMP) that reflects how well the model performs compared to all the others.

Inference in BMA is obtained by taking a weighted average of the results from all possible models, using the Posterior Model Probability (PMP) as a weight. It is worth noting that we do not estimate each of the 2^{20} regressions; instead, we employ a Model Composition Markov Chain Monte Carlo algorithm that visits models with the highest PMP and approximates the rest (see Madigan & York 1995). We implement this using the `BMS` package in `R` written by Zeugner & Feldkircher (2015). Our base specification uses a combination of uniform model prior and unit information prior for model parameters, following Eicher *et al.* (2011), but we also report results obtained under alternative priors. Detailed discussions of applications of BMA to economics can be found in Moral-Benito (2015) and Steel (2017); Koop (2003) provides an excellent technical description of the method. Another example of BMA application can be found in Fernández *et al.* (2001), who use it to combat model uncertainty in cross-country growth regressions. Havranek *et al.* (2017) use BMA in a context similar to ours, tackling model uncertainty in a meta-analysis of habit formation in consumption.

The results of the BMA estimation are reported in Figure 3. The explanatory variables shown on the left are sorted by Posterior Inclusion Probability (PIP). Each explanatory variable is present in $2^{20} - 2^{19}$ models; PIP gives the sum of posterior model probabilities of all models in which a regressor is included, assessing how likely it is that each explanatory variable belongs in the data generating process for elasticity estimates. The vertical axis of Figure 3 lists explanatory variables with the highest to lowest PIP. The horizontal axis depicts different models with highest to lowest Posterior Model Probability and plots cumulative PMP values. White color in Figure 3 indicates that the explanatory variable is not included in the selected model, blue (darker in greyscale) means that the variable is included with a positive coefficient, and red (lighter in grayscale) means that the variable is included and has a negative sign.

Figure 3: Model inclusion in Bayesian model averaging



Notes: The response variable is the estimate of the elasticity of labor supply to the firm. Each column denotes an individual model; variables are sorted in descending order by their posterior inclusion probability. The cumulative posterior model probabilities are given on the horizontal axis. Blue color (darker in greyscale) indicates that the variable is included in the model and that the estimated sign is positive. Red color (lighter in greyscale) indicates that the variable is included in the model and that the estimated sign is negative. A lack of color implies that the variable is not included in the model. The numerical results of the BMA estimation are reported in Table 3.

The signs of most explanatory variables are quite stable across models in which the variables are included. As expected, estimates converted from inverse elasticities are higher than those obtained using separations (our reference group). On the other hand, studies that directly regress labor supply on wage, or those that use structural models with production tend to come up with lower estimates. At the same time, there does not seem to be much difference between results obtained using separations and those based on recruitments, providing some evidence in support of the assumption that these two parameters are equal to one another in absolute value.

Studies that use data from developing countries seem to yield lower estimates compared to those looking at advanced economies. Among advanced economies, estimates from European data are higher when compared to the US, Canada and Australia (the geographic reference group). The use of non-linear estimation techniques (particularly hazard models, but also probit and logit) seem to result in lower elasticity estimates when compared with linear regression models. Studies that focus exclusively on the market for nurses seem to produce lower estimates of the supply elasticity, potentially indicating that this market is more monopsonistic. At the same time, the results for teachers are less stable, as most of the models that perform well do not include this explanatory variable.

Top journals seem to publish higher estimates of the supply elasticity; studies using large data sets also typically come up with higher estimates. In contrast, the most cited results are those reporting low estimates, perhaps reflecting the overall interest in evidence for monopsonistic labor markets. Our publication bias result reported in Section 4 remains intact: there is still correlation between reported estimates based on 'direct' methods and the corresponding standard error, even after we control for various aspects of study design.

There is no clear evidence of a trend in the evolution of the elasticity estimates over time: on the one hand, studies published more recently report higher estimates (other things equal); on the other hand, studies that use newer data sets typically report lower elasticity estimates, consistent with monopsony power having increased over time.

We also do not find strong evidence of the female share affecting the estimates; this, however, could be due to the fact that many studies do not report precise demographic data, and the indicator we constructed is only an approximation (see subsection 5.1 for details). The regression models in which the female share is present include it with a negative sign, which can be interpreted as some (weak) evidence of gender gaps and warrants further investigation in future research. We also observe that separation-based estimates obtained with identification strategies in place seem to be higher; however, the best models do not distinguish between whether there is an identification strategy in place for separations. Nevertheless, this result may be due to the fact that our sample only has 18 separations-based estimates that we classify as 'identified', so this particular test has low statistical power.

We present numerical results of BMA estimation in the left panel of Table 3, reporting the mean values of corresponding coefficients averaged across all models, their standard deviation and the values of posterior inclusion probabilities. Variables with PIP that exceeds 0.5 belong to the data generating process with probability of more than 50%—this can be thought of as the analogue of significance in frequentist econometrics. The right panel of Table 3 reports a robustness check of these results in which we run an OLS for variables that have PIP higher than 50%. As before, we cluster standard errors at the study level and additionally compute p-values using wild bootstrap clustering.

Table 3: Why do estimates of supply elasticity vary?

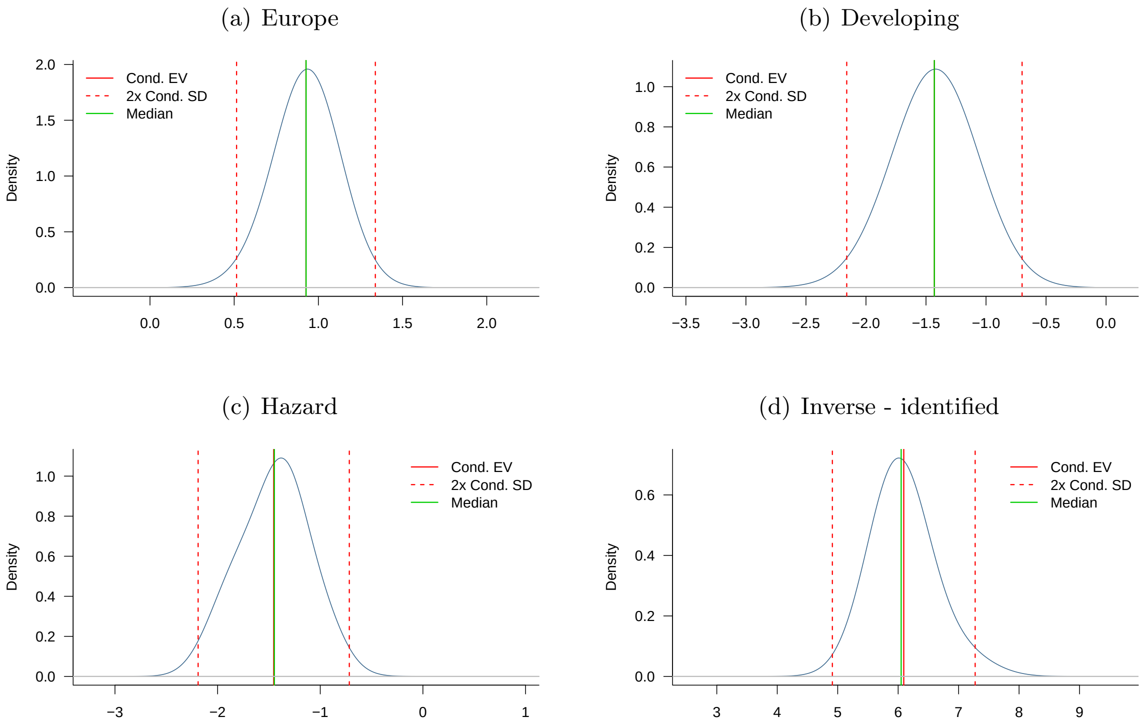
Response variable:	Bayesian model averaging			Frequentist check (OLS)			
Estimate of supply elasticity	Post. mean	Post. SD	PIP	Coef.	Std. er.	p-value	p-value (wild)
Data Characteristics							
SE non-inverse	0.157	0.108	0.759	0.173	0.051	0.001	0.154
No obs (log)	0.136	0.034	0.997	0.125	0.055	0.023	0.070
Midyear of data	-0.014	0.008	0.849	-0.015	0.006	0.010	0.440
Female share	-0.075	0.191	0.173				
Country & Industry							
Developing	-1.418	0.387	0.992	-1.267	0.559	0.023	0.254
Europe	0.923	0.212	0.997	0.899	0.358	0.012	0.054
Nurses	-2.463	0.633	1.000	-2.125	1.097	0.053	0.176
Teachers	-0.463	0.574	0.482				
Method & Identification Strategy							
Separations, id.	0.566	0.757	0.430				
Inverse, id.	6.093	0.591	1.000	5.906	1.534	0.000	0.065
Inverse, not id.	12.767	0.530	1.000	13.007	2.433	0.000	0.126
Recruitment	0.081	0.374	0.122				
L on W regression	-1.657	0.688	0.916	-1.967	0.852	0.021	0.477
Structural & other, id.	-1.986	0.816	0.929	-2.011	0.715	0.005	0.696
Structural & other, not id.	-1.131	0.461	0.938	-1.151	0.411	0.005	0.072
Estimation Technique							
Hazard	-1.452	0.371	0.999	-1.314	0.323	0.000	0.002
Probit, logit, other	-0.931	0.479	0.890	-0.844	0.442	0.056	0.155
Publication Characteristics							
Top journal	2.106	0.343	1.000	2.078	0.684	0.002	0.047
Citations	-2.641	0.491	1.000	-2.408	0.700	0.001	0.028
Pub. year	0.139	0.015	1.000	0.138	0.075	0.064	0.256
Intercept	-2.153		1.000	-2.142	2.568	0.404	0.717
Observations	787			787			

Notes: PIP denotes posterior inclusion probability; SD is the standard deviation; 'id' denotes estimates obtained with an identification strategy in place. The left panel of the table presents unconditional moments for the BMA. The right panel reports the result of the frequentist check in which we include only explanatory variables with PIP > 0.5. The standard errors in the frequentist check are clustered at the study level. 'p-value (wild)' are wild bootstrap clustered p-values. A detailed description of all variables is available in Table A1.

According to the evidence presented in Table 3, estimates converted from the inverse elasticity differ from those obtained via direct methods, and the difference is both statistically and economically significant. This gap, though still important, becomes two times smaller if the study employs an identification strategy. This result is in line with Manning (2003), who argues that this stock-based estimation may overstate the supply elasticity estimate due to, for example, unobserved supply shocks. We also find that stock-based estimates obtained from regressing labor supply on wage are lower than those obtained using separations, again in line with conjectures put forward in Manning (2003). On the other hand, there does not seem to be any significant difference between estimates obtained by using separation elasticities and those based on recruitments. As discussed in Section 2, researchers that estimate supply elasticity based on separations or recruitments have to assume steady-state equivalence between the two rates. Our results show no significant difference between the two sets of estimates and therefore speak in favor of these assumptions; they are also in line with Falch (2010) and Falch (2017), who obtains similar estimates from a recruitment and separations based approach.

The use of non-linear estimation methods (hazard, probit or logit) results in elasticities that are smaller by about 1.3 for hazards and 0.8 for probits and logits compared to those obtained using linear methods. The estimates also differ across countries: they are higher for Europe by about 0.9, and lower for developing countries by 1.3—compared to the reference group of the US, Canada and Australia. Studies that focus exclusively on the market for nurses appear to come up with stronger evidence for monopsonistic behavior compared to all other studies. Finally, estimates of the supply elasticity reported in studies published in top journals are higher by about 2.1 when compared to those reported in all other journals, an effect that, again, is both economically and statistically significant. The effect associated with publication bias is also present and shows some evidence of statistical significance. However, it is smaller in magnitude, and has a modest effect on the overall results.

Figure 4: Posterior density of selected parameter estimates



Notes: The figure displays densities of the coefficients on *Europe, Developing, Hazard* and *Inverse-ID*, calculated from all models in which the respective variable was included. Accordingly, unlike the results reported in Table 3, these are conditional moments. The posterior means of parameter estimates are more than two standard deviations away from zero, analogous to a statistical significance at the 5% level.

Figure 4 depicts the posterior densities for variables *Europe*, *Developing*, *Hazard* and *Inverse - id*. Unlike the results of Table 3, these values are conditional on variable inclusion (i.e. averaged across models in which the variable is included rather than across all models). We see that the means of all effects lie at least two standard deviations away from zero, which can be interpreted analogously to a test of statistical significance in frequentist econometrics.

Figure B1 reports coefficients and posterior inclusion probabilities under alternative prior settings; our results appear resilient to assumptions about priors. In another robustness check, Table B2 reports results for simple OLS in which we include all regressors. Even though this model is likely misspecified, the estimated magnitudes and signs of all the effects are very similar to those produced in the BMA exercise. Furthermore, we repeat our BMA estimation in a setting where we weigh each data point by the inverse of the number of estimates reported per study. This specification gives equal weight to studies that report many estimates and those reporting only a few. Figure B2 and Table B3 document the results; they appear broadly consistent with our baseline estimation. The publication bias result appears more pronounced, and the identification strategy for the separation-based approach becomes important. At the same time, results for stock-based 'direct' estimation, the use of structural methods, logit and probit become insignificant (while retaining their negative signs). The rest of the estimated effects are similar to those obtained with an unweighted specification.

Finally, we check the robustness of our treatment of outliers. We repeat the BMA estimation for our original sample of 801 estimates where outliers from each tail are winsorized at 1% and report the results in Figure B3 and Table B4. The drawback of the winsorization is that it suppresses variation within the outlier group of inverse estimates; we therefore observe less systematic differences between inverse estimates obtained with and without an identification strategy. Overall, this specification is consistent with our key results. A notable exception is that it yields some evidence suggesting that estimates based on separations are higher for studies that employ an identification strategy, compared to those that do not. Estimates based on recruitment data seem to also be larger than this latter group, and of roughly similar magnitude compared to 'identified' estimates based on separations. This may seem at odds with our previous conclusion of no systematic difference between estimates based on recruitments and separations. Note, however, that in our sample the majority of estimates obtained from recruitment data come from studies that employ an identification strategy; we can therefore conclude that, conditional on employing an identification strategy, there still seems to be little difference between estimates obtained from data on recruitments and separations.

To understand what these different methods imply about firm wage-setting power, we now compare fitted values of supply elasticity estimates across different methods and features of the data. For the final estimates to be useful to the reader, we construct what we believe are estimates associated with 'best practices' in the literature, rather than just focusing on sample means. This is done by substituting high parameter values for variables that, we believe, reflect best practice; low parameter values for those that do not; and putting sample means for cases where we are indifferent. For example, we correct for publication bias by substituting zero instead of the mean for the standard error on non-inverse estimates; we also believe that results from studies that use large data sets are probably more reliable—we therefore put the value of the 90th percentile for the *number of observations*; we also think that our readers are probably more interested in estimates that are more current (both recently published and using modern data)—we put 90th percentile values for *publication year* and *midyear of data*. We would like to rely on estimates that other economists trust as well, we therefore set the value of *top journal* to one and use the value of 90th percentile for the *number of citations*.[13]

In the top panel of Table 4, we present best practice estimates obtained from a separation elasticity based strategy, by far the most common strategy employed in studies that we examine. Three important findings are clear in the table. Most importantly, we observe small estimates, which are much more consistent with a monopsonistic labor market than they are with a perfectly competitive labor market (which requires the elasticity of labor supply to be infinite). The point estimates for both the linear and hazard approaches imply that firms are able to pay the last worker hired between 18 and 25 percent less than his or her marginal revenue product. Even the largest estimate contained in one of our 95% confidence intervals, 7.57, implies firm markdown power of nearly 12%.

The top panel of Table 4 also reveals two important dimensions of heterogeneity that we have discussed above. First, linear specifications yield estimates that are higher than those obtained from a hazard model. Second, it shows evidence of cross-country variation. Overall, we see the largest estimates coming from Europe, followed by other advanced economies, then developing countries. This suggests that important institutional factors, which vary across labor markets, may have non-negligible impacts on firm wage-setting power.

The bottom half of the table presents the best practice estimates for studies that use the inverse elasticity strategy. As discussed above, these estimates are generally higher. As endogeneity seems to be a significant concern for these estimates, we focus only on estimates obtained from the models with identification strategies. Overall, we find a point estimate slightly larger than 10, implying wage markdown of about 9%. The regional variation in the elasticity is necessarily the same as for the separations based approach, given our linear model from which we arrive at these best practice estimates. As the inverse-based estimates are generally larger, and wage-setting power is a non-linear function of the elasticity, the bottom half of the table suggests less regional variation in markdowns, varying between European and developing labor markets by 2 percentage points rather than 19 percentage points. Overall, these estimates again do provide strong evidence of firms possessing some monopsony power; the largest estimate in our 95% confidence interval here, 17.74, still suggests that firms have the power to pay workers about 5% less than they are worth. While a much smaller amount of wage-setting power than suggested by the results above, this still represents an economically significant departure from the perfectly competitive model.

The best practice exercise reported in Table 4 is based on a model that only includes variables that proved significant in our baseline BMA estimation. As another robustness check, we recalculate our best practice estimates using the OLS model shown in Table B2, which includes all 20 controls. Best practice values that we obtain are very similar to those reported in Table 4,

[13]We also compute alternative best practice estimates, using values from the 75th and 95th percentiles for *number of observations*, *publication year*, *midyear of data* and *number of citations*. We report the results in Table B5; they are in line with estimates reported here.

though the confidence intervals are wider on account of over-controlling. Nevertheless, this exercise provides interesting insights into how the use of an identification strategy may affect the results: the gap between estimates based on separations and inverse elasticities shrinks when both methods employ an identification strategy (see Table B6).

Table 4: Best practice

Group	Point Estimate	95% interval	95% interval (wild)	Implied Markdown
Separations				
Linear – all	4.350	[2.59; 6.11]	[1.79; 7.57]	18.7%
Hazard – all	3.036	[1.40; 4.67]	[0.49; 6.16]	24.8%
Hazard – developing	1.484	[0.01; 2.96]	[-0.63; 3.70]	40.3%
Hazard – Europe	3.651	[1.75; 5.56]	[0.53; 7.01]	21.5%
Hazard – other advanced	2.752	[1.21; 4.29]	[0.46; 5.88]	26.7%
Inverse				
Identified – all	10.256	[5.97; 14.54]	[2.46; 16.91]	8.9%
Identified – developing	8.705	[4.82; 12.59]	[2.33; 14.93]	10.3%
Identified – Europe	10.872	[6.42; 15.33]	[2.84; 17.74]	8.4%
Identified – other advanced	9.972	[5.73; 14.22]	[2.21; 16.60]	9.1%

Notes: The table presents fitted 'best practice' estimates for various estimation techniques and data. We used the model estimated as the frequentist check of our baseline specification reported in Table 3. We report both the standard 95% confidence interval calculated for errors clustered at the study level, and the 95% confidence interval calculated with wild bootstrap clusters. The estimates of the markdown are obtained using equation (2).

This evidence of firm monopsony power can be used to reconcile some empirical puzzles arising in the labor literature. For example, in two meta-analysis, Card & Krueger (1995b) and Doucouliagos & Stanley (2009) show that increases in the minimum wage do not depress employment, and in fact sometimes have a positive effect. This finding goes against the logic of the competitive labor market framework; it can, however, be explained through presence of monopsony. Manning (2003, pp. 345-347) uses a general equilibrium version of the monopsony model to generate responses in employment to changes in minimum wages. In his example, positive or negligibly small responses are generated under elasticities of 3.3 and 5, consistent with results we report in Table 4.

Dube *et al.* (2018b) argue that firm wage-setting power may explain bunching in wages at round numbers: the lower the elasticity of labor supply to the firm, the less costly the 'wrong' wage is in terms of turnover costs. For example, elasticities of labor supply to the firm of 1 and 5 would be associated with firms forsaking 1% or 10% of profits due to bunching, respectively. Our results are broadly consistent with these numbers.

Card *et al.* (2018a) provide micro-foundations for the static monopsony model, assuming heterogeneity in worker preferences across different work environments. This leads to workers distinguishing between different employers on the basis other than wage, and gives wage-setting power to the firms. The authors show that, under the assumption of a supply elasticity of 4 (and markdown of 20%) this model can be used to explain observed dispersion of wages, and their link to firm productivity.

6 Discussion and Conclusion

Imperfect competition among employers can lead to workers being payed less than their worth to the firm. Recently, academic research on such labor market structures has made its way into policy debate. At the end of the Obama administration, the Council of Economic Advisers issued a policy brief on monopsonistic labor markets and potential policy remedies (Council of Econoimcs Advisors 2016). The arguments of Krueger & Posner (2018) were put forward to a wider audience in a *New York Times* op-ed (Posner & Krueger 2018). In late 2017, Senators Cory Booker and Elizabeth Warren wrote an open letter to Attorney General Jeff Sessions, urging enforcement of recent Department of Justice guidance that no-poach agreements are likely illegal (Warren & Booker 2017; Booker 2017). This new found interest from policy makers calls for a detailed investigation of the existing quantitative evidence for monopsonistic labor markets.

Here, we attempt to synthesize empirical evidence on the elasticity of labor supply to the firm, a parameter that captures the extent of firms' wage-setting power. We show that features pertaining to study design, data, publication quality and researcher's implicit preference combine to explain the observed variation in estimates. We also provide quantitative predictions of what supply elasticity estimates should be for different estimation techniques, conditional on employing best research practices. Our results suggest that, overall, the literature provides strong evidence for monopsonistic competition and implies sizable markdowns in wages. That being said, several caveats are in order.

First, we do not claim to explain the systematic variation in the 'true' supply elasticity parameter. Instead, our empirical exercise approximates the data generating process for supply elasticity *estimates*, conditional on the existing literature. Some of the variation that we report is likely driven by differences in the underlying parameter value (e.g. estimates for different countries), whereas other variation may arise purely due to choices made by researchers (e.g. estimation technique). Second, our results provide evidence on the elasticity of labor supply to the firm and the implied degree of firms' wage-setting power; strictly speaking, they remain silent about whether the firms are able to exercise this power. This latter question is outside the scope of our analysis, because it is outside the scope of the literature that we base our inference on. Given this concern, our results regarding implied salary markdowns can be viewed as a prediction of what these markdowns would be assuming that firms fully exploit the power they have over workers.

Finally, our inference is based upon evidence coming from a set of studies that were published at the time that we ran our search; it therefore omits some of the recent unpublished work on the topic.[14] Given the rising interest toward the monopsony literature, we are certain that the number of studies offering estimates of supply elasticity will increase in the coming years. We leave this new evidence to future research.

[14] For example, the works of Crawford *et al.* (2015), Hirsch *et al.* (2018a), Félix & Portugal (2016), Bellon (2016), Tucker (2017), Garin & Silverio (2017), Dube *et al.* (Forthcoming) and Webber (2018) were not published at the time of our search.

References

Acemoglu, D. & J.-S. Pischke (1999): "Beyond Becker: Training in imperfect labour markets." *The economic journal* **109(453)**: pp. 112–142.

Ashenfelter, O., C. Harmon, & H. Oosterbeek (1999): "A review of estimates of the schooling/earnings relationship, with tests for publication bias." *Labour Economics* **6(4)**: pp. 453–470.

Azar, J., I. Marinescu, & M. I. Steinbaum (2017): "Labor market concentration." *Working Paper 24147*, National Bureau of Economic Research.

Balasubramanian, N., J. W. Chang, M. Sakakibara, J. Sivadasan, & E. Starr (2018): "Locked in? The enforceability of covenants not to compete and the careers of high-tech workers." *Working Paper No. CES-WP-17-09*, US Census Bureau Center for Economic Studies.

Bellon, M. (2016): "The characteristics of worker flows by firm growth: Empirical evidence from a matched firm-worker dataset from france." *Working Paper.*

Benmelech, E., N. Bergman, & H. Kim (2018): "Strong employers and weak employees: How does employer concentration affect wages?" *Working Paper 24307*, National Bureau of Economic Research.

Berry, S., J. Levinsohn, & A. Pakes (1995): "Automobile prices in market equilibrium." *Econometrica* **63(4)**: pp. 841–890.

Boal, W. M. & M. R. Ransom (1997): "Monopsony in the labor market." *Journal of Economic Literature* **35(1)**: pp. 86–112.

Bodah, M., J. Burkett, & L. Lardaro (2003): "Ix. employment relations for health care workers." In "Proceedings of the Annual Meeting–Industrial Relations Research Association," p. 199. IRRA.

Booker (2017): "Booker, Warren sound alarm on collusive 'no-poach' agreements." *Senate Press Release* November 21, 2017.

Booth, A. L. & P. Katic (2011): "Estimating the wage elasticity of labour supply to a firm: What evidence is there for monopsony?" *Economic Record* **87(278)**: pp. 359–369.

Brummund, P. (2011): "Variation in monopsonistic behavior across establishments: Evidence from the Indonesian labor market." *Working Paper.*

Burdett, K. & D. T. Mortensen (1998): "Wage differentials, employer size, and unemployment." *International Economic Review* **39(2)**: pp. 257–273.

Cameron, A. C., J. B. Gelbach, & D. L. Miller (2008): "Bootstrap-based improvements for inference with clustered errors." *The Review of Economics and Statistics* **90(3)**: pp. 414–427.

Card, D., A. R. Cardoso, J. Heining, & P. Kline (2018a): "Firms and labor market inequality: Evidence and some theory." *Journal of Labor Economics* **36(S1)**: pp. S13–S70.

Card, D., J. Kluve, & A. Weber (2018b): "What works? A meta analysis of recent active labor market program evaluations." *Journal of the European Economic Association* **(forthcoming)**.

Card, D. & A. 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)**: pp. 772–793.

Card, D. & A. B. Krueger (1995a): *Myth and measurement.* Princeton University Press Princeton, NJ.

Card, D. & A. B. Krueger (1995b): "Time-series minimum-wage studies: A meta-analysis." *American Economic Review* **85(2)**: pp. 238–43.

Chetty, R., A. Guren, D. Manoli, & A. Weber (2013): "Does indivisible labor explain the difference between micro and macro elasticities? A meta-analysis of extensive margin elasticities." *NBER Macroeconomics Annual* **27(1)**: pp. 1–56.

Council of Econoimcs Advisors (2016): "Labor market monopsony: Trends, consequences, and policy responses." *Issue Brief* obamawhitehouse.archives.gov.

Crawford, R., R. Disney, & C. Emmerson (2015): "The short run elasticity of National Health Service nurses? Labour supply in Great Britain." *IFS Working Papers W15/04*, Institute for Fiscal Studies.

Dal Bó, E., F. Finan, & M. A. Rossi (2013): "Strengthening state capabilities: The role of financial incentives in the call to public service." *The Quarterly Journal of Economics* **128(3)**: pp. 1169–1218.

Depew, B. & T. A. Sørensen (2013): "The elasticity of labor supply to the firm over the business cycle." *Labour Economics* **24**: pp. 196–204.

Dobbelaere, S. & J. Mairesse (2013): "Panel data estimates of the production function and product and labor market imperfections." *Journal of Applied Econometrics* **28(1)**: pp. 1–46.

Doucouliagos, H. & T. D. Stanley (2009): "Publication selection bias in minimum-wage research? A meta-regression analysis." *British Journal of Industrial Relations* **47(2)**: pp. 406–428.

DUBE, A., L. GIULIANO, & J. LEONARD (Forthcoming): "Fairness and frictions: The impact of unequal raises on quit behavior." *American Economic Review* .

DUBE, A., J. JACOBS, S. NAIDU, & S. SURI (2018a): "Monopsony in online labor markets." *American Economic Review: Insights* **(forthcoming)**.

DUBE, A., A. MANNING, & S. NAIDU (2018b): "Monopsony and employer mis-optimization explain why wages bunch at round numbers." *Working Paper 24991*, National Bureau of Economic Research.

EFENDIC, A., G. PUGH, & N. ADNETT (2011): "Institutions and economic performance: A meta-regression analysis." *European Journal of Political Economy* **27(3)**: pp. 586–599.

EGGER, M., G. D. SMITH, M. SCHEIDER, & C. MINDER (1997): "Bias in meta-analysis detected by a simple, graphical test." *British Medical Journal* **316**: pp. 629–634.

EICHER, T. S., C. PAPAGEORGIOU, & A. E. RAFTERY (2011): "Default priors and predictive performance in bayesian model averaging, with application to growth determinants." *Journal of Applied Econometrics* **26(1)**: pp. 30–55.

FAKHFAKH, F. & F. FITZROY (2006): "Dynamic monopsony: Evidence from a French establishment panel." *Economica* **73(291)**: pp. 533–545.

FALCH, T. (2010): "The elasticity of labor supply at the establishment level." *Journal of Labor Economics* **28(2)**: pp. 237–266.

FALCH, T. (2017): "Wages and recruitment: evidence from external wage changes." *ILR Review* **70(2)**: pp. 483–518.

FELDKIRCHER, M. (2012): "Forecast combination and bayesian model averaging: A prior sensitivity analysis." *Journal of Forecasting* **31(4)**: pp. 361–376.

FELDKIRCHER, M. & S. ZEUGNER (2012): "The impact of data revisions on the robustness of growth determinants—a note on determinants of economic growth: Will data tell?" *Journal of Applied Econometrics* **27(4)**: pp. 686–694.

FÉLIX, S. & P. PORTUGAL (2016): "Labor market imperfections and the firm's wage setting policy." *IZA Discussion Papers 10241*, Institute for the Study of Labor (IZA).

FERNANDEZ, C., E. LEY, & M. F. J. STEEL (2001): "Benchmark priors for bayesian model averaging." *Journal of Econometrics* **100(2)**: pp. 381–427.

FERNÁNDEZ, C., E. LEY, & M. F. J. STEEL (2001): "Model uncertainty in cross-country growth regressions." *Journal of Applied Econometrics* **16(5)**: pp. 563–576.

FLEISHER, B. M. & X. WANG (2004): "Skill differentials, return to schooling, and market segmentation in a transition economy: the case of mainland China." *Journal of Development Economics* **73(1)**: pp. 315–328.

GALIZZI, M. (2001): "Gender and labor attachment: Do within-firms' relative wages matter?" *Industrial Relations: A Journal of Economy and Society* **40(4)**: pp. 591–619.

GARIN, A. & F. SILVERIO (2017): "How does firm performance affect wages? Evidence from idiosyncratic export shocks." *2017 Papers pga940*, Job Market Papers.

HAVRANEK, T. (2010): "Rose effect and the Euro: Is the magic gone?" *Review of World Economics* **146(2)**: pp. 241–261.

HAVRANEK, T. (2015): "Measuring intertemporal substitution: The importance of method choices and selective reporting." *Journal of the European Economic Association* **13(6)**: pp. 1180–1204.

HAVRANEK, T., M. RUSNAK, & A. SOKOLOVA (2017): "Habit formation in consumption: A meta-analysis." *European Economic Review* **95**: pp. 142 – 167.

HAVRANEK, T. & A. SOKOLOVA (2018): "How does consumption respond to expected income changes? Evidence from 144 studies." Czech National Bank and Charles University, Prague.

HIRSCH, B., E. JAHN, A. MANNING, & M. OBERFICHTNER (2018a): "The urban wage premium in imperfect labour markets." *Working Paper*.

HIRSCH, B. & E. J. JAHN (2015): "Is there monopsonistic discrimination against immigrants?" *ILR Review* **68(3)**: pp. 501–528.

HIRSCH, B., E. J. JAHN, & C. SCHNABEL (2018b): "Do employers have more monopsony power in slack labor markets?" *ILR Review* **71(3)**: pp. 676–704.

HIRSCH, B., T. SCHANK, & C. SCHNABEL (2010): "Differences in labor supply to monopsonistic firms and the gender pay gap: An empirical analysis using linked employer-employee data from Germany." *Journal of Labor Economics* **28(2)**: pp. 291–330.

KOOP, G. (2003): *Bayesian Econometrics*. John Wiley & Sons.

KRUEGER, A. B. & O. ASHENFELTER (2018): "Theory and evidence on employer collusion in the franchise sector." *Working Paper 24831*, National Bureau of Economic Research.

KRUEGER, A. B. & E. A. POSNER (2018): "A proposal for protecting low-income workers from monopsony and collusion." *Policy Proposal 2018-05*, The Hamilton Project.

Krugman, P. (1980): "Scale economies, product differentiation, and the pattern of trade." *The American Economic Review* **70(5)**: pp. 950–959.

Lewis, J. B. & D. A. Linzer (2005): "Estimating regression models in which the dependent variable is based on estimates." *Political Analysis* **13(4)**: pp. 345–364.

Ley, E. & M. F. Steel (2009): "On the effect of prior assumptions in Bayesian model averaging with applications to growth regressions." *Journal of Applied Econometrics* **24(4)**: pp. 651–674.

Madigan, D. & J. York (1995): "Bayesian graphical models for discrete data." *International Statistical Review* **63(2)**: pp. 215–232.

Manning, A. (2003): *Monopsony in motion: Imperfect competition in labor markets*. Princeton University Press.

Manning, A. (2011): "Imperfect competition in the labor market." In "Handbook of Labor Economics, Volume 4B," chapter 11, pp. 973–1035. North Holland.

Matsudaira, J. D. (2014): "Monopsony in the low-wage labor market? Evidence from minimum nurse staffing regulations." *Review of Economics and Statistics* **96(1)**: pp. 92–102.

Melitz, M. J. (2003): "The impact of trade on intra-industry reallocations and aggregate industry productivity." *Econometrica* **71(6)**: pp. 1695–1725.

Moral-Benito, E. (2015): "Model averaging in economics: An overview." *Journal of Economic Surveys* **29(1)**: pp. 46–75.

Naidu, S. (2010): "Recruitment restrictions and labor markets: Evidence from the postbellum US south." *Journal of Labor Economics* **28(2)**: pp. 413–445.

Naidu, S., Y. Nyarko, & S.-Y. Wang (2016): "Monopsony power in migrant labor markets: evidence from the United Arab Emirates." *Journal of Political Economy* **124(6)**: pp. 1735–1792.

Naidu, S. & E. Posner (2018): "Labor market monopsony and the limits of the law." *Mimeo*.

Naidu, S., E. A. Posner, & E. G. Weyl (2018): "Antitrust remedies for labor market power." *Harvard Law Review* **(forthcoming)**.

Naidu, S. & N. Yuchtman (2013): "Coercive contract enforcement: Law and the labor market in nineteenth century industrial Britain." *American Economic Review* **103(1)**: pp. 107–44.

Ogloblin, C. & G. Brock (2005): "Wage determination in urban russia: Underpayment and the gender differential." *Economic Systems* **29(3)**: pp. 325–343.

Posner, E. & A. B. Krueger (2018): "Corporate america is suppressing wages for many workers." *op-ed*, New York Times, February 28, 2018.

Ransom, M. R. & R. L. Oaxaca (2010): "New market power models and sex differences in pay." *Journal of Labor Economics* **28(2)**: pp. 267–289.

Ransom, M. R. & D. P. Sims (2010): "Estimating the firm's labor supply curve in a "new monopsony" framework: Schoolteachers in Missouri." *Journal of Labor Economics* **28(2)**: pp. 331–355.

Ransom, T. (2018): "Labor market frictions and moving costs of the employed and unemployed." *Working Paper*.

Rinz, K. *et al.* (2018): "Labor market concentration, earnings inequality, and earnings mobility." *Working paper*, Center for Economic Studies, US Census Bureau.

Robinson, J. (1933): *The economics of imperfect competition*. MacMillin.

Roodman, D. (2018): "Boottest: Stata module to provide fast execution of the wild bootstrap with null imposed."

Rose, A. K. & T. D. Stanley (2005): "A meta-analysis of the effect of common currencies on international trade." *Journal of Economic Surveys* **19(3)**: pp. 347–365.

Rusnak, M., T. Havranek, & R. Horvath (2013): "How to solve the price puzzle? A meta-analysis." *Journal of Money, Credit and Banking* **45(1)**: pp. 37–70.

Staiger, D. O., J. Spetz, & C. S. Phibbs (2010): "Is there monopsony in the labor market? evidence from a natural experiment." *Journal of Labor Economics* **28(2)**: pp. 211–236.

Stanley, T. D. (2001): "Wheat from chaff: Meta-analysis as quantitative literature review." *Journal of Economic Perspectives* **15(3)**: pp. 131–150.

Stanley, T. D. (2005): "Beyond publication bias." *Journal of Economic Surveys* **19(3)**: pp. 309–345.

Stanley, T. D. (2008): "Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection." *Oxford Bulletin of Economics and Statistics* **70(1)**: pp. 103–127.

Stanley, T. D. & H. Doucouliagos (2015): "Neither fixed nor random: Weighted least squares meta-analysis." *Statistics in Medicine* **34(13)**: pp. 2116–27.

Steel, M. F. J. (2017): "Model Averaging and its Use in Economics." *MPRA Paper 81568*, University Library of Munich, Germany.

Sulis, G. (2011): "What can monopsony explain of the gender wage differential in Italy?" *International Journal of Manpower* **32(4)**: pp. 446–470.

Tucker, L. (2017): "Monopsony for whom? Evidence from Brazilian administrative data." *Working Paper*.

Warren, E. & C. A. Booker (2017): "Open letter to Attorney General Jeff Sessions." November 2017.

Webber, D. A. (2015): "Firm market power and the earnings distribution." *Labour Economics* **35**: pp. 123–134.

Webber, D. A. (2018): "Employment adjustment over the business cycle: The impact of competition in the labor market." *Working Paper 11887*, Institute for the Study of Labor (IZA).

Zeugner, S. & M. Feldkircher (2015): "Bayesian model averaging employing fixed and flexible priors – the BMS package for R." *Journal of Statistical Software* **68(4)**: pp. 1–37.

Appendix A: Description of Variables

Table A1: Definitions and summary statistics of explanatory variables

Variable	Description	Mean (all)	SD (all)	Mean (98%)	SD (98%)
Data characteristics					
SE non-inverse	An interaction between standard error and a dummy for whether the estimate is obtained through 'direct' (not inverse) estimation.	0.45	1.08	0.45	1.08
No obs (log)	The logarithm of the number of observations.	9.91	3.42	9.91	3.43
Midyear of data	The average year of the data used minus 1919 (the earliest midyear in the sample).	70.86	20.33	70.81	20.48
Female share	The share of female workers in the study's data set; 0.5 if sample stats not reported.	0.47	0.28	0.47	0.28
*Country & industry**					
Developing	=1 for data coming from countries classified as 'Emerging and Developing economies' by IMF classification in 2018.	0.07	0.26	0.07	0.26
Europe	=1 for data coming from countries in Europe.	0.42	0.49	0.42	0.49
Nurses	=1 for data that exclusively covers the market of medical workers.	0.10	0.30	0.10	0.29
Teachers	=1 for data that exclusively covers the market of teachers.	0.13	0.33	0.13	0.34
*[Reference category for COUNTRY: Other advanced economies.] [Reference category for INDUSTRY: estimates that do not exclusively relate to either market of nurses or teachers]					
*Method & identification***					
Separations, id.	=1 if estimate is based on separation rate AND is obtained through either IV or randomized identification strategy.	0.02	0.15	0.02	0.15
Inverse, id.	=1 if estimate converted from inverse elasticity AND is obtained through IV identification strategy.	0.09	0.29	0.08	0.28
Inverse, not id.	=1 if estimate converted from inverse elasticity AND the authors do not use IV.	0.03	0.18	0.03	0.17
Recruitment	=1 if estimate is based on recruitment rate.	0.08	0.27	0.08	0.27
L on W regression	=1 if estimate is obtained via stock-based estimation through regressing labor on wage.	0.06	0.24	0.06	0.24
Structural & other, id.	=1 if estimated obtained from structural model with production, or any other method not based on separations and not covered by specification controls above AND is obtained through either IV or randomized identification strategy.	0.01	0.12	0.01	0.12
Structural & other, not id.	=1 if estimated obtained from structural model with production, or any other method not based on separations and not covered by specification controls above AND the authors do not use IV or randomize.	0.10	0.31	0.11	0.31
**[Reference category: estimates based on separations AND not identified (no IV or randomized identification)]					
*Estimation technique****					
Hazard	=1 if study uses hazard model (reference category: linear techniques).	0.32	0.47	0.32	0.47
Probit, logit, other	=1 if study uses probit, logit or any other non-linear technique not previously classified (reference category: linear techniques).	0.22	0.42	0.22	0.42
***[Reference category: estimates based on linear techniques]					

Continued on next page

Table A1: Definitions and summary statistics of explanatory variables (continued)

Variable	Description	Mean (all)	SD (all)	Mean (98%)	SD (98%)
Publication characteristics					
Top journal	=1 if the study was published in one of the top five general interest journals in economics or the top field journal in labor.	0.12	0.32	0.12	0.32
Citations	The logarithm of the number of per-year citations of the study in Google Scholar (data for August 2018).	0.44	0.24	0.44	0.24
Pub. year	The year of publication of the study minus 1977, the year when the first study in our sample was published.	34.06	8.28	34.13	8.25

Notes: Data was collected from published studies estimating the elasticity of labor supply to the firm. When indicator variables form groups, we state the reference category. We report means and standard deviations for the full sample of 801 observations, as well as for the truncated subsample of 787 estimates.

Appendix B: Robustness checks

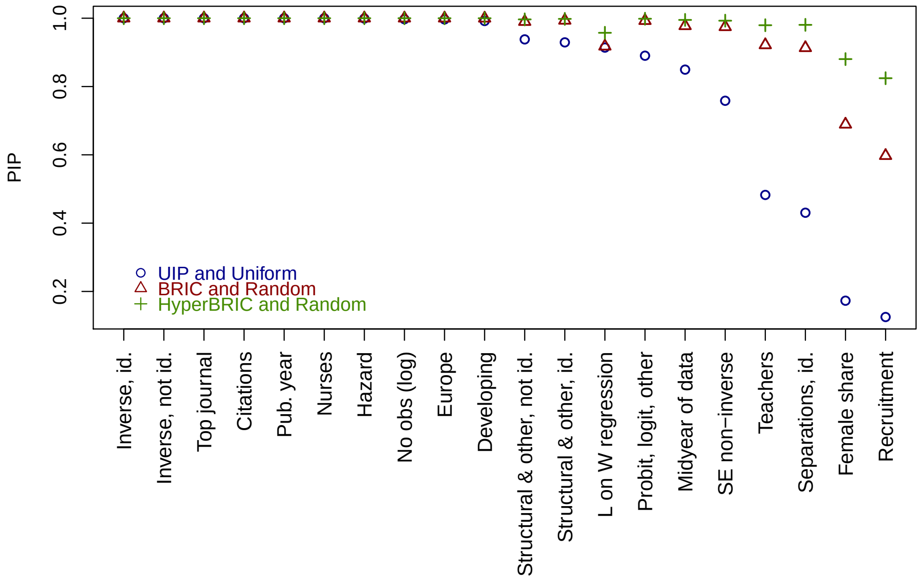
Table B1: Testing for publication bias

	OLS	FE	BE	Precision	Study	IV
SE	0.423	0.118	4.016	1.996	0.971	2.674
	(0.001)	(0.000)	(0.000)	(0.000)	(0.048)	(0.165)
	[0.123]	[0.313]	[0.000]	[0.016]	[0.000]	[0.395]
Constant	1.255	1.409	0.481	0.459	1.246	0.116
	(0.000)	(0.000)	(0.468)	(0.002)	(0.000)	(0.870)
	[0.000]	.	[0.290]	[0.091]	[0.000]	[0.848]
Studies	33	33	33	33	33	33
Observations	699	699	33	699	699	699

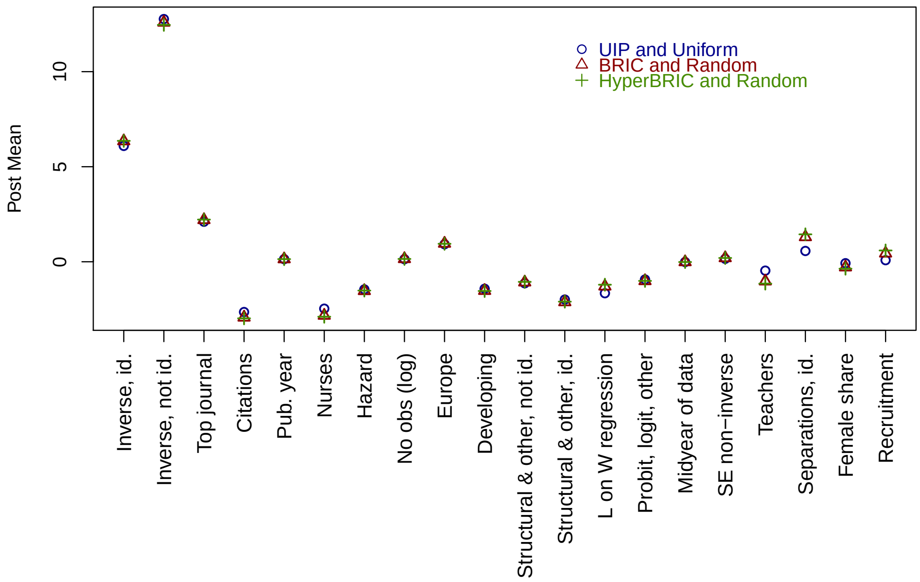
Notes: The table presents results from the following regression: $\hat{\epsilon}_{ij} = \epsilon_0 + \lambda \cdot SE(\hat{\epsilon}_{ij}) + u_{ij}$, where $\hat{\epsilon}_{ij}$ is the i-th estimate from the j-th study, $SE(\hat{\epsilon}_{ij})$ is the standard error of the estimate, and u_{ij} captures the unobservables in the regression. Standard errors from the regression are clustered at the study level and p-values are shown in parenthesis. We report p-values from wild bootstrap clustering in square brackets. This is implemented via the `boottest` command in `Stata` (see Roodman 2018). We use Rademacher weights and 9999 replications. The package does not allow for computation of a bootstrapped p-value for the constant term in the fixed effects specification. 'OLS' denotes ordinary least squares, 'FE' is study-level fixed effects, 'BE' is study-level between effects, 'Precision' is a specification with precision weights, and 'Study' is a specification with weights based on the inverse number of estimates reported in the study, 'IV' denotes an instrumental variables estimation with the number of observations as the instrument.

Figure B1: Robustness checks: alternative priors

(a) Posterior inclusion probabilities under alternative priors



(b) Posterior means under alternative priors



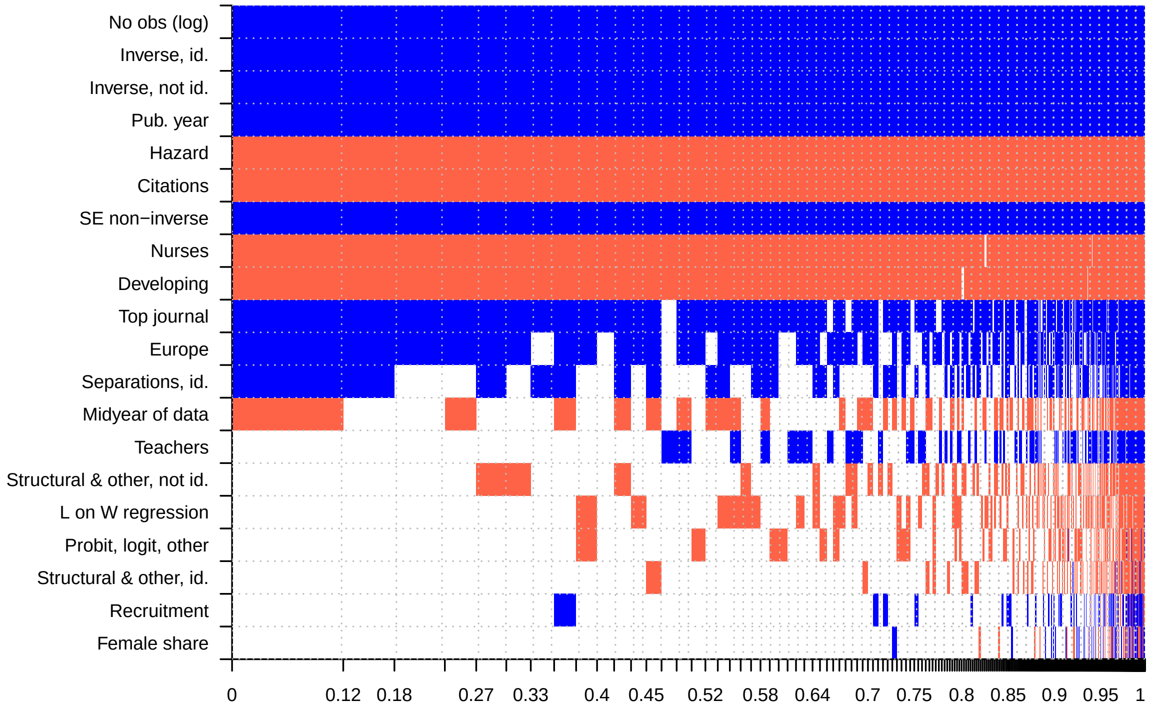
Notes: 'UIP and Uniform' denote unit information prior for parameters and the uniform model prior for model space; these are the priors that we use to obtain Figure 3 and Table 3. UIP is a data-dependent prior which conveys the amount of information equivalent to one observation. The uniform prior for model space effectively gives more weight to average model size. Eicher *et al.* (2011) demonstrates that these priors work well for predictive estimations. BRIC and Random denote a benchmark g-prior for paramters (that has been shown to have a very small effect on the posterior inference; see Fernandez *et al.* 2001) and the beta-binomial model prior for the model space (which gives equal weight to each model size; see Ley & Steel 2009). HyperBRIC indicates a data-dependent hyper-g prior for model parameters that should be more resilient to noise (see Feldkircher 2012 and Feldkircher & Zeugner 2012).

Table B2: Why do estimates of supply elasticity vary? OLS robustness check

Response variable:	OLS			
Estimate of supply elasticity	Coef.	Std. er.	p-value	p-value (wild)
Data Characteristics				
SE non-inverse	0.201	0.060	0.001	0.326
No obs (log)	0.151	0.067	0.024	0.137
Midyear of data	-0.014	0.006	0.014	0.420
Female share	-0.416	0.204	0.041	0.075
Country & Industry				
Developing	-1.567	0.598	0.009	0.196
Europe	0.946	0.466	0.042	0.248
Nurses	-2.968	1.161	0.011	0.139
Teachers	-1.239	0.557	0.026	0.080
Method & Identification				
Separations, id.	1.527	0.486	0.002	0.105
Inverse, id.	6.461	1.414	0.000	0.077
Inverse, not id.	12.581	2.499	0.000	0.153
Recruitment	0.716	0.781	0.359	0.464
L on W regression	-1.185	0.957	0.216	0.332
Structural & other, id.	-2.112	0.920	0.022	0.786
Structural & other, not id.	-1.058	0.456	0.020	0.128
Estimation Technique				
Hazard	-1.520	0.549	0.006	0.155
Probit, logit, other	-1.018	0.487	0.037	0.212
Publication Characteristics				
Top journal	2.271	0.766	0.003	0.040
Citations	-3.050	0.735	0.000	0.001
Pub. year	0.140	0.073	0.056	0.267
Intercept	-1.899	2.747	0.489	0.799
Observations	787			

Notes: 'id' denotes estimates obtained with an identification strategy in place. The standard errors are clustered at the study level. 'p-value (wild)' are wild bootstrap clustered p-values. A detailed description of all variables is available in Table A1.

Figure B2: Model inclusion in Bayesian model averaging: weighed by inverse of number of estimates



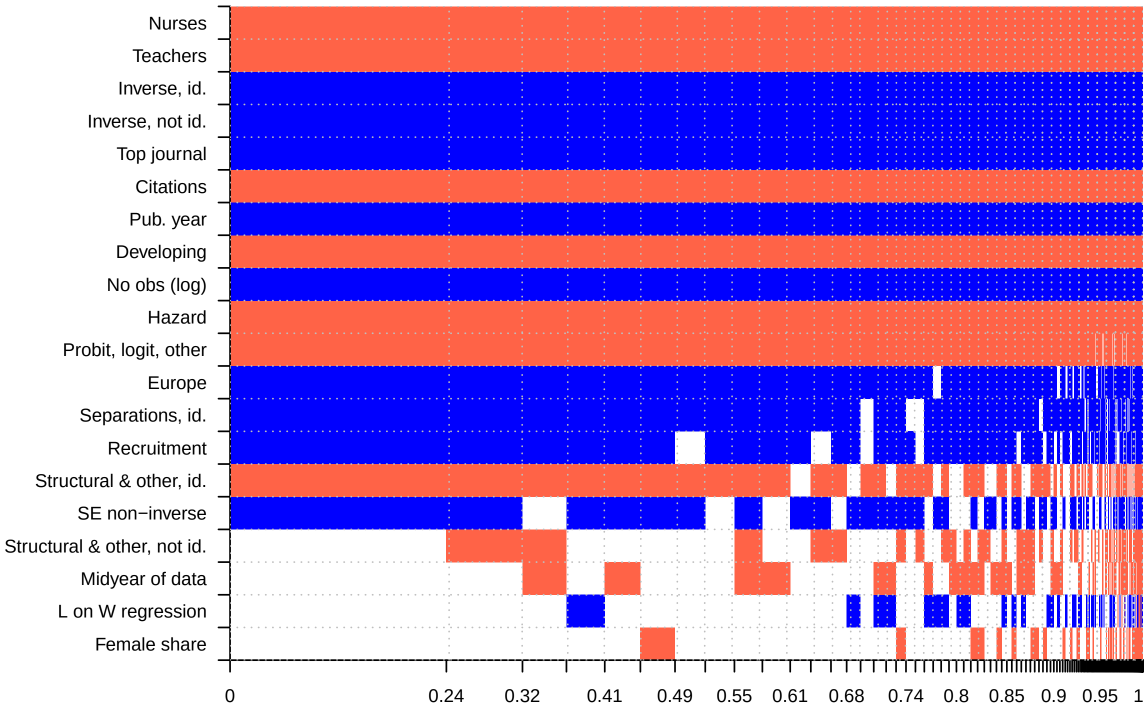
Notes: The response variable is the estimate of the elasticity of labor supply to the firm. Each column denotes an individual model; variables are sorted in descending order by their posterior inclusion probability. The cumulative posterior model probabilities are given on the horizontal axis. Blue color (darker in greyscale) indicates that the variable is included in the model and that the estimated sign is positive. Red color (lighter in greyscale) indicates that the variable is included in the model and that the estimated sign is negative. A lack of color implies that the variable is not included in the model. The numerical results of the BMA estimation are reported in Table B3.

Table B3: Why do estimates of supply elasticity vary? Weighed by inverse of number of estimates

Response variable:	Bayesian model averaging			Frequentist check (WLS)			
Estimate of supply elasticity	Post. mean	Post. SD	PIP	Coef.	Std. er.	p-value	p-value (wild)
Data Characteristics							
SE non-inverse	0.640	0.146	0.999	0.620	0.181	0.001	0.028
No obs (log)	0.254	0.047	1.000	0.252	0.111	0.023	0.086
Midyear of data	-0.007	0.009	0.424				
Female share	0.002	0.062	0.038				
Country & Industry							
Developing	-1.937	0.484	0.992	-2.047	0.791	0.010	0.076
Europe	0.595	0.386	0.782	0.673	0.481	0.162	0.233
Nurses	-2.406	0.525	0.995	-2.656	0.952	0.005	0.063
Teachers	0.196	0.421	0.224				
Method & Identification							
Separations, id.	0.676	0.756	0.511	1.209	0.722	0.094	0.178
Inverse, id.	5.748	0.496	1.000	5.734	1.905	0.003	0.239
Inverse, not id.	8.362	0.555	1.000	8.170	4.223	0.053	0.053
Recruitment	0.044	0.208	0.079				
L on W regression	-0.236	0.517	0.219				
Structural & other, id.	-0.095	0.406	0.085				
Structural & other, not id.	-0.202	0.435	0.223				
Estimation Technique							
Hazard	-2.004	0.357	1.000	-2.069	0.676	0.002	0.006
Probit, logit, other	-0.080	0.229	0.149				
Publication Characteristics							
Top journal	0.767	0.336	0.915	0.781	0.522	0.135	0.229
Citations	-3.144	0.455	1.000	-3.227	0.996	0.001	0.103
Pub. year	0.103	0.018	1.000	0.087	0.053	0.104	0.246
Intercept	-0.384		1.000	-0.368	0.173	0.034	0.136
Observations	787			787			

Notes: PIP denotes posterior inclusion probability; SD is the standard deviation; 'id' denotes estimates obtained with an identification strategy in place. The left panel of the table presents unconditional moments for the BMA. The right panel reports the result of the frequentist check in which we include only explanatory variables with PIP > 0.5. The standard errors in the frequentist check are clustered at the study level. 'p-value (wild)' are wild bootstrap clustered p-values. A detailed description of all variables is available in Table A1.

Figure B3: Model inclusion in Bayesian model averaging: winsorized outliers



Notes: The response variable is the estimate of the elasticity of labor supply to the firm. Each column denotes an individual model; variables are sorted in descending order by their posterior inclusion probability. The cumulative posterior model probabilities are given on the horizontal axis. Blue color (darker in greyscale) indicates that the variable is included in the model and that the estimated sign is positive. Red color (lighter in greyscale) indicates that the variable is included in the model and that the estimated sign is negative. A lack of color implies that the variable is not included in the model. The numerical results of the BMA estimation are reported in Table B4.

Table B4: Why do estimates of supply elasticity vary? Winsorized outliers.

Response variable:	Bayesian model averaging			Frequentist check (OLS)			
Estimate of supply elasticity	Post. mean	Post. SD	PIP	Coef.	Std. er.	p-value	p-value (wild)
Data characteristics							
SE non-inverse	0.203	0.138	0.759	0.275	0.093	0.003	0.173
No obs (log)	0.219	0.045	1.000	0.229	0.100	0.023	0.070
Midyear of data	-0.004	0.008	0.310				
Female share	-0.063	0.200	0.123				
Country & industry							
Developing	-2.423	0.493	1.000	-2.401	0.736	0.001	0.094
Europe	0.961	0.325	0.969	0.938	0.470	0.046	0.143
Nurses	-6.611	0.664	1.000	-6.627	1.365	0.000	0.026
Teachers	-3.403	0.663	1.000	-3.586	0.946	0.000	0.062
Method & identification							
Separations, id.	2.266	0.902	0.943	2.233	0.589	0.000	0.074
Inverse, id.	9.587	0.654	1.000	9.505	2.043	0.000	0.077
Inverse, not id.	11.495	0.631	1.000	11.406	2.812	0.000	0.239
Recruitment	1.763	0.911	0.881	1.802	0.768	0.019	0.164
L on W regression	0.264	0.645	0.188				
Structural & other, id.	-2.206	1.184	0.851	-2.624	1.106	0.018	0.811
Structural & other, not id.	-0.342	0.550	0.335				
Estimation technique							
Hazard	-2.250	0.427	1.000	-2.412	0.613	0.000	0.005
Probit, logit, other	-1.913	0.456	0.993	-2.178	0.501	0.000	0.022
Publication Characteristics							
Top journal	2.930	0.437	1.000	3.129	0.809	0.000	0.022
Citations	-4.671	0.660	1.000	-4.739	0.965	0.000	0.004
Pub. year	0.160	0.019	1.000	0.158	0.085	0.064	0.264
Intercept	-2.852		1.000	-3.102	2.975	0.297	0.609
Observations	801			801			

Notes: PIP denotes posterior inclusion probability; SD is the standard deviation; 'id' denotes estimates obtained with an identification strategy in place. The left panel of the table presents unconditional moments for the BMA. The right panel reports the result of the frequentist check in which we include only explanatory variables with PIP > 0.5. The standard errors in the frequentist check are clustered at the study level. 'p-value (wild)' are wild bootstrap clustered p-values. A detailed description of all variables is available in Table A1.

Table B5: Best practice

Panel A: *75th percentile for best practice*	Point Estimate	95% interval	95% interval (wild)	Implied Markdown
Separations				
Linear – all	4.510	[2.74; 6.28]	[1.66; 8.03]	18.1%
Hazard – all	3.196	[1.51; 4.89]	[0.55; 6.54]	23.8%
Hazard – developing	1.645	[0.16; 3.13]	[-0.36; 4.06]	37.8%
Hazard – Europe	3.812	[1.87; 5.75]	[0.66; 7.28]	20.8%
Hazard – other advanced	2.912	[1.30; 4.53]	[0.46; 6.23]	25.6%
Inverse				
Identified – all	10.416	[6.046; 14.79]	[2.37; 17.27]	8.8%
Identified – developing	8.865	[4.91; 12.83]	[2.14; 15.33]	10.1%
Identified – Europe	11.032	[6.50; 15.56]	[2.76; 18.09]	8.3%
Identified – other advanced	10.133	[5.80; 14.47]	[2.11; 16.97]	9.0%
Panel B: *95th percentile for best practice*	Point Estimate	95% interval	95% interval (wild)	Implied Markdown
Separations				
Linear – all	4.866	[2.88; 6.85]	[2.22; 8.40]	17.0%
Hazard – all	3.552	[1.75; 5.35]	[0.74; 6.92]	22.0%
Hazard – developing	2.001	[0.47; 3.54]	[-0.24; 4.39]	33.3%
Hazard – Europe	4.167	[2.08; 6.26]	[0.93; 7.91]	19.4%
Hazard – other advanced	3.268	[1.58; 4.96]	[0.76; 6.45]	23.4%
Inverse				
Identified – all	10.772	[6.42; 15.13]	[2.86; 17.48]	8.5%
Identified – developing	9.221	[5.31; 13.13]	[2.80; 15.45]	9.8%
Identified – Europe	11.388	[6.85; 15.93]	[3.20; 18.37]	8.1%
Identified – other advanced	10.488	[6.18; 14.79]	[2.62; 17.18]	8.7%

Notes: The table presents fitted 'best practice' estimates for various estimation techniques and data. We used the model estimated as the frequentist check of our baseline specification reported in Table 3. We report both the standard 95% confidence interval calculated for errors clustered at the study level, and the 95% confidence interval calculated with wild bootstrap clusters. The estimates of the markdown are obtained using equation (2).

Table B6: Best practice (all varibles; sep id)

Group	Point Estimate	95% interval	95% interval (wild)	Implied Markdown
Separations				
Linear – Not Identified	4.405	[1.88;6.93]	[0.05;8.84]	18.5%
Linear – Identified	5.932	[3.63;8.23]	[0.43;11.99]	14.4%
Hazard – Not Identified	2.885	[1.08;4.69]	[-0.25;6.39]	25.7%
Hazard – Identified	4.412	[2.73;6.09]	[-0.27;8.39]	18.5%
Inverse				
Not Identified	16.506	[11.86;21.15]	[-1.88;20.60]	5.7%
Identified	10.386	[5.95;14.83]	[0.46;18.83]	8.8%

Notes: The table presents fitted 'best practice' estimates for various estimation techniques and data. Here, we use the frequentist model including all of our explanatory variables, reported in Table B2. We report both the standard 95% confidence interval calculated for errors clustered at the study level, and the 95% confidence interval calculated with wild bootstrap clusters. The estimates of the markdown are obtained using equation (2).

Appendix C Studies Used in Meta-analysis

We used the following search query to find the relevant studies:

> Our search querry is: *("monopsony" OR "monopsonistic" OR "elasticity of labor supply to the firm" OR "separation elasticity" OR "recruitment elasticity") AND ("estimate" "elasticity")*

Papers in Study

Bachmann, R. & H. Frings (2017): "Monopsonistic competition, low-wage labour markets, and minimum wages–an empirical analysis." *Applied Economics* **49(51)**: pp. 5268–5286.

Barth, E. & H. Dale-Olsen (2009): "Monopsonistic discrimination, worker turnover, and the gender wage gap." *Labour Economics* **16(5)**: pp. 589–597.

Van den Berg, G. J. & G. Ridder (1998): "An empirical equilibrium search model of the labor market." *Econometrica* **66(5)**: pp. 1183–1221.

Blau, F. D. & L. M. Kahn (1981): "Race and sex differences in quits by young workers." *ILR Review* **34(4)**: pp. 563–577.

Bodah, M., J. Burkett, & L. Lardaro (2003): "Ix. employment relations for health care workers." In "Proceedings of the Annual Meeting–Industrial Relations Research Association," p. 199. IRRA.

Booth, A. L. & P. Katic (2011): "Estimating the wage elasticity of labour supply to a firm: What evidence is there for monopsony?" *Economic Record* **87(278)**: pp. 359–369.

Campbell III, C. M. (1993): "Do firms pay efficiency wages? evidence with data at the firm level." *Journal of Labor Economics* **11(3)**: pp. 442–470.

Currie, J. (1991): "Employment determination in a unionized public-sector labor market: the case of ontario's school teachers." *Journal of Labor Economics* **9(1)**: pp. 45–66.

Dal Bó, E., F. Finan, & M. A. Rossi (2013): "Strengthening state capabilities: The role of financial incentives in the call to public service." *The Quarterly Journal of Economics* **128(3)**: pp. 1169–1218.

Depew, B., P. Norlander, & T. A. Sørensen (2017): "Inter-firm mobility and return migration patterns of skilled guest workers." *Journal of Population Economics* **30(2)**: pp. 681–721.

Depew, B. & T. A. Sørensen (2013): "The elasticity of labor supply to the firm over the business cycle." *Labour Economics* **24**: pp. 196–204.

Dobbelaere, S. & J. Mairesse (2013): "Panel data estimates of the production function and product and labor market imperfections." *Journal of Applied Econometrics* **28(1)**: pp. 1–46.

Dube, A., J. Jacobs, S. Naidu, & S. Suri (2018): "Monopsony in online labor markets." *American Economic Review: Insights* **(forthcoming)**.

Fakhfakh, F. & F. FitzRoy (2006): "Dynamic monopsony: Evidence from a French establishment panel." *Economica* **73(291)**: pp. 533–545.

Falch, T. (2010): "The elasticity of labor supply at the establishment level." *Journal of Labor Economics* **28(2)**: pp. 237–266.

Falch, T. (2011): "Teacher mobility responses to wage changes: Evidence from a quasi-natural experiment." *American Economic Review* **101(3)**: pp. 460–65.

Falch, T. (2017): "Wages and recruitment: evidence from external wage changes." *ILR Review* **70(2)**: pp. 483–518.

Fleisher, B. M. & X. Wang (2004): "Skill differentials, return to schooling, and market segmentation in a transition economy: the case of mainland China." *Journal of Development Economics* **73(1)**: pp. 315–328.

Galizzi, M. (2001): "Gender and labor attachment: Do within-firms' relative wages matter?" *Industrial Relations: A Journal of Economy and Society* **40(4)**: pp. 591–619.

Hirsch, B. & E. J. Jahn (2015): "Is there monopsonistic discrimination against immigrants?" *ILR Review* **68(3)**: pp. 501–528.

Hirsch, B., E. J. Jahn, & C. Schnabel (2018): "Do employers have more monopsony power in slack labor markets?" *ILR Review* **71(3)**: pp. 676–704.

Hirsch, B., T. Schank, & C. Schnabel (2010): "Differences in labor supply to monopsonistic firms and the gender pay gap: An empirical analysis using linked employer-employee data from Germany." *Journal of Labor Economics* **28(2)**: pp. 291–330.

Hotchkiss, J. L. & M. Quispe-Agnoli (2013): "The expected impact of state immigration legislation on labor market outcomes." *Journal of Policy Analysis and Management* **32(1)**: pp. 34–59.

Howes, C. (2005): "Living wages and retention of homecare workers in san francisco." *Industrial Relations: A Journal of Economy and Society* **44(1)**: pp. 139–163.

Manning, A. (2003): "The real thin theory: monopsony in modern labour markets." *Labour Economics* **10(2)**: pp. 105–131.

Matsudaira, J. D. (2014): "Monopsony in the low-wage labor market? Evidence from minimum nurse staffing regulations." *Review of Economics and Statistics* **96(1)**: pp. 92–102.

Meitzen, M. E. (1986): "Differences in male and female job-quitting behavior." *Journal of Labor Economics* **4(2)**: pp. 151–167.

Naidu, S., Y. Nyarko, & S.-Y. Wang (2016): "Monopsony power in migrant labor markets: evidence from the United Arab Emirates." *Journal of Political Economy* **124(6)**: pp. 1735–1792.

Ogloblin, C. & G. Brock (2005): "Wage determination in urban russia: Underpayment and the gender differential." *Economic Systems* **29(3)**: pp. 325–343.

Ransom, M. R. & R. L. Oaxaca (2010): "New market power models and sex differences in pay." *Journal of Labor Economics* **28(2)**: pp. 267–289.

Ransom, M. R. & D. P. Sims (2010): "Estimating the firm's labor supply curve in a "new monopsony" framework: Schoolteachers in Missouri." *Journal of Labor Economics* **28(2)**: pp. 331–355.

Staiger, D. O., J. Spetz, & C. S. Phibbs (2010): "Is there monopsony in the labor market? evidence from a natural experiment." *Journal of Labor Economics* **28(2)**: pp. 211–236.

Sulis, G. (2011): "What can monopsony explain of the gender wage differential in Italy?" *International Journal of Manpower* **32(4)**: pp. 446–470.

Sullivan, D. (1989): "Monopsony power in the market for nurses." *The Journal of Law and Economics* **32(2, Part 2)**: pp. S135–S178.

Vick, B. (2017): "Measuring links between labor monopsony and the gender pay gap in Brazil." *IZA Journal of Development and Migration* **7(1)**: p. 10.

Wasylenko, M. J. (1977): "Some evidence of the elasticity of supply of policemen and firefighters." *Urban Affairs Quarterly* **12(3)**: pp. 365–382.

Webber, D. A. (2015): "Firm market power and the earnings distribution." *Labour Economics* **35**: pp. 123–134.

Webber, D. A. (2016): "Firm-level monopsony and the gender pay gap." *Industrial Relations: A Journal of Economy and Society* **55(2)**: pp. 323–345.