Monopsony in Labor Markets: A Meta-Analysis

Todd Sorensen and Anna Sokolova

Appendix A: Description of Variables

Table A1: Definitions and summary statistics of explanatory variables

Variable	Description	Mean (all)	SD (all)	N (all)	Mean (95%)	SD (95%)	N (95%)
Data characteristics							
SE non-inverse	An interaction between standard error and a dummy for whether the estimate is obtained through 'direct' (not inverse) estimation.	1.90	3.96	1320	1.95	3.95	1254
No obs (log)	The logarithm of the number of observations.	10.23	3.21	1320	10.16	3.17	1254
Midyear of data	The average year of the data used minus 1919 (the earliest midyear in the sample).	75.50	17.39	1320	75.62	17.01	1254
Female share	The share of female workers in the study's data set; 0.5 if sample stats not reported.	0.51	0.28	1320	0.52	0.28	1254
*Country & occupation**							
Developing	=1 for data coming from countries classified as 'Emerging and Developing economies' by IMF classification in 2018.	0.10	0.30	1320	0.08	0.27	1254
Europe	=1 for data coming from countries in Europe.	0.26	0.44	1320	0.27	0.44	1254
Nurses	=1 for data that exclusively covers the market of nurses.	0.06	0.24	1320	0.06	0.24	1254
Teachers	=1 for data that exclusively covers the market of teachers.	0.08	0.27	1320	0.08	0.27	1254
*[Reference category for COUNTRY: other advanced economies.] [Reference category for OCCUPATION: 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.	0.20	0.40	1320	0.21	0.41	1254
Inverse, id.	=1 if estimate converted from inverse elasticity AND is obtained through IV identification strategy.	0.10	0.30	1320	0.08	0.28	1254
Inverse, not id.	=1 if estimate converted from inverse elasticity AND the authors do not use IV.	0.04	0.19	1320	0.02	0.15	1254
Recruitment, id.	=1 if estimate is based on recruitment rate AND is obtained through IV or other identification strategy.	0.06	0.25	1320	0.07	0.25	1254
Recruitment, not id.	if estimate based on recruitment rate AND the authors do not use IV.	0.01	0.07	1320	0.00	0.07	1254
L on W regression, id.	=1 if estimate is obtained via stock-based estimation through regressing labor on wage AND is obtained through either IV or randomized identification strategy.	0.05	0.22	1320	0.05	0.22	1254
Structural & other, id.	=1 if estimate obtained from structural model with production, or any other method not based on separations and not covered by specification regressor above AND is obtained through either IV or randomized identification strategy.	0.02	0.15	1320	0.02	0.15	1254
Structural & other, not id.	=1 if estimate obtained from structural model with production, or any other method not based on separations and not covered by specification regressor above AND the authors do not use IV or randomize.	0.06	0.24	1320	0.07	0.25	1254
**[Reference category: estimates based on separations AND not identified (no IV or randomized identification)]							

Continued on next page

Table A1: Definitions and summary statistics of explanatory variables

Variable	Description	Mean (all)	SD (all)	N (all)	Mean (95%)	SD (95%)	N (95%)
*Estimation technique******							
Hazard	=1 if study uses hazard model (reference category: linear techniques).	0.23	0.42	1320	0.24	0.42	1254
Probit, logit, other	=1 if study uses probit, logit or any other non-linear technique not previously classified (reference category: linear techniques).	0.16	0.36	1320	0.16	0.37	1254
***[Reference category: estimates based on linear techniques]							
*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.	0.26	0.44	1320	0.27	0.44	1254
Citations	The logarithm of the number of per-year citations of the study in Google Scholar (data for May 2019).	0.48	0.40	1320	0.48	0.41	1254
Pub. year (google)	The year the paper first appeared on Google Scholar minus 1977, the year when the first study in our sample was published.	33.98	7.08	1320	33.88	7.05	1254
NBER or IZA	=1 if estimate comes from an unpublished NBER or IZA working paper.	0.06	0.24	1320	0.06	0.24	1254
Working Other	=1 if estimate comes from other unpublished working paper.	0.09	0.29	1320	0.07	0.26	1254
****[Reference category for PUBLICATION STATUS: estimates that are published]							

Notes: Data was collected from published studies estimating η. When indicator variables form groups, we state the reference category. We report means and standard deviations for the full sample of 1320 observations, as well as for the truncated subsample of 1254 estimates.

Appendix B Publication bias. Additional results

Publication bias using Andrews and Kasy (2019)

In this section we will explicitly model selectivity under publication bias using techniques developed by Andrews and Kasy (2019) ('AK 2019', for brevity). We will then estimate relative publication probabilities of different results and the unbiased means of the population of latent studies (i.e. the mean 'true' effect). In the next paragraphs we will briefly explain the method following closely the discussion presented in *Sections IIB and IIIC* of Andrews and Kasy (2019); we will also estimate the model featured in their *Section IIIC* on our data.[1]

Consider studies estimating the supply elasticity parameter. In the AK 2019 setup, each study's underlying 'true' elasticity is drawn from some distribution. The authors make specific assumptions about this distribution's shape (e.g. normal, t-distribution) and later estimate the associated parameters (e.g. the mean of the distribution). A latent study then produces estimates of elasticity that are drawn from a normal distribution with the study's 'true' underlying elasticity as a mean, and with a standard error that is independent of the 'true' elasticity. Out of the estimates produced, some will be reported, while others will be discarded. The probability of reporting, $p(Z)$, may depend on the value of the estimate normalized by its standard error, Z.[2] The probability function $p(Z)$ may depend on the statistical significance (captured by $|Z|$) and the sign of the results (captured by the sign of Z). Absent selectivity, the observed distribution of reported results with high standard errors should reflect the distribution of results with low standard errors—plus noise. AK 2019 identify $p(Z)$ by comparing distributions of results with different standard errors.

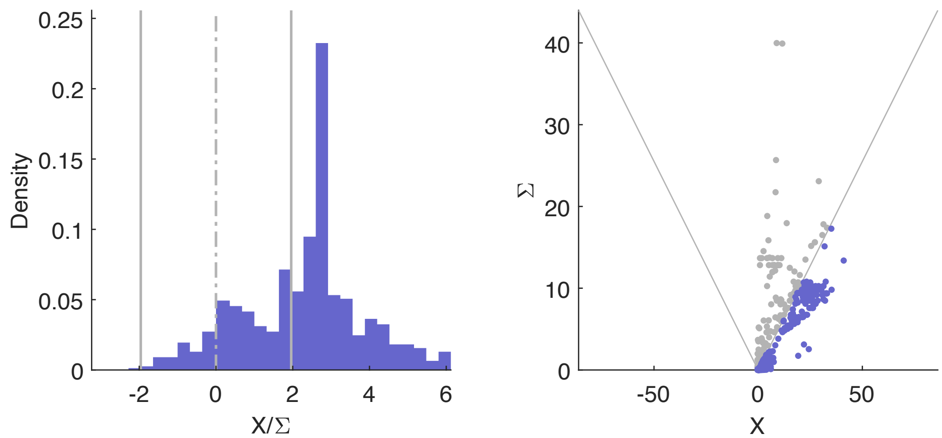
Following AK 2019, we start with a visual diagnostic test for our data (Figure B1). Adopting the notation of AK 2019, we denote our elasticity estimates with X and their standard errors with Σ. The panel on the left presents a histogram of estimates of elasticity normalized by their standard errors. In the absence of selectivity, we would expect to see a smooth distribution. For our data, we notice that the density seems to be jumping around the cutoff of 0, and also roughly around 2. This suggests that the sign and significance of the latent estimates may affect their likelihood of being reported. The right panel plots estimates against their standard errors. If there was no selectivity at play here, the mean of the observed elasticity estimates X would not depend on the level of precision. Visually, this implies that if one was to draw two horizontal lines at different levels of Σ, then the mean values of X points plotted along those lines should be approximately equal. For our data, standard errors around 10 seem to be associated with higher mean reported estimates of the elasticity compared to standard errors close to 1, for which the estimates seem to cluster relatively close to zero. This, too, points

[1] Andrews and Kasy (2019) discuss two major applications of their method: an application that utilizes estimates from replication studies and the one designed for the meta-study context, that only employs the results reported in original studies. Due to the nature of our data we will only use the latter approach; this is also the approach we will refer to throughout the text when using the 'AK 2019' notation.

[2] Throughout the paper, AK 2019 refer to $p(Z)$ as the 'publication probability', but also note that selectivity may not necessarily occur as a result of the publication process; it may be driven by researcher's decisions not to report certain results. We will refer to $p(Z)$ as a probability of the result being reported, as our application features data from both published and unpublished work.

towards selectivity in the data, as there seem to be substantial structural differences between the distributions of observed estimates with high and low standard errors.

Figure B1: 'Direct' estimates: AK2019 methodology



Notes: The figures present a visual diagnostic test for selective reporting that follows the discussion in AK 2019. We produce this diagram using the `Matlab` code accompanying AK 2019 that replicates their Figure 9 (p.2785). Using notation from AK 2019, we denote 'direct' elasticity estimates with X and their standard errors with Σ. The solid grey lines mark thresholds for estimates being significant at a 5% level, i.e. $X/\Sigma = -1.96$ and $X/\Sigma = 1.96$.

We will now estimate the version of the AK 2019 model discussed in *Section IIIC* of the body of their paper, in which the authors use data from the Wolfson and Belman (2019) meta-analysis of the elasticity of employment to changes in the minimum wage. The authors assume the distribution of the 'true' underlying elasticity to be a t-distribution with degrees of freedom $\tilde{\nu}$, scale parameter $\tilde{\tau}$ and the location parameter $\bar{\theta}$—the mean 'true' elasticity. The relative probability of results being reported is then modeled using a step function:

$$p(Z) \propto \begin{cases} \beta_{p,1} & \text{if } Z < -1.96 \\ \beta_{p,2} & \text{if } -1.96 \leq Z < 0 \\ \beta_{p,3} & \text{if } 0 \leq Z < 1.96 \\ 1 & \text{if } Z \geq 1.96 \end{cases} \tag{10}$$

where the probability of reporting a positive result significant at 5% is normalized to 1, while the relative reporting probabilities of significant negative ($\beta_{p,1}$), insignificant negative ($\beta_{p,2}$) and insignificant positive ($\beta_{p,3}$) results are estimated via maximum likelihood.

The results of estimating this model on our data are reported in Table B1. *Panel A* reports results for the full data set of 'direct' estimates in which we include results reported in both published and unpublished work. The methodology featured here is identical to that employed to obtain *Table 3* in AK 2019. For our data, there is evidence that the reporting probability function does indeed depend on Z. The reporting probability for a negative estimate is dramat-

Table B1: Testing for publication bias using Andrews and Kasy (2019)

$\bar{\theta}$	$\tilde{\tau}$	$\tilde{\nu}$	$\beta_{p,1}$	$\beta_{p,2}$	$\beta_{p,3}$
Panel A: All estimates					
0.157	0.648	1.570	0.005	0.036	0.111
(0.001)	(0.006)	(0.028)	(0.003)	(0.048)	(0.074)
Panel B: Published Estimates Only					
-0.269	0.548	1.435	0.002	0.020	0.073
(0.295)	(0.717)	(0.600)	(0.002)	(0.024)	(0.063)

Notes: This table presents results of estimating the model discussed in *Section IIIC* of AK 2019, using our sample of 'direct' estimates (*Panel A*) and the sub-sample of 'direct' estimates that were published (*Panel B*). The model estimated assumes that the 'true' underlying elasticity Ω^* is distributed according to $\Omega^* \sim \bar{\theta} + t(\tilde{\nu}) \cdot \tilde{\tau}$, and that the reporting probability is proportional to those featured by a step function in (10). See p.2784 of AK 2019 for more details. We produce our results using the `Matlab` code accompanying AK 2019 that replicates their *Table 3*. Similar to our Table 2, *Panel A* reports results for the sample of 1118 'direct' estimates, both published and unpublished. We repeat this exercise under different outlier treatments and report the results in Table D3 of Online Appendix D; the results are similar. *Panel B* restricts the sample to the 995 'direct' estimates that are published; these results are also not very sensitive to outlier treatments (see Table D4 of Online Appendix D).

ically lower compared to an estimate that is positive and significant on a 5% level. Specifically, a negative insignificant result is about 28 times less likely to be reported, and a negative significant result would be reported even less often. A positive result with a Z lower than 1.96 is about nine times less likely to be reported compared to a result with Z over 1.96. The point estimate of $\beta_{p,3}$ is somewhat larger than $\beta_{p,1}$ and $\beta_{p,2}$, suggesting that positive insignificant results may be relatively more likely to be reported compared to the negative results (although this result is not precise enough for a statistical rejection of parameters being equal). The estimate of $\bar{\theta}$—the mean of the distribution of the 'true' underlying elasticity across studies—is at 0.157 which is much smaller compared to the mean elasticity we observe in the truncated sample of reported estimates. We follow AK 2019 and repeat this exercise using the sub-sample of published studies; *Panel B* of Table B1 reports the results. Compared to the full sample, the point estimates of relative probabilities are somewhat smaller. The point estimate of the unbiased mean of the 'true' effect is also smaller, but much less precise. We therefore conclude that the results are roughly similar across the two samples, with the sub-sample of published studies being associated with somewhat stronger selectivity. All these pieces of evidence suggest that selectivity is indeed very prominent in the monopsony literature.

Appendix C Heterogeneity. Additional results

Appendix C.1 Addressing model uncertainty: Bayesian Model Averaging

The model with all 23 regressors included that we have studied in Table 3 is only one out of 2^{23} possible combinations of our chosen explanatory variables. Here we attempt to take into account the remaining $2^{23} - 1$ possible combinations of regressors and address model uncertainty using Bayesian Model Averaging (BMA). Sequential t-testing would discard the information coming from regressors that appear insignificant in the broad specification of Table 3 and Table 4; BMA offers an alternative approach: instead of selecting and estimating one model, it traverses through the space of all 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.

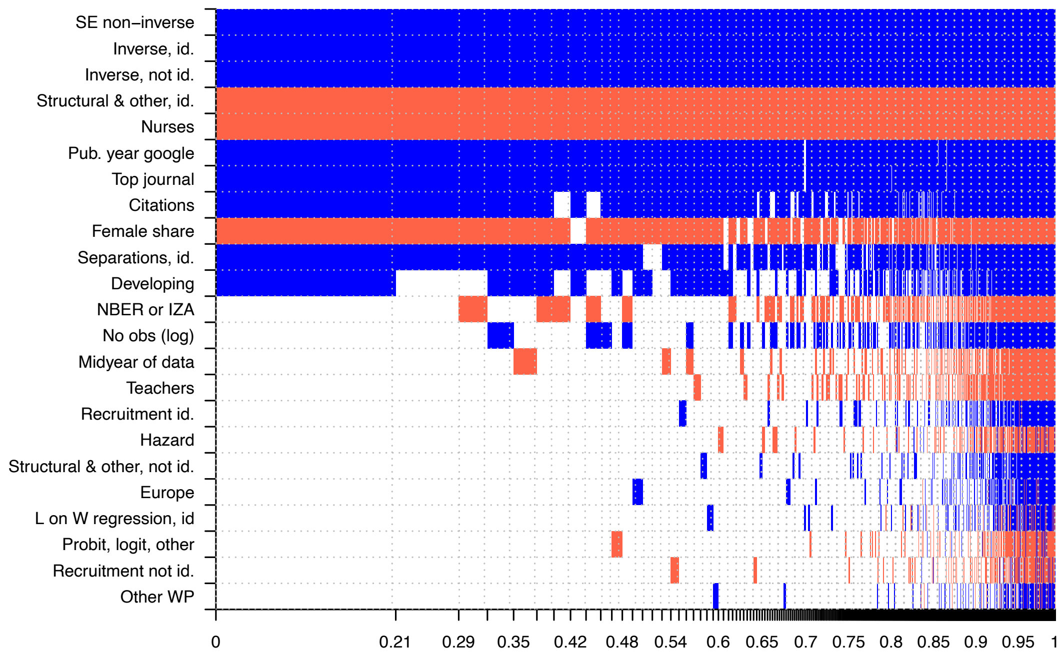
Inference in BMA is obtained by taking a weighted average of the results from all possible models, using the Posterior Model Probability (PMP) as a weight. It is worth noting that we do not estimate each of the 2^{23} regressions; instead, we employ a Model Composition Markov Chain Monte Carlo algorithm that visits models with the highest PMP and approximates the rest (see Madigan and York 1995). We implement this using the `BMS` package in `R` written by Zeugner and 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.

BMA estimation results for the full sample of 1254 estimates are reported in Figure C1. The explanatory variables shown on the left are sorted by Posterior Inclusion Probability (PIP). Each explanatory variable is present in $2^{23} - 2^{22}$ 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 C1 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 C1 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.

We observe that the signs of most explanatory variables are quite stable across models in which the variables are included; they are also broadly consistent with evidence reported in Table 3. We present numerical results of BMA estimation in the left panel of Table C1, 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 C1 reports a frequentist robustness check in which we run an OLS with variables that have PIP higher than 50%. We cluster standard errors at the study level and additionally compute p-values using wild bootstrap clustering.

Figure C1: 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 C1.

Table C1: Why do estimates of supply elasticity vary?
Bayesian Model Averaging

	BMA			OLS with selected variables			
Response variable:	Post. Mean	Post. SD	PIP	Coef.	SE	P-value	P-value (wild)
Data Characteristics							
SE non-inverse	0.977	0.073	1.000	0.983	0.298	0.001	0.002
No obs (log)	0.044	0.086	0.260				
Midyear of data	-0.003	0.008	0.124				
Female share	-2.167	1.056	0.880	-2.499	1.840	0.174	0.383
Country & Occupation							
Developing	1.358	1.320	0.580	2.384	2.959	0.420	0.503
Europe	0.024	0.156	0.048				
Nurses	-6.366	1.247	1.000	-6.346	4.875	0.193	0.301
Teachers	-0.193	0.679	0.111				
Method & Identification							
Separations, id.	2.970	1.540	0.866	3.469	3.672	0.345	0.499
Inverse, id.	15.836	1.043	1.000	15.436	7.376	0.036	0.109
Inverse, not id.	19.528	1.337	1.000	19.513	2.088	0.000	0.006
Recruitment, id.	0.110	0.543	0.068				
Recruitment, not id.	-0.087	0.698	0.039				
L on W regression, id	0.055	0.468	0.048				
Structural & other, id.	-8.596	1.745	1.000	-8.901	4.848	0.066	0.047
Structural & other, not id.	0.055	0.314	0.054				
Estimation Technique							
Hazard	-0.052	0.294	0.061				
Probit, logit, other	-0.028	0.193	0.047				
Publication Characteristics							
Top journal	3.656	1.048	0.989	3.305	1.733	0.057	0.026
Citations	2.073	0.944	0.896	2.408	1.440	0.094	0.255
Pub. year (google)	0.229	0.060	0.990	0.202	0.102	0.047	0.112
NBER or IZA	-0.634	1.164	0.276				
WP other	0.013	0.202	0.030				
Constant	-6.883		1.000	-5.842	3.653	0.110	0.173
N	1254	.	.	1254	.	.	.

Notes: Here we present results of Bayesian Model Averaging estimation. 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. Here we only use estimates that remain after we apply the outlier treatment strategy discussed in Section 2 (i.e. cutting 2.5% of outliers from each tail).

As before, we see strong support for our conjecture about publication bias: the posterior inclusion probability corresponding to the standard error of 'direct' estimates is at 100%, and its correlation with the reported estimates is high and statistically significant in the OLS robustness check. Similarly, we observe that estimates converted from inverse elasticities are markedly higher compared to those obtained using separations, while estimates obtained using structural models with an identification strategy are lower.

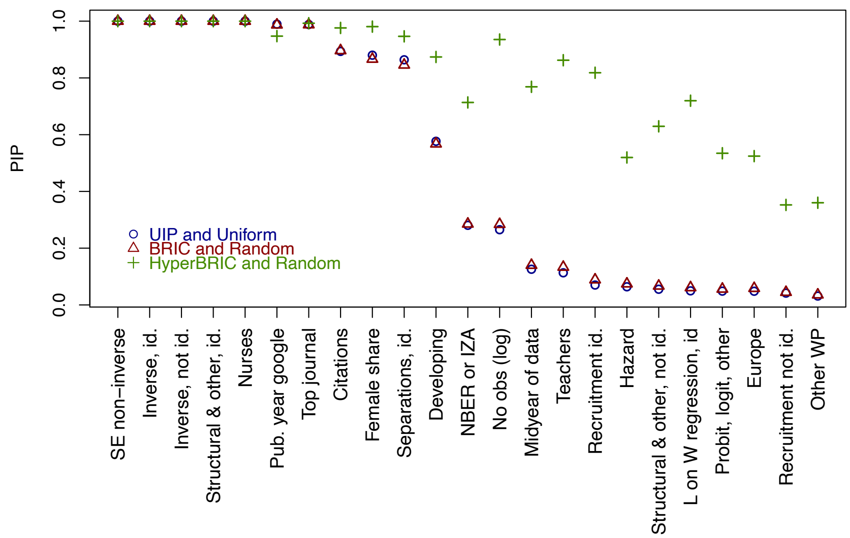
Compared to the results reported in Table 3, we see somewhat stronger support for the negative relationship between the elasticity estimates and the female shares in the associated data sets: BMA estimates the posterior inclusion probability for this variable to be around 88%. At the same time, the OLS robustness check in the right panel still does not have enough power to establish statistical significance, even though the magnitude and sign of the coefficient is consistent with both BMA estimation and results reported in Table 3. We observe a similar pattern looking at the estimated coefficient for nurses: the BMA suggests that this variable is likely part of the 'true' model with a negative associated effect; at the same time, the OLS robustness check does not show strong statistical significance for this variable, albeit it does report a similar parameter estimate. In a similar vein, we see weak evidence of differences in monopsony power across advanced non-European and developing countries, as well as between separations-based estimates obtained with and without an identification strategy.

Figure C2 compares coefficients and posterior inclusion probabilities estimated by BMA under alternative prior settings; the results discussed above appear resilient to assumptions about priors, as the posterior inclusion probabilities and the associated coefficient estimates are similar under different prior assumptions. At the same time, for some of the variables that our baseline BMA did not find to be likely belonging to the 'true' model, the results obtained under HyperBRIC and Random priors suggest higher likelihood of inclusion (while at the same time reporting roughly similar coefficient estimates).

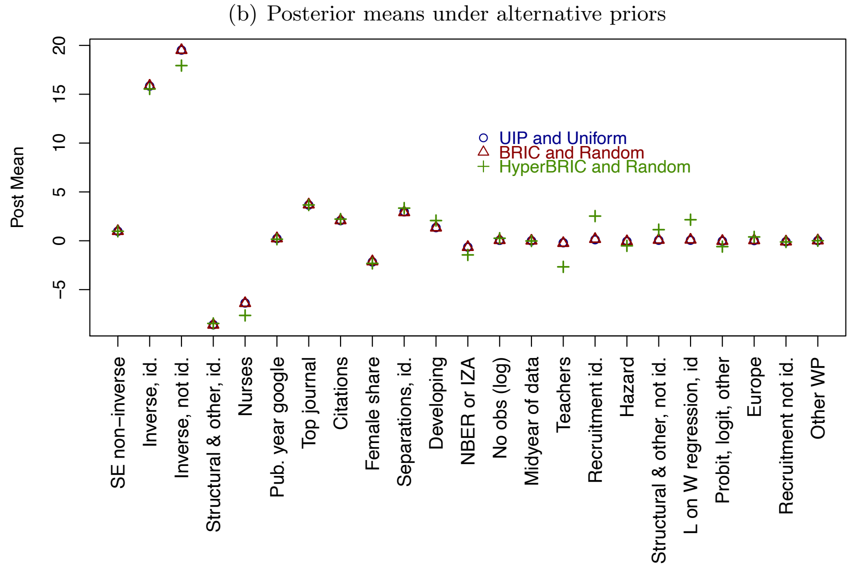
Finally, Table C2 reports the quantitative results of applying BMA to a sub-sample of the 549 identified estimates. Many of the results are similar to those discussed before. In addition, the table shows somewhat stronger evidence for high monopsony power being associated with larger shares of female workers. Unlike in Table 4 reporting OLS results for identified estimates, here we do not document a meaningful discrepancy between separation- and recruitment-based estimates. On the other hand, both BMA and OLS robustness check results suggest a sizable differences between estimates obtained using separations and those converted from inverse elasticities, obtained with a stock-base regression of labor on wage or derived using structural models. We additionally evaluate the sensitivity of BMA estimation to the outlier treatment and report the results in Online Appendix E and particularly Appendix E.3. The results are broadly consistent.

Figure C2: 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 C1 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 parameters (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 C2: Why do estimates of supply elasticity vary?
BMA, identified estimates only.

	BMA			OLS with selected variables			
Response variable:	Post. Mean	Post. SD	PIP	Coef.	SE	P-value	P-value (wild)
Data Characteristics							
SE non-inverse	0.863	0.107	1.000	0.904	0.211	0.000	0.067
No obs (log)	0.532	0.425	0.685	0.825	0.542	0.128	0.259
Midyear of data	-1.302	0.149	1.000	-1.383	0.248	0.000	0.008
Female share	-20.909	2.981	1.000	-20.652	10.601	0.051	0.224
Country & Occupation							
Developing	0.538	1.778	0.127				
Europe	-4.911	3.001	0.801	-5.858	1.121	0.000	0.025
Nurses	-0.792	2.827	0.124				
Teachers	-1.721	2.708	0.350				
Method & Identification							
Inverse	11.263	3.413	0.993	10.893	5.898	0.065	0.201
Recruitment	0.665	2.475	0.156				
L on W regression	5.547	3.600	0.800	6.551	3.417	0.055	0.183
Structural & other, id.	-9.336	4.403	0.893	-9.091	4.069	0.025	0.119
Estimation Technique							
Probit, logit, other	0.114	1.064	0.068				
Publication Characteristics							
Top journal	-0.039	0.819	0.079				
Citations	4.187	1.260	0.982	4.163	1.365	0.002	0.022
Pub. year (google)	1.571	0.211	1.000	1.600	0.312	0.000	0.031
NBER or IZA	5.281	4.665	0.631	8.818	4.448	0.047	0.252
WP other	-1.373	3.398	0.212				
Constant	60.989		1.000	62.748	16.420	0.000	0.014
N	549	.	.	549	.	.	.

Notes: Here we present results of Bayesian Model Averaging estimation using a sub-sample of identified estimates. 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. Here we only use estimates that remain after we apply the outlier treatment strategy discussed in Section 2 (i.e. cutting 2.5% of outliers from each tail).

Appendix C.2 Addressing model uncertainty: LASSO

In this subsection we try to tackle model uncertainty by implementing LASSO. The intuition behind the LASSO approach can be summarized as follows. We think that a good model explaining variation in supply elasticity estimates should be sparse, i.e. the key variation in the elasticity parameter could be captured by a smaller subset of the 23 explanatory variables that we introduced. However, we experience difficulties selecting the subset of variables that should be included. The OLS procedure performed on all 23 variables does not assign exact zeros to any of the coefficient estimates—by construction, as OLS solves an unconstrained minimization problem. LASSO introduced by Tibshirani (1996) amends the OLS approach by adding to the minimization problem a constraint that demands the sum of absolute values of the variable coefficients to be smaller or equal to an upper bound, t (that is smaller than the sum of absolute values of coefficients in an unconstrained OLS). Unlike OLS, the LASSO procedure would often yield corner solutions that assign exact zeros to coefficients corresponding to the weaker predictors, achieving sparsity. The specific value of the upper bound t is typically chosen through cross-validation, which is the approach we will also follow here. Further details on LASSO implementation can be found in Hastie et al. (2015).

We implement LASSO with cross-validation using the `cvlasso` command in `STATA`. We employ 10-fold cross-validation and choose t that minimizes the mean-squared prediction error. The left panel of Table C3 reports coefficient estimates obtained with LASSO. The fact that LASSO forces the sum of absolute values of regression coefficients to lie within a specific upper limit causes individual coefficients to shrink towards zero, which results in a bias (see Belloni et al. 2012). To correct for this shrinkage bias, we follow a post-LASSO estimation procedure discussed in Belloni et al. 2012 which discards variables not selected by the LASSO with cross-validation and runs an OLS on variables that remain. We report the post-LASSO estimation results in the right panel of Table C3.

The two variables discarded by the LASSO procedure are the variable for working papers that came out in outlets other than IZA or NBER, and a regressor for non-identified recruitment-based estimates. This is consistent with the results discussed so far, as neither of these variables showed statistical significance in any of the specifications we studied. The coefficients on the variables that remain in the post-LASSO estimation are very similar to those reported in Table 3, showing significant effects associated with some method and identification choices (e.g. converting estimates from inverse elasticities, using identified recruitment elasticities or structural models with an identification strategy). We also see some weak evidence suggesting that certain occupations (i.e. nurses and teachers) and demographic features (i.e. high shares of female employees) could be associated with higher monopsony power. Finally, we observe a significant positive correlation between estimates and their standard errors which, once again, we interpret as strong evidence of selective reporting in the monopsony literature.

Focusing on a subset of identified estimates yields results that echo those obtained using the full sample (see Table C4). We observe relatively strong effects associated with method and identification choices, as well as evidence pointing to the importance of publication bias.

The contribution of occupation and demographics appears slightly more prominent, as the negative coefficients on nurses and teachers become statistically significant, and the negative effect associated with female shares increases in magnitude. As before, we present the estimation results obtained under an alternative outlier treatment in Online Appendix E (see specifically Appendix E.4), noting that the results are broadly consistent.

Table C3: Why do estimates of supply elasticity vary? LASSO.

	LASSO	OLS using selected variables			
Response variable:	Coef.	Coef.	SE	P-value	P-value (wild)
Data Characteristics					
SE non-inverse	0.979	0.983	0.307	0.001	0.001
No obs (log)	0.321	0.333	0.250	0.183	0.352
Midyear of data	-0.025	-0.027	0.018	0.129	0.312
Female share	-2.315	-2.361	1.800	0.190	0.458
Country & Occupation					
Developing	2.388	2.429	3.089	0.432	0.586
Europe	0.581	0.618	1.040	0.552	0.632
Nurses	-7.943	-8.170	6.052	0.177	0.392
Teachers	-3.379	-3.619	2.259	0.109	0.179
Method & Identification					
Separations, id.	3.465	3.465	3.321	0.297	0.459
Inverse, id.	15.394	15.533	7.121	0.029	0.116
Inverse, not id.	17.458	17.466	3.117	0.000	0.000
Recruitment, id.	2.973	3.220	1.765	0.068	0.101
Recruitment, not id.	0.000	0.000	.	.	.
L on W regression, id	2.886	3.156	3.045	0.300	0.488
Structural & other, id.	-8.525	-8.663	4.562	0.058	0.053
Structural & other, not id.	1.827	1.952	1.790	0.276	0.474
Estimation Technique					
Hazard	-0.929	-0.957	1.762	0.587	0.704
Probit, logit, other	-1.226	-1.291	1.542	0.402	0.569
Publication Characteristics					
Top journal	3.550	3.577	1.410	0.011	0.024
Citations	2.281	2.361	1.657	0.154	0.318
Pub. year (google)	0.148	0.144	0.121	0.234	0.312
NBER or IZA	-1.747	-1.780	1.805	0.324	0.443
WP other	0.000	0.000	.	.	.
Constant	-5.137	-5.053	3.885	0.193	0.310
N	1254	1254	.	.	.

Notes: The left panel presents estimates obtained using LASSO with the penalty value selected to minimize mean-squared prediction error through cross-validation. We implement this in `STATA` using the `cvlasso` routine. Variables with zero coefficient values are excluded under the optimal penalty parameter value. The right panel shows results of estimating OLS using the subset of variables selected by LASSO. We report regular p-values and p-values from wild bootstrap clustering; 'id' denotes estimates obtained with an identification strategy in place. A detailed description of all variables is available in Table A1. Here we only use estimates that remain after we apply the outlier treatment strategy discussed in Section 2 (i.e. cutting 2.5% of outliers from each tail).

Table C4: Why do estimates of supply elasticity vary? LASSO
Identified estimates only

	LASSO	OLS using selected variables			
Response variable:	Coef.	Coef.	SE	P-value	P-value (wild)
Data Characteristics					
SE non-inverse	0.928	0.948	0.227	0.000	0.093
No obs (log)	0.804	0.750	0.576	0.193	0.343
Midyear of data	-1.221	-1.409	0.284	0.000	0.040
Female share	-17.559	-18.935	11.778	0.108	0.457
Country & Occupation					
Developing	4.914	4.924	7.059	0.485	0.667
Europe	-3.631	-5.292	2.549	0.038	0.178
Nurses	-7.572	-11.516	5.399	0.033	0.139
Teachers	-5.261	-7.338	1.543	0.000	0.006
Method & Identification					
Inverse	7.755	9.274	5.038	0.066	0.081
Recruitment	1.969	5.912	2.740	0.031	0.030
L on W regression	6.984	10.550	3.990	0.008	0.117
Structural & other	-11.584	-10.848	3.652	0.003	0.014
Estimation Technique					
Probit, logit, other	0.000	0.000	.	.	.
Publication Characteristics					
Top journal	-0.246	-0.619	2.430	0.799	0.776
Citations	4.494	5.358	1.084	0.000	0.003
Pub. year (google)	1.051	1.086	0.418	0.009	0.059
NBER or IZA	6.566	5.983	4.313	0.165	0.269
WP other	0.000	0.000	.	.	.
Constant	68.364	82.987	20.631	0.000	0.060
N	549	549	.	.	.

Notes: Here we employ the sub-sample of the identified estimates. The left panel presents estimates obtained using LASSO with the penalty value selected to minimize mean-squared prediction error through cross-validation. We implement this in `STATA` using the `cvlasso` routine. Variables with zero coefficient values are excluded under the optimal penalty parameter value. The right panel shows results of estimating OLS using the subset of variables selected by LASSO. We report regular p-values and *p*-values from wild bootstrap clustering; 'id' denotes estimates obtained with an identification strategy in place. A detailed description of all variables is available in Table A1. We only use estimates that remain after we apply the outlier treatment strategy discussed in Section 2 (i.e. cutting 2.5% of outliers from each tail).

Appendix C.3 Heterogeneity and country-specific variables

So far our analysis did not uncover any stable relationship between estimates and the geographical origin of the data. Taken at face value, this result could imply that there are no notable systematic differences in monopsony power across the regions that we studied. Alternatively, this could imply that our method of splitting data into country groups failed to reflect some key cross-country dimensions that govern the size of the elasticity parameter.

Here we take an alternative approach to studying cross-country differences in monopsony power. Instead of using region dummies, we collect country-specific information on factors that, we believe, could affect the wage-setting power of firms: country-specific labor and product market conditions, as well as the general level of economic development. To capture labor market conditions, we use data on collective bargaining coverage, strictness of employee protection and active labor market program expenditures. We capture product market conditions with data on product market regulation. Finally, we proxy for the level of economic development using GDP per capita. A detailed description of variables and data sources is available in Table C5. Our strategy is similar to that of Foged et al. (2019), who conduct a meta-analysis of the effect of immigration on natives' labor market outcomes.[3]

Having collected this data, we attempt to match our observations of supply elasticities with country-year information on our five chosen variables. We are able to match each of our estimates with the exact corresponding GDP per capita. Unfortunately, data on the rest of the country variables is more scarce, and for some of the elasticity estimates we do not have the corresponding country-years of the labor or product market variable. When we had any data on these variables for a given country, we imputed using the `ipolate` command in `Stata`.[4]

We repeat the exercise of Table 3, substituting the region dummies with the new set of country-specific variables. We report results obtained on a larger sample that uses imputed data, as well as a smaller subsample that only includes estimates for which we were able to find the exact matches of cross-country variables. We do not repeat this exercise for the subsample of identified estimates, as for this smaller subsample we do not have enough variation to estimate the effects of the explanatory variables. Table C6 presents estimation results. For brevity, we only report the effects associated with the cross-country variables, and the F-tests for cross-country variable groups.

[3] We use all institutional factors employed by Foged et al. (2019), with the exception of their measure of job tenure, which may be endogenous to the estimation of η, as many of the estimates are obtained using the separation approach which is based upon duration at a given job spell.

[4] Our dataset includes estimates from 16 countries: Australia, Belgium, Brazil, Canada, China, Colombia, France, Germany, Italy, Mexico, The Netherlands, Norway, Russia, the UAE, UK and US. Our measure of employment protection was missing in all years for the UAE. The product market regulation variable was missing in all years for China, Colombia, Russia and the UAE. The active labor market program and collective bargaining variables were missing for these countries and Brazil.

Table C5: Definitions and summary statistics: country-specific variables

Variable	Description	Imputed country data Mean (95%)	SD (95%)	N (all)	Raw country data Mean (95%)	SD (95%)	N (95%)
Col. bargaining coverage	Collective bargaining coverage, measures percentage of employees with the right to bargain.	36.07	28.41	1154	29.24	23.20	817
Strictness of emp. protect.	Strictness of employment protection – individual dismissals (regular contracts) indicator.	1.01	0.96	1251	1.12	1.00	980
ALMP expenditure	Public expenditure on active labor market programs as a percentage of GDP	1.76	1.65	1154	0.75	0.56	473
Product market reg.	Product market regulation indicator	2.13	0.75	1227	1.72	0.30	809
GDP p. c.	Real GDP Per-Capita	44966.2	18772.2	1254	44966.2	18772.2	1254

Notes: Data on labor market institutions is taken from 'Labour' section of stats.oecd.org; product market regulation data is from 'Public Sector, Taxation and Market Regulation' section of stats.oecd.org. Pre-war U.S. GDP data is from Williamson (2014). All other GDP data is taken from World Bank: databank.worldbank.org/source/gender-statistics. GDP was deflated by USD using data from: www.multpl.com/gdp-deflator/table/by-year

Overall, we are not able to capture strong effects associated with labor market conditions. This, however, does not necessarily imply that labor market conditions are unimportant: the two samples we study are characterized by high degrees of multicollinearity which could be inflating the associated standard errors. At the same time, we do observe a relatively stable effect associated with product market regulations: this variable is significant in some of the specifications, and in most of them the respective coefficient is positive. This may indicate that restrictive product market regulations have the effect of decreasing firm size, thus increasing the number of firms and reducing the risk of a labor market becoming oligopsonistic.

Table C6: Why do estimates of supply elasticity vary? Country-specific variables

	OLS, unweighted							
	Imputed country data				Raw country data			
Response variable:	(1)	(2)	(3)	(4)	(1')	(2')	(3')	(4')
Col. bargaining coverage	-0.022	.	.	-0.006	0.054	.	.	-0.061
	(0.252)	.	.	(0.893)	(0.044)	.	.	(0.233)
	[0.503]	.	.	[0.910]	[0.087]	.	.	[0.622]
Strictness of emp. protect.	.	0.346	.	-0.158	.	-2.501	.	5.722
	.	(0.721)	.	(0.904)	.	(0.028)	.	(0.105)
	.	[0.815]	.	[0.927]	.	[0.398]	.	[0.561]
ALMP expenditure	.	.	-0.349	-0.196	.	.	0.440	-6.703
	.	.	(0.189)	(0.609)	.	.	(0.505)	(0.175)
	.	.	[0.409]	[0.631]	.	.	[0.601]	[0.605]
Product market reg.	2.185	0.600	1.817	2.094	-27.628	1.292	28.540	40.935
	(0.069)	(0.759)	(0.039)	(0.110)	(0.021)	(0.907)	(0.008)	(0.042)
	[0.202]	[0.813]	[0.222]	[0.185]	[0.180]	[0.942]	[0.267]	[0.425]
GDP p. c.	0.046	-0.050	0.036	0.043	-0.214	-0.779	1.025	0.932
	(0.157)	(0.510)	(0.210)	(0.198)	(0.480)	(0.026)	(0.002)	(0.001)
	[0.345]	[0.692]	[0.363]	[0.385]	[0.629]	[0.237]	[0.121]	[0.365]
F-test (labor):	1.313	0.128	1.725	1.945	4.051	4.854	0.444	3.179
	(0.252)	(0.721)	(0.189)	(0.584)	(0.044)	(0.028)	(0.505)	(0.365)
F-test (all country vars):	3.794	1.618	5.180	5.441	8.418	16.043	10.891	19.774
	(0.285)	(0.655)	(0.159)	(0.364)	(0.038)	(0.001)	(0.012)	(0.001)
N	1154	1227	1154	1154	698	809	406	406

	OLS, study weights							
	Imputed country data				Raw country data			
Response variable:	(1)	(2)	(3)	(4)	(1')	(2')	(3')	(4')
Col. bargaining coverage	0.000	.	.	0.018	0.019	.	.	-0.034
	(0.990)	.	.	(0.478)	(0.653)	.	.	0.551
	[0.992]	.	.	[0.603]	[0.739]	.	.	[0.806]
Strictness of emp. protect.	.	0.291	.	-1.077	.	-2.541	.	6.996
	.	(0.705)	.	(0.203)	.	(0.032)	.	(0.000)
	.	[0.751]	.	[0.366]	.	[0.230]	.	[0.167]
ALMP expenditure	.	.	0.054	0.326	.	.	0.482	-10.538
	.	.	(0.864)	(0.514)	.	.	(0.634)	(0.004)
	.	.	[0.921]	[0.706]	.	.	[0.805]	[0.213]
Product market reg.	3.097	2.753	3.034	3.654	-7.365	7.783	40.504	54.022
	(0.059)	(0.094)	(0.032)	(0.026)	(0.400)	(0.439)	(0.000)	(0.000)
	[0.097]	[0.098]	[0.216]	[0.060]	[0.578]	[0.624]	[0.343]	[0.393]
GDP p. c.	0.015	-0.080	0.013	0.014	0.043	-0.749	1.476	1.026
	(0.651)	(0.386)	(0.675)	(0.663)	(0.892)	(0.001)	(0.000)	(0.000)
	[0.724]	[0.691]	[0.739]	[0.729]	[0.935]	[0.191]	[0.246]	[0.380]
F-test (labor):	0.000	0.143	0.029	2.040	0.203	4.603	0.227	14.307
	(0.990)	(0.705)	(0.864)	(0.564)	(0.653)	(0.032)	(0.634)	(0.003)
F-test (all country vars):	6.202	8.112	7.149	7.545	1.104	92.987	25.775	22.846
	(0.102)	(0.044)	(0.067)	(0.183)	(0.776)	(0.000)	(0.000)	(0.000)
N	1154	1227	1154	1154	698	809	406	406

Notes: We investigate the effects of country-specific variables on elasticity estimates. We employ the set of all explanatory variables used to obtain Table 3 in which we replace the variables *Developing* and *Europe* with the country-specific variables reflecting labor market conditions, product market regulations and the level of economic development. We use the resulting set of explanatory variables to run an OLS estimation (top panel) and the specification in which we use weights based on the inverse of the number of estimates reported in each study (bottom panel). We use imputed values of the country variables (left panel), as well as the raw country-level data with no imputations done (right panel). For brevity, we only present coefficient estimates for the country-specific variables. We report regular p-values and *p*-values from wild bootstrap clustering. We also report results of the F-tests for joint significance of the subset of labor market variables, and for the set of all country variables. A detailed description of all variables is available in Table C5. We only use estimates that remain after we apply the outlier treatment strategy discussed in Section 2 (i.e. cutting 2.5% of outliers from each tail).

Appendix C.4 Best practice

Table C7: Best Practice Estimates
Identified estimates only.

Group	Point Estimate	95% interval	95% interval (wild)	Implied Markdown
Separations: Model				
Linear model	-1.487	[-6.49; 3.51]	[-5.44 ; 11.75]	-
BMA	-0.825	[-5.26; 3.61]	[-6.24; 10.61]	-
LASSO	-1.359	[-6.36; 3.64]	[-5.12; 12.34]	-
Separations: Gender				
Women	-10.654	[-23.02; 1.71]	[-36.06; 17.65]	-
Men	8.014	[-4.61; 20.64]	[-22.34; 38.23]	11.1
Inverse				
Inverse	7.480	[-3.21; 18.17]	[-3.85; 20.51]	11.8

Notes: The table presents fitted 'best practice' estimates using alternative models and data. Estimates in rows 1-3 are obtained using models reported in Table 4, frequentist check in Table C2 and the post-LASSO results of Table C4. The rest of the results are obtained using the linear model. 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 D Publication Bias Robustness Checks

Table D1: Testing for publication bias: robustness to treatment of outliers, all estimates

	OLS	FE	BE	Precision	Study	IV
			Full sample			
SE	1.366	0.225	1.960	1.930	0.555	2.003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.068)	(0.000)
	[0.000]	[0.06]	[0.000]	[0.000]	[0.000]	[0.143]
Constant	1.761	4.278	1.213	0.516	1.821	0.355
	(0.006)	(0.000)	(0.306)	(0.002)	(0.000)	(0.000)
	[0.000]	.	[0.337]	[0.021]	[0.000]	[0.073]
Studies	46	46	46	46	46	46
Observations	1140	1140	46	1140	1140	1140
			Full sample, winsorized at 2%			
SE	1.367	0.228	1.960	1.928	0.555	2.004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.068)	(0.000)
	[0.000]	[0.060]	[0.000]	[0.000]	[0.000]	[0.144]
Constant	1.760	4.271	1.212	0.513	1.821	0.355
	(0.006)	(0.000)	(0.306)	(0.003)	(0.000)	(0.000)
	[0.000]	.	[0.337]	[0.021]	[0.000]	[0.073]
Studies	46	46	46	46	46	46
Observations	1140	1140	46	1140	1140	1140
			Full sample, winsorized at 5%			
SE	1.499	0.258	1.889	1.965	0.768	2.038
	(0.000)	(0.000)	(0.000)	(0.000)	(0.039)	(0.000)
	[0.000]	[0.061]	[0.000]	[0.000]	[0.000]	[0.133]
Constant	1.517	4.220	1.256	0.509	1.613	0.343
	(0.004)	(0.000)	(0.268)	(0.003)	(0.000)	(0.000)
	[0.000]	.	[0.302]	[0.020]	[0.000]	[0.078]
Studies	46	46	46	46	46	46
Observations	1140	1140	46	1140	1140	1140
			Cut sample, 2% of outliers dropped			
SE	1.410	0.321	1.640	1.955	0.558	2.033
	(0.000)	(0.000)	(0.000)	(0.000)	(0.070)	(0.000)
	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]	[0.156]
Constant	1.708	4.088	1.772	0.515	1.822	0.344
	(0.005)	(0.000)	(0.159)	(0.002)	(0.000)	(0.000)
	[0.000]	.	[0.080]	[0.021]	[0.000]	[0.073]
Studies	46	46	46	46	46	46
Observations	1137	1137	46	1137	1137	1137
			Cut sample, 5% of outliers dropped			
SE	1.443	0.400	1.258	1.986	0.562	2.089
	(0.000)	(0.001)	(0.000)	(0.000)	(0.072)	(0.000)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.179]
Constant	1.733	4.009	2.175	0.550	1.837	0.325
	(0.004)	(0.000)	(0.055)	(0.003)	(0.000)	(0.000)
	[0.000]	.	[0.000]	[0.013]	[0.000]	[0.072]
Studies	46	46	46	46	46	46
Observations	1118	1118	46	1118	1118	1118

Notes: This table checks the robustness of the results presented in *Panel A* of Table 2 against the treatment of outliers. We report results obtained for untreated full sample ('*Full sample*'); for full sample where observations are winsorized at 1% at each tail of the distribution ('*Full sample, winsorized at 2%*') and at 2.5% at each tail ('*Full sample, winsorized at 5%*'); for the sample where 1% of outliers is dropped from each tail ('*Cut sample, 2% of outliers dropped*'); for the sample where 2.5% of outliers are dropped from each tail ('*Cut sample, 5% of outliers dropped*') — our preferred treatment reported in *Panel A* of Table 2. We repeat this exercise for our five specifications ('*OLS*', '*FE*', '*BE*', '*Precision*' and '*Study*') and report regular p-values in parenthesis and p-values from wild bootstrap clustering in square brackets, see notes for Table 2 for detailed description. In addition, we report results for the specification in which we use the number of observations to instrument for the standard error ('*IV*'). We do not report this in the main text as the first-stage results are insignificant.

Table D2: Testing for publication bias: robustness to treatment of outliers, published estimates

	OLS	FE	BE	Precision	Study	IV
			Full sample			
SE	1.677	0.220	2.101	2.070	1.772	2.176
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.125]	[0.000]	[0.011]	[0.000]	[0.288]
Constant	1.423	4.698	1.264	0.540	1.101	0.302
	(0.012)	(0.000)	(0.017)	(0.005)	(0.000)	(0.000)
	[0.000]	.	[0.000]	[0.043]	[0.000]	[0.254]
Studies	38	38	38	38	38	38
Observations	1016	1016	38	1016	1016	1016
			Full sample, winsorized at 2%			
SE	1.679	0.226	2.101	2.067	1.772	2.177
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.125]	[0.000]	[0.010]	[0.000]	[0.289]
Constant	1.421	4.687	1.263	0.536	1.101	0.301
	(0.012)	(0.000)	(0.017)	(0.006)	(0.000)	(0.000)
	[0.000]	.	[0.000]	[0.041]	[0.000]	[0.255]
Studies	38	38	38	38	38	38
Observations	1016	1016	38	1016	1016	1016
			Full sample, winsorized at 5%			
SE	1.681	0.235	2.101	2.067	1.774	2.181
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.127]	[0.000]	[0.010]	[0.000]	[0.294]
Constant	1.425	4.677	1.264	0.535	1.103	0.300
	(0.011)	(0.000)	(0.017)	(0.007)	(0.000)	(0.000)
	[0.000]	.	[0.000]	[0.039]	[0.000]	[0.256]
Studies	38	38	38	38	38	38
Observations	1016	1016	38	1016	1016	1016
			Cut sample, 2% of outliers dropped			
SE	1.746	0.367	2.109	2.099	1.807	2.212
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.003]	[0.000]	[0.010]	[0.000]	[0.313]
Constant	1.326	4.397	1.262	0.538	1.080	0.288
	(0.007)	(0.000)	(0.017)	(0.005)	(0.000)	(0.000)
	[0.000]	.	[0.000]	[0.043]	[0.000]	[0.253]
Studies	38	38	38	38	38	38
Observations	1013	1013	38	1013	1013	1013
			Cut sample, 5% of outliers dropped			
SE	1.800	0.491	2.125	2.135	1.832	2.276
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.000]	[0.015]	[0.000]	[0.370]
Constant	1.322	4.231	1.272	0.578	1.083	0.266
	(0.004)	(0.000)	(0.016)	(0.007)	(0.000)	(0.002)
	[0.000]	.	[0.000]	[0.039]	[0.000]	[0.258]
Studies	38	38	38	38	38	38
Observations	995	995	38	995	995	995

Notes: This table checks the robustness of the results presented in *Panel B* of Table 2 against the treatment of outliers. We report results obtained on a subset of published studies with estimates coming from an untreated full sample ('*Full sample*'); from the full sample where observations are winsorized at 1% at each tail of the distribution ('*Full sample, winsorized at 2%*') and at 2.5% at each tail ('*Full sample, winsorized at 5%*'); from the sample where 1% of outliers is dropped from each tail ('*Cut sample, 2% of outliers dropped*'); from the sample where 2.5% of outliers are dropped from each tail ('*Cut sample, 5% of outliers dropped*') — our preferred treatment reported in *Panel B* of Table 2. We repeat this exercise for our five specifications ('*OLS*', '*FE*', '*BE*', '*Precision*' and '*Study*') and report regular p-values in parenthesis and p-values from wild bootstrap clustering in square brackets, see notes for Table 2 for detailed description. In addition, we report results for the specification in which we use the number of observations to instrument for the standard error ('*IV*'). We do not report this in the main text as the first-stage results are insignificant.

Table D3: Testing for publication bias using Andrews and Kasy (2019): robustness to treatment of outliers, all estimates.

$\bar{\theta}$	$\tilde{\tau}$	$\tilde{\nu}$	$\beta_{p,1}$	$\beta_{p,2}$	$\beta_{p,3}$
Full sample					
0.166	1.026	1.645	0.011	0.066	0.161
(0.001)	(0.011)	(0.068)	(0.006)	(0.086)	(0.127)
Full sample, winsorized at 2%					
-0.120	0.698	1.456	0.007	0.040	0.110
(0.020)	(0.043)	(0.078)	(0.004)	(0.044)	(0.068)
Full sample, winsorized at 5%					
-0.114	0.698	1.460	0.005	0.042	0.110
(0.051)	(0.116)	(0.204)	(0.003)	(0.047)	(0.070)
Cut sample, 2% of outliers dropped					
0.158	1.098	1.572	0.011	0.068	0.174
(0.000)	(0.006)	(0.039)	(0.007)	(0.093)	(0.140)
Cut sample, 5% of outliers dropped					
0.157	0.648	1.570	0.005	0.036	0.111
(0.001)	(0.006)	(0.028)	(0.003)	(0.048)	(0.074)

Notes: This table checks the robustness of the results presented in *Panel A* of Table B1 against the treatment of outliers. We report results obtained from an untreated full sample of 'direct' estimates ('*Full sample*'); from the full sample where observations are winsorized at 1% at each tail of the distribution ('*Full sample, winsorized at 2%*') and at 2.5% at each tail ('*Full sample, winsorized at 5%*'); from the sample where 1% of outliers is dropped from each tail ('*Cut sample, 2% of outliers dropped*'); from the sample where 2.5% of outliers are dropped from each tail ('*Cut sample, 5% of outliers dropped*') — our preferred treatment reported in *Panel A* of Table B1. See notes of Table B1 for details regarding the estimated specification.

Table D4: Testing for publication bias using Andrews and Kasy (2019): robustness to treatment of outliers, published estimates

Full sample					
$\bar{\theta}$	$\tilde{\tau}$	$\tilde{\nu}$	$\beta_{p,1}$	$\beta_{p,2}$	$\beta_{p,3}$
-0.314	0.488	1.411	0.004	0.023	0.067
(0.302)	(0.819)	(0.657)	(0.006)	(0.031)	(0.069)
Full sample, winsorized at 2%					
$\bar{\theta}$	$\tilde{\tau}$	$\tilde{\nu}$	$\beta_{p,1}$	$\beta_{p,2}$	$\beta_{p,3}$
-0.314	0.486	1.410	0.004	0.023	0.067
(0.313)	(0.815)	(0.632)	(0.006)	(0.031)	(0.070)
Full sample, winsorized at 5%					
$\bar{\theta}$	$\tilde{\tau}$	$\tilde{\nu}$	$\beta_{p,1}$	$\beta_{p,2}$	$\beta_{p,3}$
-0.300	0.481	1.391	0.003	0.024	0.068
(0.341)	(0.855)	(0.651)	(0.005)	(0.034)	(0.074)
Cut sample, 2% of outliers dropped					
$\bar{\theta}$	$\tilde{\tau}$	$\tilde{\nu}$	$\beta_{p,1}$	$\beta_{p,2}$	$\beta_{p,3}$
-0.314	0.482	1.404	0.004	0.022	0.067
(0.335)	(0.868)	(0.673)	(0.007)	(0.031)	(0.073)
Cut sample, 5% of outliers dropped					
$\bar{\theta}$	$\tilde{\tau}$	$\tilde{\nu}$	$\beta_{p,1}$	$\beta_{p,2}$	$\beta_{p,3}$
-0.269	0.548	1.435	0.002	0.020	0.073
(0.295)	(0.717)	(0.600)	(0.002)	(0.024)	(0.063)

Notes: This table checks the robustness of the results presented in *Panel B* of Table B1 against the treatment of outliers. We report results obtained from an untreated full sample of published 'direct' estimates ('*Full sample*'); from the full sample where observations are winsorized at 1% at each tail of the distribution ('*Full sample, winsorized at 2%*') and at 2.5% at each tail ('*Full sample, winsorized at 5%*'); from the sample where 1% of outliers is dropped from each tail ('*Cut sample, 2% of outliers dropped*'); from the sample where 2.5% of outliers are dropped from each tail ('*Cut sample, 5% of outliers dropped*') — our preferred treatment reported in *Panel B* of Table B1. See notes of Table B1 for details regarding the estimated specification.

Appendix E Heterogeneity Robustness Checks

Appendix E.1 Heterogeneity: model with all regressors, full sample

Table E1: Why do estimates of supply elasticity vary? No outlier treatment.

	OLS, unweighted				OLS, study weights			
Response variable:	Coef.	SE	P-value	P-value (wild)	Coef.	SE	P-value	P-value (wild)
Data Characteristics								
SE non-inverse	0.800	0.463	0.084	0.175	1.020	0.461	0.027	0.052
No obs (log)	0.331	1.006	0.742	0.798	1.125	0.835	0.178	0.237
Midyear of data	-0.056	0.048	0.244	0.243	-0.081	0.068	0.230	0.223
Female share	-14.297	7.239	0.048	0.033	-7.182	5.430	0.186	0.390
F-test (group 1):	5.692	.	0.223	.	6.186	.	0.186	.
Country & Occupation								
Developing	11.679	11.784	0.322	0.502	17.638	11.586	0.128	0.298
Europe	2.859	6.563	0.663	0.732	12.018	9.312	0.197	0.506
Nurses	6.632	36.175	0.855	0.892	-9.231	20.058	0.645	0.746
Teachers	1.645	15.357	0.915	0.918	-6.139	15.083	0.684	0.800
F-test (group 2):	4.099	.	0.393	.	4.093	.	0.394	.
Method & Identification								
Separations, id.	-2.328	8.012	0.771	0.820	1.604	7.312	0.826	0.866
Inverse, id.	8.801	39.047	0.822	0.886	18.832	26.282	0.474	0.815
Inverse, not id.	88.368	53.018	0.096	0.168	38.791	31.608	0.220	0.310
Recruitment, id.	-1.494	7.664	0.845	0.880	-7.776	8.513	0.361	0.435
Recruitment, not id.	-1.970	7.407	0.790	0.816	-14.157	7.777	0.069	0.236
L on W regression, id	-11.833	25.710	0.645	0.717	6.108	17.741	0.731	0.804
Structural & other, id.	-27.833	19.224	0.148	0.110	-26.100	16.122	0.105	0.377
Structural & other, not id.	6.858	7.528	0.362	0.527	-12.875	8.733	0.140	0.178
F-test (group 3):	23.484	.	0.003	.	9.455	.	0.305	.
Estimation Technique								
Hazard	-4.284	5.797	0.460	0.480	-10.635	7.038	0.131	0.177
Probit, logit, other	-6.179	5.317	0.245	0.360	1.267	4.581	0.782	0.801
F-test (group 4):	4.088	.	0.130	.	2.304	.	0.316	.
Publication Characteristics								
Top journal	11.365	7.718	0.141	0.217	3.084	6.215	0.620	0.688
Citations	5.848	5.994	0.329	0.359	4.025	3.186	0.206	0.384
Pub. year (google)	0.569	0.843	0.500	0.687	0.037	0.374	0.920	0.930
NBER or IZA	2.007	6.444	0.755	0.771	9.706	7.205	0.178	0.242
WP other	22.827	30.125	0.449	0.684	3.895	19.483	0.842	0.851
F-test (group 5):	7.361	.	0.195	.	2.388	.	0.793	.
Constant	-13.674	28.052	0.626	0.794	-4.148	12.919	0.748	0.804
N	1320	.	.	.	1320	.	.	.

Notes: Here we repeat the exercise presented in Table 3 using the full sample without implementing any outlier treatment. We present the results of the OLS estimation (left panel) and the specification in which we use weights based on the inverse of the number of estimates reported in each study (right panel). We report regular p-values and *p*-values from wild bootstrap clustering; 'id' denotes estimates obtained with an identification strategy in place. We also report results of the F-test for joint significance for each group of explanatory variables. A detailed description of all variables is available in Table A1.

Table E2: Why do estimates of supply elasticity vary? Outliers winsorized at 1% (each tail).

	OLS, unweighted				OLS, study weights			
Response variable:	Coef.	SE	P-value	P-value (wild)	Coef.	SE	P-value	P-value (wild)
Data Characteristics								
SE non-inverse	0.833	0.397	0.036	0.128	0.884	0.394	0.025	0.080
No obs (log)	0.448	0.534	0.402	0.483	0.640	0.481	0.183	0.259
Midyear of data	-0.038	0.034	0.255	0.339	-0.055	0.049	0.254	0.313
Female share	-10.579	5.562	0.057	0.063	-5.451	3.971	0.170	0.377
F-test (group 1):	6.440	.	0.169	.	5.944	.	0.203	.
Country & Occupation								
Developing	11.453	8.521	0.179	0.331	14.920	9.822	0.129	0.300
Europe	1.530	3.271	0.640	0.709	6.092	4.455	0.172	0.393
Nurses	0.295	13.827	0.983	0.988	-0.978	9.252	0.916	0.942
Teachers	0.121	6.181	0.984	0.987	-0.528	6.890	0.939	0.951
F-test (group 2):	3.270	.	0.514	.	3.790	.	0.435	.
Method & Identification								
Separations, id.	-1.163	6.071	0.848	0.884	1.593	4.261	0.708	0.769
Inverse, id.	11.584	15.953	0.468	0.724	11.818	11.728	0.314	0.668
Inverse, not id.	51.880	23.920	0.030	0.015	28.800	15.176	0.058	0.066
Recruitment, id.	0.049	3.622	0.989	0.989	-6.877	5.826	0.238	0.344
Recruitment, not id.	-0.984	5.893	0.867	0.898	-9.607	5.624	0.088	0.303
L on W regression, id	-7.940	10.913	0.467	0.612	-0.158	8.210	0.985	0.989
Structural & other, id.	-24.591	14.662	0.093	0.112	-21.919	12.227	0.073	0.348
Structural & other, not id.	7.372	5.055	0.145	0.397	-9.719	7.055	0.168	0.224
F-test (group 3):	21.347	.	0.006	.	10.768	.	0.215	.
Estimation Technique								
Hazard	-3.974	3.336	0.233	0.237	-5.718	3.938	0.146	0.226
Probit, logit, other	-6.566	3.978	0.099	0.265	0.552	3.104	0.859	0.882
F-test (group 4):	6.128	.	0.047	.	2.337	.	0.311	.
Publication Characteristics								
Top journal	9.413	4.998	0.060	0.141	2.980	3.281	0.364	0.465
Citations	5.022	4.782	0.294	0.397	3.127	2.347	0.183	0.382
Pub. year (google)	0.338	0.453	0.455	0.618	0.087	0.186	0.641	0.669
NBER or IZA	1.141	4.454	0.798	0.822	6.883	5.785	0.234	0.343
WP other	20.580	15.035	0.171	0.499	7.117	11.434	0.534	0.667
F-test (group 5):	6.675	.	0.246	.	2.640	.	0.755	.
Constant	-9.124	15.079	0.545	0.740	-3.624	7.331	0.621	0.750
N	1320	.	.	.	1320	.	.	.

Notes: Here we repeat the exercise presented in Table 3 using the full sample of elasticity estimates in which we winsorize the outliers in each tail at 1%. We present the results of the OLS estimation (left panel) and the specification in which we use weights based on the inverse of the number of estimates reported in each study (right panel). We report regular p-values and p-values from wild bootstrap clustering; 'id' denotes estimates obtained with an identification strategy in place. We also report results of the F-test for joint significance for each group of explanatory variables. A detailed description of all variables is available in Table A1.

Table E3: Why do estimates of supply elasticity vary? Outliers winsorized at 2.5% (each tail).

	OLS, unweighted				OLS, study weights			
Response variable:	Coef.	SE	P-value	P-value (wild)	Coef.	SE	P-value	P-value (wild)
Data Characteristics								
SE non-inverse	0.873	0.321	0.007	0.016	0.753	0.304	0.013	0.088
No obs (log)	0.435	0.273	0.112	0.201	0.479	0.290	0.098	0.192
Midyear of data	-0.030	0.021	0.159	0.314	-0.035	0.030	0.245	0.358
Female share	-4.944	2.815	0.079	0.166	-2.602	1.944	0.181	0.353
F-test (group 1):	11.241	.	0.024	.	7.536	.	0.110	.
Country & Occupation								
Developing	5.107	4.387	0.244	0.400	7.436	5.259	0.157	0.347
Europe	1.250	1.580	0.429	0.526	3.663	2.256	0.104	0.256
Nurses	-6.891	5.787	0.234	0.398	-2.329	4.572	0.610	0.745
Teachers	-3.149	2.406	0.191	0.289	-0.982	3.250	0.762	0.856
F-test (group 2):	3.558	.	0.469	.	3.834	.	0.429	.
Method & Identification								
Separations, id.	2.002	3.940	0.611	0.709	3.155	2.725	0.247	0.366
Inverse, id.	15.544	7.018	0.027	0.225	10.067	5.650	0.075	0.296
Inverse, not id.	26.712	6.373	0.000	0.006	16.539	5.714	0.004	0.040
Recruitment, id.	2.465	2.021	0.223	0.239	-2.893	2.861	0.312	0.407
Recruitment, not id.	-0.432	3.211	0.893	0.922	-5.681	3.210	0.077	0.306
L on W regression, id	0.266	4.026	0.947	0.960	1.171	4.054	0.773	0.844
Structural & other, id.	-13.208	7.498	0.078	0.091	-12.290	6.419	0.056	0.287
Structural & other, not id.	3.545	2.563	0.167	0.409	-5.591	3.826	0.144	0.204
F-test (group 3):	35.843	.	0.000	.	19.475	.	0.013	.
Estimation Technique								
Hazard	-2.263	2.049	0.269	0.381	-3.345	2.359	0.156	0.260
Probit, logit, other	-2.988	2.133	0.161	0.350	0.879	1.763	0.618	0.678
F-test (group 4):	2.882	.	0.237	.	3.024	.	0.221	.
Publication Characteristics								
Top journal	5.663	2.547	0.026	0.082	1.861	1.650	0.259	0.362
Citations	3.126	2.475	0.207	0.336	1.786	1.468	0.224	0.448
Pub. year (google)	0.201	0.202	0.320	0.490	0.041	0.107	0.704	0.761
NBER or IZA	-0.992	2.679	0.711	0.777	2.858	3.229	0.376	0.490
WP other	6.146	5.852	0.294	0.565	2.053	5.318	0.700	0.779
F-test (group 5):	9.564	.	0.089	.	3.154	.	0.676	.
Constant	-6.785	6.647	0.307	0.532	-2.221	3.770	0.556	0.708
N	1320	.	.	.	1320	.	.	.

Notes: Here we repeat the exercise presented in Table 3 using the full sample of elasticity estimates in which we winsorize the outliers in each tail at 2.5%. We present the results of the OLS estimation (left panel) and the specification in which we use weights based on the inverse of the number of estimates reported in each study (right panel). We report regular p-values and p-values from wild bootstrap clustering; 'id' denotes estimates obtained with an identification strategy in place. We also report results of the F-test for joint significance for each group of explanatory variables. A detailed description of all variables is available in Table A1.

Table E4: Why do estimates of supply elasticity vary? Outliers dropped, 1% (each tail).

	OLS, unweighted				OLS, study weights			
Response variable:	Coef.	SE	P-value	P-value (wild)	Coef.	SE	P-value	P-value (wild)
Data Characteristics								
SE non-inverse	0.865	0.357	0.015	0.086	0.775	0.347	0.025	0.092
No obs (log)	0.241	0.274	0.378	0.359	0.227	0.353	0.519	0.588
Midyear of data	-0.022	0.024	0.346	0.449	-0.036	0.036	0.317	0.407
Female share	-6.483	3.641	0.075	0.185	-4.024	2.889	0.164	0.354
F-test (group 1):	8.566	.	0.073	.	5.962	.	0.202	.
Country & Occupation								
Developing	8.530	6.411	0.183	0.347	11.670	7.585	0.124	0.283
Europe	0.269	1.996	0.893	0.912	3.325	2.613	0.203	0.371
Nurses	1.606	8.106	0.843	0.887	3.578	5.164	0.488	0.597
Teachers	0.678	3.674	0.854	0.873	1.765	4.020	0.661	0.733
F-test (group 2):	2.201	.	0.699	.	3.422	.	0.490	.
Method & Identification								
Separations, id.	-0.001	4.889	1.000	1.000	1.558	3.233	0.630	0.694
Inverse, id.	9.097	10.067	0.366	0.472	10.232	7.518	0.174	0.452
Inverse, not id.	39.779	14.833	0.007	0.002	27.216	10.761	0.011	0.011
Recruitment, id.	0.294	2.642	0.911	0.913	-5.893	4.422	0.183	0.324
Recruitment, not id.	0.206	3.996	0.959	0.969	-6.503	4.412	0.140	0.370
L on W regression, id	-7.123	6.094	0.242	0.364	-3.307	4.480	0.461	0.555
Structural & other, id.	-19.069	11.722	0.104	0.197	-17.115	9.096	0.060	0.299
Structural & other, not id.	5.965	3.751	0.112	0.392	-7.256	5.661	0.200	0.283
F-test (group 3):	23.370	.	0.003	.	17.593	.	0.024	.
Estimation Technique								
Hazard	-2.049	2.161	0.343	0.381	-2.282	2.727	0.403	0.511
Probit, logit, other	-4.872	3.064	0.112	0.332	0.973	2.361	0.680	0.755
F-test (group 4):	3.348	.	0.188	.	1.203	.	0.548	.
Publication Characteristics								
Top journal	7.739	3.514	0.028	0.072	3.161	2.372	0.183	0.284
Citations	3.398	3.519	0.334	0.461	2.349	1.831	0.200	0.380
Pub. year (google)	0.305	0.315	0.333	0.467	0.198	0.170	0.244	0.302
NBER or IZA	-0.287	3.272	0.930	0.942	4.357	4.484	0.331	0.465
WP other	15.751	9.230	0.088	0.418	6.020	7.653	0.431	0.645
F-test (group 5):	6.928	.	0.226	.	4.301	.	0.507	.
Constant	-8.466	10.535	0.422	0.614	-5.358	6.111	0.381	0.562
N	1294	.	.	.	1294	.	.	.

Notes: Here we repeat the exercise presented in Table 3 using the sample of elasticity estimates in which we drop 1% of outliers from each tail. We present the results of the OLS estimation (left panel) and the specification in which we use weights based on the inverse of the number of estimates reported in each study (right panel). We report regular p-values and p-values from wild bootstrap clustering; 'id' denotes estimates obtained with an identification strategy in place. We also report results of the F-test for joint significance for each group of explanatory variables. A detailed description of all variables is available in Table A1.

Appendix E.2 Heterogeneity: model with all regressors, subsample of ‘identified’ estimates

Table E5: Why do estimates of supply elasticity vary? Identified estimates only. No outlier treatment.

	OLS, unweighted				OLS, study weights				OLS, precision weights			
Response variable:	Coef.	SE	P-value	P-value (wild)	Coef.	SE	P-value	P-value (wild)	Coef.	SE	P-value	P-value (wild)
Data Characteristics												
SE non-inverse	0.583	0.370	0.115	0.262	1.515	0.782	0.053	0.169	1.706	0.166	0.000	0.002
No obs (log)	-0.127	0.964	0.895	0.920	1.289	1.193	0.280	0.497	0.226	0.215	0.292	0.488
Midyear of data	-1.961	0.522	0.000	0.059	-0.603	0.701	0.389	0.595	-0.463	0.220	0.035	0.057
Female share	-27.943	14.195	0.049	0.210	-8.841	7.375	0.231	0.317	-6.909	4.845	0.154	0.278
F-test (group 1):	52.325	.	0.000	.	7.784	.	0.100	.	376.247	.	0.000	.
Country & Occupation												
Developing	-7.945	12.109	0.512	0.663	5.824	8.467	0.492	0.640	-0.967	0.803	0.229	0.314
Europe	22.784	17.433	0.191	0.404	18.731	15.762	0.235	0.512	-4.901	3.073	0.111	0.210
Nurses	-45.706	19.252	0.018	0.209	-43.830	20.521	0.033	0.303	-8.190	4.672	0.080	0.167
Teachers	-47.265	19.626	0.016	0.150	-27.899	14.452	0.054	0.350	-3.237	1.854	0.081	0.111
F-test (group 2):	8.337	.	0.080	.	6.752	.	0.150	.	6.216	.	0.184	.
Method & Identification												
Inverse	48.419	19.813	0.015	0.114	38.545	22.920	0.093	0.512	8.333	1.651	0.000	0.001
Recruitment	30.282	18.634	0.104	0.358	3.815	8.360	0.648	0.756	5.359	2.127	0.012	0.046
L on W regression	24.944	18.935	0.188	0.372	31.610	22.955	0.169	0.466	8.481	3.998	0.034	0.078
Structural & other	-15.662	16.578	0.345	0.637	-18.361	12.219	0.133	0.395	-1.450	1.385	0.295	0.576
F-test (group 3):	9.892	.	0.042	.	7.371	.	0.118	.	27.226	.	0.000	.
Estimation Technique												
Probit, logit, other	13.718	16.316	0.400	0.546	7.714	17.181	0.653	0.789	-0.701	1.799	0.697	0.752
F-test (group 4):	0.707	.	0.400	.	0.202	.	0.653	.	0.152	.	0.697	.
Publication Characteristics												
Top journal	11.516	14.103	0.414	0.546	-3.364	11.631	0.772	0.866	-2.079	2.317	0.369	0.551
Citations	10.984	4.539	0.016	0.204	8.902	4.713	0.059	0.378	1.189	1.060	0.262	0.343
Pub. year (google)	1.142	0.732	0.119	0.240	-0.210	1.086	0.846	0.901	0.422	0.174	0.015	0.028
NBER or IZA	0.447	18.949	0.981	0.987	7.819	13.420	0.560	0.701	-3.792	1.677	0.024	0.115
WP other	-10.269	16.798	0.541	0.687	-25.635	21.537	0.234	0.606	-6.335	2.961	0.032	0.129
F-test (group 5):	17.465	.	0.004	.	4.309	.	0.506	.	10.762	.	0.056	.
Constant	126.702	44.051	0.004	0.160	49.280	42.181	0.243	0.444	27.438	16.345	0.093	0.202
N	576	.	.	.	576	.	.	.	576	.	.	.

Notes: Here we repeat the exercise presented in Table 4 under an alternative outlier treatment. As in Table 4, we restrict our analysis to the sub-sample of 'identified' estimates. Unlike in Table 4, here we use all available 'identified' estimates without implementing any outlier treatment. The panel on the left presents the results of the OLS estimation; the middle panel reports results from a weighted specification that uses inverse of the number of estimates per study as weights; the panel on the right reports results from a specification that uses precision weights. We report regular p-values and p-values from wild bootstrap clustering. We also report results of the F-test for joint significance for each group of explanatory variables. A detailed description of all variables is available in Table A1.

Table E6: Why do estimates of supply elasticity vary? Identified estimates only. Outliers winsorized at 1% (each tail).

	OLS, unweighted				OLS, study weights				OLS, precision weights			
Response variable:	Coef.	SE	P-value	P-value (wild)	Coef.	SE	P-value	P-value (wild)	Coef.	SE	P-value	P-value (wild)
Data Characteristics												
SE non-inverse	0.733	0.285	0.010	0.103	1.184	0.544	0.030	0.124	1.697	0.165	0.000	0.002
No obs (log)	0.612	0.801	0.444	0.523	0.754	0.639	0.238	0.357	0.211	0.218	0.333	0.558
Midyear of data	-2.230	0.489	0.000	0.036	-1.026	0.618	0.097	0.354	-0.465	0.222	0.036	0.057
Female share	-39.049	21.094	0.064	0.359	-15.341	11.863	0.196	0.340	-7.081	4.994	0.156	0.284
F-test (group 1):	56.094	.	0.000	.	7.761	.	0.101	.	375.903	.	0.000	.
Country & Occupation												
Developing	3.938	13.642	0.773	0.862	13.757	11.304	0.224	0.525	-1.078	0.787	0.170	0.249
Europe	5.205	9.479	0.583	0.680	7.974	8.751	0.362	0.634	-4.775	3.050	0.117	0.216
Nurses	-19.128	8.483	0.024	0.098	-15.148	9.830	0.123	0.367	-7.628	4.610	0.098	0.199
Teachers	-21.517	6.224	0.001	0.031	-11.131	5.640	0.048	0.241	-3.139	1.845	0.089	0.120
F-test (group 2):	14.031	.	0.007	.	7.770	.	0.100	.	6.718	.	0.152	.
Method & Identification												
Inverse	24.113	8.758	0.006	0.064	20.666	9.494	0.029	0.190	8.590	1.602	0.000	0.000
Recruitment	16.342	8.220	0.047	0.090	0.423	6.159	0.945	0.962	5.199	2.114	0.014	0.053
L on W regression	15.396	10.479	0.142	0.270	15.735	12.683	0.215	0.490	8.283	3.953	0.036	0.080
Structural & other	-17.368	6.562	0.008	0.124	-17.999	7.590	0.018	0.206	-1.392	1.447	0.336	0.633
F-test (group 3):	33.842	.	0.000	.	14.256	.	0.007	.	31.676	.	0.000	.
Estimation Technique												
Probit, logit, other	3.737	10.617	0.725	0.802	3.063	10.599	0.773	0.862	-0.420	1.750	0.810	0.837
F-test (group 4):	0.124	.	0.725	.	0.084	.	0.773	.	0.058	.	0.810	.
Publication Characteristics												
Top journal	6.470	9.348	0.489	0.567	-1.113	7.863	0.887	0.934	-2.000	2.287	0.382	0.559
Citations	9.317	2.630	0.000	0.067	7.314	3.465	0.035	0.313	1.133	1.074	0.291	0.376
Pub. year (google)	1.735	0.448	0.000	0.054	0.643	0.605	0.288	0.463	0.459	0.169	0.006	0.012
NBER or IZA	10.939	11.667	0.348	0.405	13.387	11.426	0.241	0.525	-3.852	1.689	0.023	0.117
WP other	4.653	8.851	0.599	0.685	-5.404	11.696	0.644	0.791	-6.647	2.894	0.022	0.105
F-test (group 5):	30.955	.	0.000	.	6.309	.	0.277	.	13.063	.	0.023	.
Constant	131.463	34.032	0.000	0.064	60.460	35.830	0.092	0.352	26.496	16.539	0.109	0.236
N	576	.	.	.	576	.	.	.	576	.	.	.

Notes: Here we repeat the exercise presented in Table 4 under an alternative outlier treatment. As in Table 4, we restrict our analysis to the sub-sample of 'identified' estimates. Unlike in Table 4, here we winsorize the outliers in each tail at 1% (prior to sub-sampling). The panel on the left presents the results of the OLS estimation; the middle panel reports results from a weighted specification that uses inverse of the number of estimates per study as weights; the panel on the right reports results from a specification that uses precision weights. We report regular p-values and p-values from wild bootstrap clustering. We also report results of the F-test for joint significance for each group of explanatory variables. A detailed description of all variables is available in Table A1.

Table E7: Why do estimates of supply elasticity vary? Identified estimates only. Outliers winsorized at 2.5% (each tail).

	OLS, unweighted				OLS, study weights				OLS, precision weights			
Response variable:	Coef.	SE	P-value	P-value (wild)	Coef.	SE	P-value	P-value (wild)	Coef.	SE	P-value	P-value (wild)
Data Characteristics												
SE non-inverse	0.777	0.269	0.004	0.086	0.976	0.432	0.024	0.161	1.696	0.165	0.000	0.002
No obs (log)	0.792	0.596	0.184	0.334	0.614	0.454	0.177	0.308	0.196	0.208	0.346	0.573
Midyear of data	-1.622	0.342	0.000	0.010	-0.632	0.433	0.145	0.409	-0.449	0.211	0.033	0.055
Female share	-23.499	13.620	0.084	0.498	-9.697	7.849	0.217	0.416	-6.688	4.718	0.156	0.294
F-test (group 1):	79.018	.	0.000	.	7.397	.	0.116	.	365.783	.	0.000	.
Country & Occupation												
Developing	3.177	9.305	0.733	0.836	9.308	7.782	0.232	0.528	-1.057	0.755	0.161	0.264
Europe	0.216	6.328	0.973	0.978	3.852	5.829	0.509	0.738	-4.720	2.920	0.106	0.208
Nurses	-12.150	5.710	0.033	0.112	-9.502	6.381	0.136	0.368	-7.242	4.379	0.098	0.203
Teachers	-12.489	2.948	0.000	0.019	-5.678	3.114	0.068	0.232	-2.886	1.698	0.089	0.116
F-test (group 2):	25.074	.	0.000	.	8.966	.	0.062	.	7.149	.	0.128	.
Method & Identification												
Inverse	15.469	4.861	0.001	0.031	11.924	5.268	0.024	0.147	8.425	1.518	0.000	0.000
Recruitment	9.475	4.049	0.019	0.037	-1.131	4.352	0.795	0.855	4.977	1.986	0.012	0.051
L on W regression	11.073	6.987	0.113	0.234	8.200	8.299	0.323	0.568	7.971	3.757	0.034	0.082
Structural & other	-12.149	3.806	0.001	0.029	-13.667	5.387	0.011	0.161	-1.232	1.399	0.379	0.656
F-test (group 3):	49.266	.	0.000	.	16.178	.	0.003	.	34.844	.	0.000	.
Estimation Technique												
Probit, logit, other	1.945	7.318	0.790	0.846	1.544	7.193	0.830	0.901	-0.372	1.673	0.824	0.847
F-test (group 4):	0.071	.	0.790	.	0.046	.	0.830	.	0.049	.	0.824	.
Publication Characteristics												
Top journal	3.921	6.896	0.570	0.642	-0.203	5.665	0.971	0.985	-1.953	2.184	0.371	0.551
Citations	6.224	1.620	0.000	0.044	4.569	2.350	0.052	0.337	1.076	1.017	0.290	0.370
Pub. year (google)	1.309	0.290	0.000	0.036	0.358	0.393	0.362	0.526	0.455	0.162	0.005	0.009
NBER or IZA	8.499	8.351	0.309	0.387	8.460	8.010	0.291	0.557	-3.806	1.613	0.018	0.111
WP other	2.503	6.620	0.705	0.753	-2.354	7.631	0.758	0.850	-6.539	2.784	0.019	0.104
F-test (group 5):	40.862	.	0.000	.	6.246	.	0.283	.	14.490	.	0.013	.
Constant	90.365	21.001	0.000	0.036	38.945	24.163	0.107	0.372	25.283	15.649	0.106	0.233
N	576	.	.	.	576	.	.	.	576	.	.	.

Notes: Here we repeat the exercise presented in Table 4 under an alternative outlier treatment. As in Table 4, we restrict our analysis to the sub-sample of 'identified' estimates. Unlike in Table 4, here we winsorize the outliers in each tail at 2.5% (prior to sub-sampling). The panel on the left presents the results of the OLS estimation; the middle panel reports results from a weighted specification that uses inverse of the number of estimates per study as weights; the panel on the right reports results from a specification that uses precision weights. We report regular p-values and p-values from wild bootstrap clustering. We also report results of the F-test for joint significance for each group of explanatory variables. A detailed description of all variables is available in Table A1.

Table E8: Why do estimates of supply elasticity vary? Identified estimates only. Outliers are dropped, 1% (each tail).

	OLS, unweighted				OLS, study weights				OLS, precision weights			
Response variable:	Coef.	SE	P-value	P-value (wild)	Coef.	SE	P-value	P-value (wild)	Coef.	SE	P-value	P-value (wild)
Data Characteristics												
SE non-inverse	0.858	0.268	0.001	0.123	0.980	0.453	0.031	0.169	1.723	0.169	0.000	0.004
No obs (log)	0.920	0.774	0.235	0.373	0.114	0.881	0.897	0.937	0.196	0.225	0.382	0.629
Midyear of data	-1.940	0.480	0.000	0.063	-0.973	0.507	0.055	0.286	-0.456	0.218	0.037	0.060
Female share	-30.278	17.301	0.080	0.381	-13.910	10.244	0.174	0.313	-7.119	4.941	0.150	0.285
F-test (group 1):	73.320	.	0.000	.	11.169	.	0.025	.	335.331	.	0.000	.
Country & Occupation												
Developing	5.427	10.506	0.605	0.774	12.083	10.355	0.243	0.575	-1.141	0.779	0.143	0.227
Europe	-1.215	6.960	0.861	0.889	4.361	5.675	0.442	0.698	-4.629	3.027	0.126	0.225
Nurses	-14.012	9.181	0.127	0.180	0.992	10.312	0.923	0.951	-7.187	4.792	0.134	0.260
Teachers	-12.976	4.118	0.002	0.020	-3.324	3.325	0.317	0.393	-3.030	1.859	0.103	0.141
F-test (group 2):	20.170	.	0.000	.	3.583	.	0.465	.	7.463	.	0.113	.
Method & Identification												
Inverse	13.777	8.185	0.092	0.155	19.267	9.005	0.032	0.172	8.688	1.552	0.000	0.000
Recruitment	11.191	5.610	0.046	0.067	-0.715	5.066	0.888	0.932	5.029	2.140	0.019	0.068
L on W regression	12.763	8.003	0.111	0.219	6.474	9.034	0.474	0.654	8.041	3.950	0.042	0.089
Structural & other	-14.703	4.778	0.002	0.032	-12.101	7.366	0.100	0.457	-1.310	1.501	0.383	0.692
F-test (group 3):	43.844	.	0.000	.	20.084	.	0.000	.	33.428	.	0.000	.
Estimation Technique												
Probit, logit, other	0.462	7.329	0.950	0.959	7.664	8.844	0.386	0.584	-0.251	1.817	0.890	0.894
F-test (group 4):	0.004	.	0.950	.	0.751	.	0.386	.	0.019	.	0.890	.
Publication Characteristics												
Top journal	3.586	7.667	0.640	0.700	1.487	5.892	0.801	0.880	-1.922	2.254	0.394	0.552
Citations	7.365	1.887	0.000	0.068	4.866	2.967	0.101	0.426	1.035	1.070	0.333	0.423
Pub. year (google)	1.401	0.621	0.024	0.064	1.237	0.654	0.059	0.157	0.473	0.167	0.005	0.010
NBER or IZA	11.604	9.816	0.237	0.332	9.991	9.835	0.310	0.595	-3.903	1.677	0.020	0.110
WP other	6.620	9.403	0.481	0.556	-5.971	10.319	0.563	0.723	-6.790	2.840	0.017	0.099
F-test (group 5):	26.288	.	0.000	.	8.657	.	0.124	.	12.846	.	0.025	.
Constant	114.787	37.074	0.002	0.149	39.983	37.349	0.284	0.648	25.373	16.814	0.131	0.309
N	562	.	.	.	562	.	.	.	562	.	.	.

Notes: Here we repeat the exercise presented in Table 4 under an alternative outlier treatment. As in Table 4, we restrict our analysis to the sub-sample of 'identified' estimates. Unlike in Table 4, here we drop 1% of outliers from each tail (prior to sub-sampling). The panel on the left presents the results of the OLS estimation; the middle panel reports results from a weighted specification that uses inverse of the number of estimates per study as weights; the panel on the right reports results from a specification that uses precision weights. We report regular p-values and p-values from wild bootstrap clustering. We also report results of the F-test for joint significance for each group of explanatory variables. A detailed description of all variables is available in Table A1.

Appendix E.3 Heterogeneity and model uncertainty: outlier treatments in BMA

Table E9: Why do estimates of supply elasticity vary?
Bayesian Model Averaging, outliers dropped, 1% (each tail).

	BMA			OLS with selected variables			
Response variable:	Post. Mean	Post. SD	PIP	Coef.	SE	P-value	P-value (wild)
Data Characteristics							
SE non-inverse	0.883	0.127	1.000	0.888	0.301	0.003	0.041
No obs (log)	0.002	0.034	0.032				
Midyear of data	-0.003	0.011	0.076				
Female share	-6.153	1.467	0.997	-6.272	3.614	0.083	0.227
Country & Occupation							
Developing	7.918	1.868	0.999	7.587	6.900	0.272	0.395
Europe	0.027	0.238	0.039				
Nurses	0.045	0.589	0.035				
Teachers	0.065	0.429	0.045				
Method & Identification							
Separations, id.	0.056	0.566	0.036				
Inverse, id.	11.292	2.295	0.998	12.115	7.159	0.091	0.189
Inverse, not id.	41.837	2.500	1.000	42.660	15.486	0.006	0.045
Recruitment, id.	0.054	0.416	0.040				
Recruitment, not id.	-0.022	0.837	0.028				
L on W regression, id	-7.614	2.897	0.942	-7.781	4.383	0.076	0.179
Structural & other, id.	-17.188	3.321	1.000	-18.492	10.220	0.070	0.143
Structural & other, not id.	1.830	2.812	0.347				
Estimation Technique							
Hazard	-0.081	0.430	0.059				
Probit, logit, other	-1.760	2.310	0.433				
Publication Characteristics							
Top journal	8.548	1.590	0.998	8.444	3.183	0.008	0.053
Citations	1.992	1.977	0.573	2.983	2.908	0.305	0.428
Pub. year (google)	0.327	0.095	0.982	0.331	0.220	0.133	0.204
NBER or IZA	-0.084	0.560	0.049				
WP other	14.672	2.378	1.000	14.149	9.006	0.116	0.391
Constant	-8.646		1.000	-9.468	7.536	0.209	0.293
N	1294	.	.		.	.	.

Notes: Here we repeat the exercise presented in Table C1 using the sample of elasticity estimates in which we drop 1% of outliers from each tail. 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 E10: Why do estimates of supply elasticity vary?
BMA, identified estimates only, outliers dropped, 1% (each tail).

	BMA			OLS with selected variables			
Response variable:	Post. Mean	Post. SD	PIP	Coef.	SE	P-value	P-value (wild)
Data Characteristics							
SE non-inverse	0.794	0.147	1.000	0.822	0.224	0.000	0.082
No obs (log)	0.674	0.564	0.660	1.038	0.649	0.109	0.202
Midyear of data	-1.762	0.189	1.000	-1.804	0.336	0.000	0.011
Female share	-32.718	3.911	1.000	-31.704	14.916	0.034	0.093
Country & Occupation							
Developing	0.599	2.031	0.119				
Europe	-0.930	2.331	0.186				
Nurses	-0.732	3.111	0.106				
Teachers	-5.873	3.214	0.874	-6.249	1.190	0.000	0.017
Method & Identification							
Inverse	16.419	3.486	0.998	16.447	8.450	0.052	0.180
Recruitment	1.191	3.920	0.163				
L on W regression	8.025	4.198	0.870	9.610	4.512	0.033	0.107
Structural & other, id.	-13.911	4.072	0.972	-14.417	5.054	0.004	0.145
Estimation Technique							
Probit, logit, other	0.242	1.600	0.068				
Publication Characteristics							
Top journal	0.376	1.596	0.100				
Citations	6.345	1.478	0.994	6.471	1.910	0.001	0.049
Pub. year (google)	2.100	0.226	1.000	2.124	0.417	0.000	0.029
NBER or IZA	10.004	6.125	0.798	13.265	6.020	0.028	0.231
WP other	-0.063	1.789	0.070				
Constant	83.257		1.000	81.287	24.511	0.001	0.031
N	562	.	.	562	.	.	.

Notes: Here we repeat the exercise presented in Table C2 using the sample of elasticity estimates in which we drop 1% of outliers from each tail. 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.

Appendix E.4 Heterogeneity and model uncertainty: outlier treatments in LASSO

Table E11: Why do estimates of supply elasticity vary? LASSO, outliers dropped, 1% (each tail).

	LASSO	OLS using selected variables			
Response variable:	Coef.	Coef.	SE	P-value	P-value (wild)
Data Characteristics					
SE non-inverse	0.825	0.876	0.352	0.013	0.075
No obs (log)	0.140	0.258	0.268	0.336	0.354
Midyear of data	-0.004	-0.018	0.019	0.355	0.410
Female share	-5.462	-6.392	3.443	0.063	0.132
Country & Occupation					
Developing	8.801	8.557	6.123	0.162	0.306
Europe	0.000	0.000	.	.	.
Nurses	0.000	0.000	.	.	.
Teachers	0.000	0.000	.	.	.
Method & Identification					
Separations, id.	1.071	-0.256	4.851	0.958	0.967
Inverse, id.	8.753	9.480	8.182	0.247	0.324
Inverse, not id.	38.976	39.588	15.101	0.009	0.000
Recruitment, id.	0.000	0.000	.	.	.
Recruitment, not id.	0.000	0.000	.	.	.
L on W regression, id	-5.680	-6.316	4.096	0.123	0.247
Structural & other, id.	-16.623	-19.049	11.335	0.093	0.152
Structural & other, not id.	4.069	6.199	3.254	0.057	0.447
Estimation Technique					
Hazard	-1.540	-2.373	1.765	0.179	0.249
Probit, logit, other	-4.344	-5.244	3.060	0.087	0.303
Publication Characteristics					
Top journal	6.441	7.595	3.301	0.021	0.070
Citations	2.786	3.521	2.887	0.223	0.366
Pub. year (google)	0.234	0.275	0.276	0.320	0.374
NBER or IZA	0.000	0.000	.	.	.
WP other	14.936	15.315	8.766	0.081	0.384
Constant	-6.342	-7.673	8.687	0.377	0.456
N	1294	.	.	.	.

Notes: Here we repeat the exercise presented in Table C3 using the sample of elasticity estimates in which we drop 1% of outliers from each tail. The left panel presents estimates obtained using LASSO with the penalty value selected to minimize mean-squared prediction error through cross-validation. We implement this in `STATA` using the `cvlasso` routine. Variables with zero coefficient values are excluded under the optimal penalty parameter value.The right panel shows results of estimating the OLS using the subset of variables selected by LASSO. We report regular p-values and p-values from wild bootstrap clustering; 'id' denotes estimates obtained with an identification strategy in place. A detailed description of all variables is available in Table A1.

Table E12: Why do estimates of supply elasticity vary? LASSO
Identified estimates only, outliers dropped, 1% (each tail).

	LASSO	OLS using selected variables			
Response variable:	Coef.	Coef.	SE	P-value	P-value (wild)
Data Characteristics					
SE non-inverse	0.886	0.858	0.268	0.001	0.123
No obs (log)	1.029	0.920	0.774	0.235	0.373
Midyear of data	-1.643	-1.940	0.480	0.000	0.063
Female share	-28.483	-30.278	17.301	0.080	0.381
Country & Occupation					
Developing	6.623	5.427	10.506	0.605	0.774
Europe	-0.633	-1.215	6.960	0.861	0.889
Nurses	-8.928	-14.012	9.181	0.127	0.180
Teachers	-8.360	-12.976	4.118	0.002	0.020
Method & Identification					
Inverse	10.794	13.777	8.185	0.092	0.155
Recruitment	2.269	11.191	5.610	0.046	0.067
L on W regression	8.602	12.763	8.003	0.111	0.219
Structural & other	-16.008	-14.703	4.778	0.002	0.032
Estimation Technique					
Probit, logit, other	0.456	0.462	7.329	0.950	0.959
Publication Characteristics					
Top journal	0.212	3.586	7.667	0.640	0.700
Citations	6.163	7.365	1.887	0.000	0.068
Pub. year (google)	1.402	1.401	0.621	0.024	0.064
NBER or IZA	11.656	11.604	9.816	0.237	0.332
WP other	2.444	6.620	9.403	0.481	0.556
Constant	92.891	114.787	37.074	0.002	0.149
N	562	.	.	.	.

Notes: Here we repeat the exercise presented in Table C4 using the sample of elasticity estimates in which we drop 1% of outliers from each tail. The left panel presents estimates obtained using LASSO with the penalty value selected to minimize mean-squared prediction error through cross-validation. We implement this in `STATA` using the `cvlasso` routine. Variables with zero coefficient values are excluded under the optimal penalty parameter value.The right panel shows results of estimating the OLS using the subset of variables selected by LASSO. We report regular p-values and *p*-values from wild bootstrap clustering; 'id' denotes estimates obtained with an identification strategy in place. A detailed description of all variables is available in Table A1.

Appendix E.5 Heterogeneity and country-specific variables: outlier treatments

Table E13: Why do estimates of supply elasticity vary?
Country-specific variables, outliers dropped, 1% (each tail).

	OLS, unweighted							
	Imputed country data				Raw country data			
Response variable:	(1)	(2)	(3)	(4)	(1')	(2')	(3')	(4')
Col. bargaining coverage	-0.010	.	.	0.052	0.053	.	.	-0.061
	(0.697)	.	.	(0.455)	(0.056)	.	.	(0.231)
	[0.804]	.	.	[0.602]	[0.106]	.	.	[0.620]
Strictness of emp. protect.	.	0.220	.	-1.629	.	-3.081	.	5.683
	.	(0.871)	.	(0.416)	.	(0.014)	.	(0.109)
	.	[0.892]	.	[0.600]	.	[0.404]	.	[0.564]
ALMP expenditure	.	.	-0.209	-0.066	.	.	0.454	-6.606
	.	.	(0.557)	(0.898)	.	.	(0.492)	(0.182)
	.	.	[0.699]	[0.910]	.	.	[0.591]	[0.607]
Product market reg.	2.471	0.993	2.386	2.617	-28.098	-2.344	28.415	40.683
	(0.077)	(0.785)	(0.020)	(0.117)	(0.027)	(0.860)	(0.008)	(0.044)
	[0.183]	[0.828]	[0.188]	[0.207]	[0.182]	[0.912]	[0.268]	[0.422]
GDP p. c.	0.025	-0.123	0.022	0.029	-0.229	-1.081	1.023	0.931
	(0.589)	(0.326)	(0.581)	(0.516)	(0.465)	(0.022)	(0.002)	(0.001)
	[0.696]	[0.443]	[0.661]	[0.644]	[0.618]	[0.236]	[0.123]	[0.366]
F-test (labor):	0.152	0.026	0.344	1.561	3.639	6.080	0.472	3.129
	0.697	0.871	0.557	0.668	0.056	0.014	0.492	0.372
F-test (all country vars):	4.007	1.927	5.614	6.835	7.261	13.046	10.945	19.396
	0.261	0.588	0.132	0.233	0.064	0.005	0.012	0.002
N	1174	1264	1174	1174	706	833	412	412

	OLS, study weights							
	Imputed country data				Raw country data			
Response variable:	(1)	(2)	(3)	(4)	(1')	(2')	(3')	(4')
Col. bargaining coverage	0.022	.	.	0.044	0.019	.	.	-0.033
	(0.521)	.	.	(0.246)	(0.653)	.	.	(0.553)
	[0.701]	.	.	[0.449]	[0.739]	.	.	[0.805]
Strictness of emp. protect.	.	0.114	.	-2.053	.	-3.548	.	7.013
	.	(0.921)	.	(0.130)	.	(0.012)	.	(0.000)
	.	[0.927]	.	[0.313]	.	[0.192]	.	[0.167]
ALMP expenditure	.	.	0.417	0.839	.	.	0.480	-10.578
	.	.	(0.419)	(0.280)	.	.	(0.636)	(0.004)
	.	.	[0.682]	[0.552]	.	.	[0.805]	[0.216]
Product market reg.	3.512	4.984	3.791	4.711	-7.027	9.269	40.549	54.145
	(0.116)	(0.060)	(0.034)	(0.032)	(0.414)	(0.440)	(0.000)	(0.000)
	[0.181]	[0.062]	[0.193]	[0.072]	[0.584]	[0.633]	[0.343]	[0.392]
GDP p. c.	0.003	-0.176	0.002	-0.000	0.047	-1.024	1.477	1.027
	(0.951)	(0.263)	(0.970)	(0.998)	(0.882)	(0.000)	(0.000)	(0.000)
	[0.961]	[0.593]	[0.974]	[0.998]	[0.928]	[0.132]	[0.247]	[0.379]
F-test (labor):	0.413	0.010	0.653	2.507	0.203	6.250	0.224	14.399
	(0.521)	(0.921)	(0.419)	(0.474)	(0.653)	(0.012)	(0.636)	(0.002)
F-test (all country vars):	5.453	5.512	6.526	6.961	1.066	108.827	25.712	22.884
	(0.141)	(0.138)	(0.089)	(0.224)	(0.785)	(0.000)	(0.000)	(0.000)
N	1174	1264	1174	1174	706	833	412	412

Notes: Here we repeat the exercise presented in Table C6 using the sample of elasticity estimates in which we drop 1% of outliers from each tail. We investigate the effects of country-specific variables on elasticity estimates. We employ the set of all explanatory variables used to obtain Table 3 in which we replace the variables *Developing* and *Europe* with the country-specific variables reflecting labor market conditions, product market regulations and the level of economic development. We use the resulting set of explanatory variables to run an OLS estimation (top panel) and the specification in which we use weights based on the inverse of the number of estimates reported in each study (bottom panel). We use imputed values of the country variables (left panel), as well as the raw country-level data with no imputations done (right panel). For brevity, we only present coefficient estimates for the country-specific variables. We report regular p-values and p-values from wild bootstrap clustering. We also report results of the F-tests for joint significance of the subset of labor market variables, and for the set of all country variables. A detailed description of all variables is available in Table C5.

Appendix E.6 Heterogeneity and best practice: outlier treatments

Table E14: Best Practice Estimates
outliers dropped, 1% (each tail).

Group	Point Estimate	95% interval	95% interval (wild)	Implied Markdown
Separations: Model				
Linear model	11.995	[4.37; 19.62]	[0.90; 22.18]	7.7
BMA	10.020	[4.21; 15.83]	[1.59; 17.27]	9.1
LASSO	11.825	[4.58; 19.07]	[1.26; 21.65]	7.8
Separations: Gender				
Women	8.811	[2.01; 15.61]	[-0.23; 17.49]	10.2
Men	15.294	[5.48; 25.11]	[0.52; 28.28]	6.1
Separations vs. Inverse				
Separations - Not identified	11.995	[3.48; 20.51]	[-0.21; 23.43]	7.7
Separations - Identified	11.994	[3.36; 20.63]	[0.87; 22.01]	7.7
Inverse - Not identified	51.774	[19.62; 83.93]	[20.59; 109.68]	1.9
Inverse - Identified	21.091	[0.23; 41.96]	[-18.68; 46.80]	4.5

Notes: Here we repeat the exercise presented in Table 5 using the sample of elasticity estimates in which we drop 1% of outliers from each tail. Estimates in rows 1-3 are obtained using models reported in Table E4, frequentist check in Table E9 and the post-LASSO results of Table E11. The rest of the results are obtained using the linear model. 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 E15: Best Practice Estimates
Identified estimates only; outliers dropped, 1% (each tail).

Group	Point Estimate	95% interval	95% interval (wild)	Implied Markdown
Separations: Model				
Linear model	-1.780	[-8.57; 5.02]	[-11.29; 6.83]	-
BMA	-0.957	[-5.99; 4.07]	[-10.41; 6.21]	-
LASSO	0.615	[-4.89; 6.12]	[-9.66; 10.95]	61.9
Separations: Gender				
Women	-16.647	[-36.20; 2.91]	[-59.50; 19.09]	-
Men	13.630	[-3.19; 30.45]	[-33.09; 49.21]	6.8
Inverse				
Inverse	11.997	[-2.42; 26.41]	[-5.88; 34.13]	7.7

Notes: Here we repeat the exercise presented in Table C7 using the sample of elasticity estimates in which we drop 1% of outliers from each tail. Estimates in rows 1-3 are obtained using models reported in Table E8, frequentist check in Table E10 and the post-LASSO results of Table E12. The rest of the results are obtained using the linear model. 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 F 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")*

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 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. The specific survey pieces whose references we checked were Boal and Ransom (1997) and Manning (2011). We first ran the search on November 12th 2017, saved the .html files for the first 100 pages listed and downloaded the .pdf files, when available, for the first 50 pages covering 500 papers. After receiving a request for revisions from this journal, we updated our search by repeating the Google scholar queries on May 12th 2019, focusing on papers from 2017, 2018 and 2019. We downloaded the first 100 papers from each year's search. In addition, we also attempted to find relevant unpublished papers searching the NBER and IZA working paper series websites. For NBER, we used our Google Scholar approach and screened 70 papers posted over the last three years. As IZA's Discussion Papers were not feasibly searchable using our search terms approach, we instead screened using JEL codes, focusing on the J42 code for monopsony, screening the first 100 hits and again studying the references of relevant papers. We obtained 797 from 38 studies during our first search, and 523 estimates from 15 studies in our second search.

Papers in Study

Bachmann, Ronald and Hanna Frings, "Monopsonistic competition, low-wage labour markets, and minimum wages–An empirical analysis," *Applied Economics*, 2017, *49* (51), 5268–5286.

Barth, Erling and Harald Dale-Olsen, "Monopsonistic discrimination, worker turnover, and the gender wage gap," *Labour Economics*, 2009, *16* (5), 589–597.

Blau, Francine D and Lawrence M Kahn, "Race and sex differences in quits by young workers," *ILR Review*, 1981, *34* (4), 563–577.

Bó, Ernesto Dal, Frederico Finan, and Martín A Rossi, "Strengthening state capabilities: The role of financial incentives in the call to public service," *The Quarterly Journal of Economics*, 2013, *128* (3), 1169–1218.

Bodah, Matthew, John Burkett, and Leonard Lardaro, "IX. EMPLOYMENT RELATIONS FOR HEALTH CARE WORKERS," in "Proceedings of the Annual Meeting–Industrial Relations Research Association" IRRA 2003, p. 199.

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

Brochu, Pierre and David A Green, "The impact of minimum wages on labour market transitions," *The Economic Journal*, 2013, *123* (573), 1203–1235.

Caldwell, Sydnee and Emily Oehlsen, "Monopsony and the Gender Wage Gap: Experimental Evidence from the Gig Economy," *Working Paper*, 2018, *MIT.*

Cho, David, "The Labor Market Effects of Demand Shocks: Firm-Level Evidence from the Recovery Act," *Working Paper*, 2018, *Princeton University.*

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

den Berg, Gerard J Van and Geert Ridder, "An empirical equilibrium search model of the labor market," *Econometrica*, 1998, *66* (5), 1183–1221.

Depew, Briggs and Todd A Sørensen, "The elasticity of labor supply to the firm over the business cycle," *Labour Economics*, 2013, *24*, 196–204.

—, Peter Norlander, and Todd A Sørensen, "Inter-firm mobility and return migration patterns of skilled guest workers," *Journal of Population Economics*, 2017, *30* (2), 681–721.

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

Dobelaere, Sabien and Mark Vancauteren, "Market imperfections and total factor productivity," *Working Paper*, 2017, *Vrije Universiteit Amsterdam.*

Dube, Arindrajit, Alan Manning, and Suresh Naidu, "Monopsony and Employer Mis-optimization Explain Why Wages Bunch at Round Numbers," Working Paper 24991, National Bureau of Economic Research September 2018.

—, Jeff Jacobs, Suresh Naidu, and Siddharth Suri, "Monopsony in Online Labor Markets," *American Economic Review: Insights*, 2018, *(forthcoming).*

—, Laura Giuliano, and Jonathan Leonard, "Fairness and frictions: The impact of unequal raises on quit behavior," *American Economic Review*, 2019, *109* (2), 620–63.

—, T William Lester, and Michael Reich, "Minimum wage shocks, employment flows, and labor market frictions," *Journal of Labor Economics*, 2016, *34* (3), 663–704.

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

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

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

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

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

Galizzi, Monica, "Gender and Labor Attachment: Do Within-Firms' Relative Wages Matter?," *Industrial Relations: A Journal of Economy and Society*, 2001, *40* (4), 591–619.

Gittings, R Kaj and Ian M Schmutte, "Getting handcuffs on an octopus: Minimum wages, employment, and turnover," *ILR Review*, 2016, *69* (5), 1133–1170.

Hirsch, Boris and Elke J Jahn, "Is there monopsonistic discrimination against immigrants?," *ILR Review*, 2015, *68* (3), 501–528.

— , — , **Alan Manning, and Michael Oberfichtner**, "The urban wage premium in imperfect labour markets," *CEP Discussion Paper*, 2019, *No 1608.*

— , **Elke J. Jahn, and Claus Schnabel**, "Do Employers Have More Monopsony Power in Slack Labor Markets?," *ILR Review*, 2018, *71* (3), 676–704.

— , **Thorsten Schank, and Claus Schnabel**, "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*, 2010, *28* (2), 291–330.

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

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

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

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar, "Who profits from patents? Rent-sharing at innovative firms," *The Quarterly Journal of Economics*, 2019, *134* (3), 1343–1404.

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

Mas, Alexandre, "Does transparency lead to pay compression?," *Journal of Political Economy*, 2017, *125* (5), 1683–1721.

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

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

Naidu, Suresh, Yaw Nyarko, and Shing-Yi Wang, "Monopsony power in migrant labor markets: evidence from the United Arab Emirates," *Journal of Political Economy*, 2016, *124* (6), 1735–1792.

Ogloblin, Constantin and Gregory Brock, "Wage determination in urban Russia: Underpayment and the gender differential," *Economic Systems*, 2005, *29* (3), 325–343.

Pörtner, Claus C and Nail Hassairi, "What Labor Supply Elasticities do Employers Face? Evidence from Field Experiments," *Working Paper*, 2018, *SSRN.*

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

— **and Ronald L Oaxaca**, "New market power models and sex differences in pay," *Journal of Labor Economics*, 2010, *28* (2), 267–289.

Staiger, Douglas O, Joanne Spetz, and Ciaran S Phibbs, "Is there monopsony in the labor market? Evidence from a natural experiment," *Journal of Labor Economics*, 2010, *28* (2), 211–236.

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

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

Tortarolo, Dario and Roman D Zarate, "Measuring Imperfect Competition in Product and Labor Markets. An Empirical Analysis using Firm-level Production Data," *Working paper*, 2018, *SSRN.*

Tucker, Lee, "Monopsony for whom? evidence from Brazilian administrative data," *Working Paper*, 2017, *Available at http://leetucker.net/docs/LeeTucker_JMP_latest.pdf.*

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

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

Webber, Douglas A, "Firm market power and the earnings distribution," *Labour Economics*, 2015, *35*, 123–134.

— , "Firm-Level Monopsony and the Gender Pay Gap," *Industrial Relations: A Journal of Economy and Society*, 2016, *55* (2), 323–345.

— , "Employment Adjustment Over the Business Cycle: The Impact of Competition in the Labor Market," *IZA Discussion Paper*, 2018, *11887.*

Appendix G Further Notes

Appendix G.1 Standard Errors

We use the delta method to approximate standard errors when the exact estimate is not available, assuming independence of parameters; this strategy is common in the meta-analyses literature, see, for example, Cavlovic et al. (2000), Havranek (2015), Havranek and Sokolova (2020); it was also employed in a labor meta-analysis context close to ours, see Evers et al. (2006).

Appendix G.2 Weighting

We explore two alternative weighting strategies to further check robustness of these results. We weighted 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 and Doucouliagos 2015 for a discussion). For our data, precision weighting yielded strong evidence for publication bias that is very similar to the OLS results. It is worth noting that this technique is not without some caveats. It is possible that some estimation methods 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 importance to studies that use IV. Furthermore, Lewis and Linzer (2005) show that for models with an estimated dependent variable, a simple OLS would often outperform the weighted estimation.[5]

We also weight some of our analysis by the inverse of the number of estimates obtained from the associated paper. This strategy essentially gives all papers equal weight, rather than assigning more weight to papers which estimated more parameters.

[5] For additional discussion of precision weights, see section 4.1 of Card et al. (2018b).

Appendix G.3 Outliers

In our main results, we follow the strategy discussed in Section 2 and use a data set in which we cut 2.5% of outliers from each tail. We also check the robustness of our treatment of outliers. We repeat the exercises of Table 3 and Table 4 under alternative outlier treatments, i.e. no outlier treatment (see Table E1 for the mixed sample and Table E5 for the subsample of identified estimates), outliers winsorized at 1% in each tail (see Table E2 and Table E6), outliers winsorized at 2.5% in each tail (see Table E3 and Table E7) and outliers cut at 1% from each tail (see Table E4 and Table E8).

On the one hand, inclusion of additional outliers introduces extra noise which then makes the overall results less precise. We see increases in magnitudes of some of the estimated coefficients (especially the coefficients for unidentified estimates converted from inverse elasticities, e.g. see Table E1); other effects can no longer be estimated precisely: for example, occupation effects disappear when we consider the full untreated sample of both identified and unidentified estimates.

On the other hand, comparing point estimates we still see some of the same patterns even in the untreated sample, namely some differences across method and identification choices, as well as the positive correlation between estimates and their standard errors consistent with selective reporting. For the subsample of identified estimates, we see strong negative effects for nurses and teachers in the untreated sample. We also observe a negative effect associated with the female share that becomes somewhat more pronounced in some specifications, albeit not always significant. Finally, comparing our baseline results in which we cut the outliers with those obtained on a sample where the previously cut outliers are winsorized (see Table E3 and Table E7), we do not observe much of a difference aside from changes in magnitudes of some of the point estimates. We therefore conclude that the results reported in Table 3 and Table 4 are broadly consistent with those obtained under alternative outlier treatments.

Robustness checks to some outlier treatments are shown regarding BMA and LASSO in subsection Appendix E.3 and subsection Appendix E.4, respectively. For an examination of outliers and best practice estimates see subsection Appendix E.6. For our cross-country analysis and outliers see subsection Appendix E.5.

Appendix G.4 Heterogeneity and Model Uncertainty

The results reported in Table 3 and Table 4 were obtained under an assumption that all of the 23 explanatory variables belong to the 'true' data generating process. It is possible that some of the 23 explanatory variables don't contribute to the observed variation in supply elasticity estimates in a meaningful way, although we have some intuition for why each of the 23 variables might contribute to determining the magnitude of elasticity estimates. In this section, we discuss our systematic addressing of model selection with two methods designed to mitigate model uncertainty: Bayesian Model Averaging (BMA) and LASSO.

The BMA methodology tackles model uncertainty by explicitly modeling and estimating probabilities that different combinations of explanatory variables represent the 'true' model. In our case, there are 2^{23} distinct variable combinations (or models) that include one or more of our 23 variables. In the main analysis, we have only estimated a tiny fraction of this model space. The BMA approach is radically different compared to what we have done in previous sections: instead of picking one specific model, BMA approximates the entire model space, assigning each of the 2^{23} possible models a metric—Posterior Model Probability—that reflects the likelihood of it being the 'true' model. It then averages parameter estimates across all models, using posterior model probabilities as weights (see subsection Appendix C.1 for more details about BMA).

LASSO provides a very different solution to the model uncertainty problem. Assuming that the 'true' model is sparse (i.e. there is only a handful of explanatory variables that have a non-zero effect on the dependent variable), LASSO amends the OLS minimization problem by introducing an extra constraint on the sum of absolute values of regression coefficients. This amended minimization problem typically yields corner solutions that assign exact zeros to coefficients on some of the less relevant explanatory variables. As a consequence, the less relevant variables get automatically excluded achieving sparsity (see subsection Appendix C.2 for more details about LASSO).

The results of BMA and LASSO estimations are presented in Appendix C.1 and Appendix C.2, along with detailed discussions. We show some robustness checks to our outlier treatments in subsection Appendix E.3 and subsection Appendix E.4. In line with the OLS results reported earlier, both of these approaches detect positive and significant correlation between 'direct' estimates and their standard errors indicative of publication bias. The effect of having estimates converted from inverse elasticities remains positive and significant in all specifications. As before, point estimates associated with 'identified' inverse elasticities are smaller than those corresponding to 'not identified', underscoring the importance of having an identification strategy. Furthermore, both methods evaluate the effect of structural identified estimates to be negative—when compared to estimates from the separations-based approach.

Once again, we find suggestive evidence that studies that analyze data with higher shares of female workers produce more evidence of monopsony power. The results regarding effects of occupation are mixed: BMA provides some evidence linking the market of nurses to higher degrees of monopsony power—but not the market for teachers; LASSO results suggest stronger negative effects for both—especially in the subset of the 'identified' estimates. The two methods also generate conflicting results with regard to recruitment-based estimates: according to BMA, results obtained using recruitments are not different from those obtained using separations, as the probability of the regressor for recruitment-based estimates belonging to the 'true' model is estimated to be below 7%. At the same time, LASSO reports a positive significant effect associated with using recruitment elasticities.

Best practice estimates, including results from our preferred BMA results, are given below.

Table G1: Best Practice Estimates With BMA Included

Group	Point Estimate	95% interval	95% interval (wild)	Implied Markdown
Separations: Model				
Linear model	7.133	[1.75; 12.51]	[-0.88; 15.07]	12.3
BMA	5.738	[2.46; 9.02]	[1.03; 10.52]	14.8
LASSO	7.177	[2.37; 11.99]	[0.41; 13.78]	12.2
Separations: Gender				
Women	5.971	[1.09; 10.86]	[-0.90; 13.13]	14.3
Men	8.336	[1.98; 14.70]	[-1.19; 17.52]	10.7
Separations vs. Inverse				
Separations - Not identified	6.429	[1.00; 11.85]	[-1.39; 14.29]	13.5
Separations - Identified	9.910	[2.08; 17.74]	[-0.89; 19.17]	9.2
Inverse - Not identified	24.674	[19.33; 30.02]	[14.61; 31.24]	3.9
Inverse - Identified	22.810	[8.29; 37.33]	[1.58; 50.09]	4.2

Notes: The table presents fitted 'best practice' estimates using alternative models and data. Estimates in rows 1-3 are obtained using models reported in Table 3, frequentist check in Table C1 and the post-LASSO results of Table C3. The rest of the results are obtained using the linear model. 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 G.5 Additional Results on Gender

In Table G2, we explore four alternative treatments to gender, expanding upon our preferred results given in Table 5. As with Table 5, this table also presents confidence intervals and markdowns. In addition the fourth and fifth columns present p-values on a test of significance for a gender effect in the specification.

The first panel of the table presents our baseline gender based best practice estimates that were presented in Table 5. While the gender effect is not statistically significant, we see that women face a larger wage markdown (14.3%) than men (10.7%). As this corresponds to women and men being paid 91.3 and 85.7 for each dollar they produce, respectively, it implies a 4.0% gender wage gap resulting from differential monopsony power.

By including direct regressors for two female dominated occupations our results may understate the effects of gender.[6] Specifically, nurses and teachers accounted for the employment of around one in eight women in the US as of 2018.[7] We address this concern in two ways. First, in the "Reduced Form" specification, we drop covariants for these occupations, allowing the female share variable to pick up any measurable effect through these channels. Overall, elasticities for both men and women are higher, consistent with higher overall monopsony power in these two occupations. Results are of similar significance to our baseline specification, and the implied gender wage gap of 4% is of similar magnitude as above. In the third panel, we include nurses and teachers, but allow them to contribute to the gender wage gap by weighting gender differences in the elasticity according to the shares (given in footnote) of the gender's employment within the occupation. In this specification, we see very little change in the male elasticity, but a drop in the female elasticity. This implies a 4.6% gender wage gap. In summary, accounting for the effects of nurses and teachers does lead to some increase in the implied monopsony based gender wage gap, though the results are still on the margins of statistical significance.

In the bottom three panels of the table, we explore whether measurement error may have impacted our estimation of gender effects. Recall that we specify the values of our female share variable in three ways: 1) by coding them as zero or one if an elasticity estimate was constructed solely on men or women (this result is also reported in Table 5), 2) by entering the gender mix, if reported, for an estimate conducted upon a sample of women and men, and 3) imputing a value of 0.5 for all other estimates. We refer to the first method as "non-imputed", and the second method as "partially-imputed". We run three more meta-regressions, restricting our sample first to "non-imputed" estimates, second to the "partially-imputed" estimates, and finally to a pooled sample of "non-imputed" and "partially-imputed" (i.e. omitting all estimates for which we imputed the female share as 0.5).

When focussing on the "non-imputed" estimates, we find lower elasticity estimates for both men and women, with a female effect that is statistically significant at 10% and 5% levels with and without bootstrapping, respectively. The markdowns of 38.5 for women and 32.1 for

[6] We thank the editor for this critique.

[7] Using the 2018 American Community Survey Ruggles et al. (2020), we find that 5.2% of employed women are nurses, compared to 0.7% of men. A similar gap exists for teachers, with 6.4% of women working in this occupation, as compared to 1.9% of men.

men imply that women are paid 61.5% of their worth to the firm, as compared to 67.9% for men. This implies a monopsony driven gender wage gap of around 9.4%, much larger than results reported above. The panel for "partially-imputed" results reports non-sensible negative estimates for both men and women, suggesting that this subsample of estimates is different enough from the overall sample, that our practice of applying sample mean values from the full sample here results in substantially issues when generating best practice estimates. This also implies that trying to infer anything about gender differences in elasticities, without possessing any information from estimates that were explicitly estimated differently by gender, is not appropriate. Finally, when pooling the "non-imputed" and "partially-imputed" samples, we find results that are more similar to those reported above with an implied gender wage gap of around 9.3%.

These results demonstrate two important things. First, when accounting for group differences in meta-regressions, it is important not to over-control for factors correlated across groups. This issue is of course similar to the familiar problem in the decompositions of wage gaps. Second, our treatment of the measurement error reveals that we are unable to say anything useful about gender differences in elasticities without information from estimates which explicitly estimated this difference. A sample share giving the percent female is not sufficient here, rather elasticity estimates which allowed for a fully interacted approach to estimating male and female elasticity were needed to say something meaningful. This should serve as a reminder to researchers that meta-regressions may not be able to answer questions not originally posed by researchers. It also highlights the value of future research in the monopsony literature on the topic of race. The authors of this article would have liked to speak to this issue here, but we found no published elasticity estimates that addressed racial differences in the elasticity of labor supply to the firm.

Table G2: Best Practice Gender Varients

Specification	Gender	Elasticity	Gender P	Gender (Wild)	95% Conf. Min	95% Conf. Max	95 % Conf. Min (Wild)	95 % Conf Max (Wild)	Markdown
Baseline	F	5.971	0.200	0.452	1.09	10.86	-0.90	13.13	14.3
	M	8.336	.	.	1.98	14.70	-1.19	17.52	10.7
Reduced Form	F	6.346	0.170	0.397	1.54	11.15	-0.45	13.52	13.6
	M	8.963	.	.	2.24	15.68	-1.26	18.87	10.0
Nurse & Teacher	F	6.074	0.145	.	1.14	11.01	-0.91	13.33	14.1
Effects	M	8.985	.	.	2.21	15.76	-1.27	18.99	10.0
Non-Imputed	F	1.599	0.059	0.027	-1.78	4.97	-2.90	12.32	38.5
	M	2.118	.	.	-1.25	5.48	-2.26	12.50	32.1
Partially-Imputed	F	-5.097	0.688	0.760	-13.36	3.16	-18.52	22.64	-24.4
	M	-3.028	.	.	-12.54	6.48	-69.45	57.59	-49.3
Non and Partially	F	2.615	0.142	0.262	-3.57	8.80	-6.95	13.30	27.7
Imputed	M	3.928	.	.	-3.06	10.91	-6.37	15.88	20.3

Appendix H: Further References

Belloni, A., D. Chen, V. Chernozhukov, and C. Hansen, "Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain," *Econometrica*, November 2012, *80* (6), 2369–2429.

Boal, William M and Michael R Ransom, "Monopsony in the labor market," *Journal of Economic Literature*, 1997, *35* (1), 86–112.

Cavlovic, Therese A., Kenneth H. Baker, Robert P. Berrens, and Kishore Gawande, "A Meta-Analysis Of Environmental Kuznets Curve Studies," *Agricultural and Resource Economics Review*, April 2000, *29* (1), 1–11.

Eicher, Theo S., Chris Papageorgiou, and Adrian E. Raftery, "Default Priors and Predictive Performance in Bayesian Model Averaging, with Application to Growth Determinants," *Journal of Applied Econometrics*, January/F 2011, *26* (1), 30–55.

Evers, Michiel, Ruud A. de Mooij, and Daniel J. van Vuuren, "What explains the Variation in Estimates of Labour Supply Elasticities?," Tinbergen Institute Discussion Papers 06-017/3, Tinbergen Institute February 2006.

Feldkircher, Martin, "Forecast Combination and Bayesian Model Averaging: A Prior Sensitivity Analysis," *Journal of Forecasting*, 07 2012, *31* (4), 361–376.

and Stefan Zeugner, "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*, 06 2012, *27* (4), 686–694.

Fernandez, Carmen, Eduardo Ley, and Mark F. J. Steel, "Benchmark Priors for Bayesian Model Averaging," *Journal of Econometrics*, February 2001, *100* (2), 381–427.

Fernández, Carmen, Eduardo Ley, and Mark F. J. Steel, "Model uncertainty in cross-country growth regressions," *Journal of Applied Econometrics*, 2001, *16* (5), 563–576.

Foged, Mette, Linea Hasager, and Vasil Yasenov, "The Role of Institutions in the Labor Market Impact of Immigration," *Working Paper*, 2019.

Hastie, Trevor, R. Tibshirani, and M. Wainwright, *Statistical learning with sparsity: The lasso and generalizations* 01 2015.

Koop, Gary, *Bayesian Econometrics*, John Wiley & Sons, 2003.

Ley, Eduardo and Mark F.J. Steel, "On the effect of prior assumptions in Bayesian model averaging with applications to growth regressions," *Journal of Applied Econometrics*, 2009, *24* (4), 651–674.

Madigan, D. and J. York, "Bayesian Graphical Models for Discrete Data," *International Statistical Review*, 1995, *63(2)*, 215–232.

Manning, Alan, "Imperfect Competition in the Labor Market," in "Handbook of Labor Economics, Volume 4B," North Holland, 2011, chapter 11, pp. 973–1035.

Moral-Benito, Enrique, "MODEL AVERAGING IN ECONOMICS: AN OVERVIEW," *Journal of Economic Surveys*, 2015, *29* (1), 46–75.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek, "The IPUMS collaboration: Integrating and disseminating the world's population microdata," *IPUMS USA*, 2020, *Version 10.0.*

Stanley, T.D. and Hristos Doucouliagos, "Neither Fixed nor Random: Weighted Least Squares Meta-Analysis," *Statistics in Medicine*, 2015, *34* (13), 2116–27.

Steel, Mark F. J., "Model Averaging and its Use in Economics," MPRA Paper 81568, University Library of Munich, Germany September 2017.

Tibshirani, Robert, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1996, *58* (1), 267–288.

Williamson, Samuel H, "What was the US GDP Then?," *Measuring Worth*, 2014.

Wolfson, Paul and Dale Belman, "15 Years of Research on US Employment and the Minimum Wage," *LABOUR*, 2019, *Forthcoming.*

Zeugner, Stefan and Martin Feldkircher, "Bayesian Model Averaging Employing Fixed and Flexible Priors – The BMS Package for R," *Journal of Statistical Software*, 2015, *68* (4), 1–37.