Econometrica Supplementary Material

SUPPLEMENT TO "DOES INDUSTRIAL COMPOSITION MATTER FOR WAGES? TEST OF SEARCH AND BARGAINING THEORY"

(*Econometrica*, Vol. 80, No. 3, May 2012, 1063–1104)

BY PAUL BEAUDRY, DAVID GREEN, AND BENJAMIN SAND[1]

In this appendix, we outline the details of our data construction (Appendix S.1) and the implementation of the selection correction procedure described in the main text (Appendix S.2). In addition, we examine the robustness of our results in a number of dimensions and provide some derivations of key equations in the main paper. In particular, we examine the implications of allowing bargaining strength, mobility, and job destruction parameters to vary by industry, and of allowing parameters to vary by education level and over time in Appendix S.3. In Appendix S.4, we provide a brief discussion and investigation of the implications of allowing on-the-job search. In Appendix S.5, we present a Monte Carlo exercise aimed at investigating whether our linearization could be leading to biased results. Appendix S.6 contains a detailed derivation of the main linear approximation in the paper, and Appendix S.7 contains an extensive presentation of the model when house prices are included. Appendix S.8 contains results from specifications in which we do not correct for self-selection of workers across cities, Appendix S.9 presents our first-stage regression results, and, finally, Appendix S.10 contains estimates of equation (10).

APPENDIX S.1: DATA CONSTRUCTION

THE CENSUS DATA were obtained with extractions done using the integrated IPUMS system (see Ruggles, Sobek, Alexander, Fitch, Goeken, Hall, King, and Ronnander (2004)). The files were the 1980 5% state (A Sample), 1990 state, 2000 5% Census PUMS, and the 2007 American Community Survey. For 1970, forms 1 and 2 were used for the metro sample. The initial extraction includes all individuals aged 20–65 not living in group quarters. All calculations are made using the sample weights provided. For the 1970 data, we adjust the weights for the fact that we combine two samples. We focus on the log of weekly wages, calculated by dividing wage and salary income by annual weeks worked. We impute incomes for top-coded values by multiplying the top code value in each year by 1.5. Since top codes vary by state in 1990 and 2000, we impose common top-code values of 140,000 in 1990 and 175,000 in 2000.

A consistent measure of education is not available for these Census years. We use indicators based on the IPUMS recoded variable EDUCREC that computes comparable categories from the 1980 Census data on years of school completed and later Census years that report categorical schooling only. To calculate potential experience (age minus years of education minus 6), we assign group mean years of education from Table 5 in Park (1994) to the categorical education values reported in the 1990 and 2000 Censuses.

[1]The authors thank A. Bowlus, M. Bombardini, D. Card, G. Dahl, M. Doms, J. Fernald, N. Fortin, G. Galipoli, J. Gelbach, R. Gordon, M. Greenstone, S. Kortum, I. King, T. Lemieux, K. Milligan, B. Meyer, F. Pelgrin, J. Pencavel, J.-M. Robin, and F. Wolak for helpful discussions.

© 2012 The Econometric Society DOI: 10.3982/ECTA8659

Census definitions of metropolitan areas are not comparable over time since, in general, the geographic areas covered by them increase over time and their definitions are updated to reflect this expansion. The definition of cities we use attempts to maximize geographic comparability over time and roughly correspond to 1990 definitions of Metropolitan Statistical Areas (MSAs) provided by the U.S. Office of Management and Budget.[2] To create geographically consistent MSAs, we follow a procedure based largely on Deaton and Lubotsky (2001) which uses the geographical equivalency files for each year to assign individuals to MSAs or Primary Metropolitan Statistical Areas (PMSAs) based on FIPs state and PUMA codes (in the case of 1990 and 2000) and county group codes (for 1970 and 1980). Each MSA label we use is essentially defined by the PUMAs it spans in 1990. Once we have this information, the equivalency files dictate what counties to include in each city for the other years. Since the 1970 county group definitions are much coarser than those in later years, the number of consistent cities we can create is dictated by the 1970 data. This process results in our having 152 MSAs that are consistent across all our sample years. Code for this exercise was generously provided by Ethan G. Lewis. Our definitions differ slightly from those in Deaton and Lubotsky (2001) so as to improve the 1970–1980–1990–2000 match.

We use an industry coding that is consistent across Censuses and is based on the IPUMS recoded variable `IND1950`, which recodes census industry codes to the 1950 definitions. This generates 144 consistent industries.[3] We have also replicated our results using data only for the period 1980–2000, where we can use 1980 industry definitions to generate a larger number of consistent industry categories.[4] We are also able to define more (231) consistent cities for that period.

Our measure of housing prices follows Moretti (2010). In particular, we use the IPUMS variable "gross monthly rent" called `RENTGRS`. This measure includes the contract rent plus utility costs, and IPUMS suggests that it is more comparable across individuals than "contract monthly rent." However, we find very similar results using either measure. As in Moretti (2010), we limit the sample to rental units with two or three bedrooms, and we correct for top coding by multiplying top-coded values by 1.3.

[2]See http://www.census.gov/population/estimates/metro-city/90mfips.txt for details.

[3]See http://usa.ipums.org/usa-action/variables/IND1950#description_tab for details.

[4] The program used to convert 1990 codes to 1980 comparable codes is available at http://www.unionstats.com/. That site is maintained by Barry Hirsch, Trinity University and David Macpherson, Florida State University. Code to convert 2000 industry codes into 1990 codes was provided by Chris Wheeler and can be found at http://research.stlouisfed.org/publications/review/past/2006. See also a complete table of 2000–1990 industry cross-walks at http://www.census.gov/hhes/www/ioindex/indcswk2k.pdf.

S.1.1. *Enclave Instrument*

The construction of the enclave instrument is similar to that of Doms and Lewis (2006) and uses their origin country groupings. The country of origin groups are (1) Mexico, (2) Central America, (3) South America, (4) Central Europe and Russia, (4) Caribbean, (5) China, (6) Southeast Asia, (7) India, (8) Canada, United Kingdom, and Australia, (10) Africa, (11) Korea and Japan, (12) Pacific Islands, (13) Israel and Northwest Europe, (14) Middle East, (15) Central Asia, (16) Cuba, and (17) Southern Europe, and can be identified from the IPUMS variable `bpl "Birthplace [general version]"`. To identify the inflow of immigrants, we use the IPUMS variable `yrimmig "Year of immigration"`. We predict the inflow of immigrants from sending country h to city c in year t by $\hat{H}_{ct} = \sum_h \lambda_{ch} \cdot H_{th}$, where λ_{ch} denotes the historical settlement of immigrants from h to c (we use the 1970 distribution of immigrants to estimate this), and H_{th} is the national inflow of immigrants from sending country h over the decade ending at t. We then form IV5 by

$$\text{IV5} = \frac{\hat{H}_{ct} - P_{ct-1}}{P_{ct-1}},$$

where P_{ct-1} denotes the population of city c at time $t-1$.

S.1.2. *Climate Instrument*

The city-level climate variables were extracted from “Sperling’s Best Places to Live” (http://www.bestplaces.net/docs/DataSource.aspx). Their data are compiled from the National Oceanic and Atmospheric Administration. The variables we use in this paper are the average daily high temperatures for July and January in degrees Fahrenheit, annual inches of rainfall, and the number of sunny days. We have also compiled climate data from an alternative source to use as a robustness check. These data come from “CityRating.com’s” historical weather data, and include variables on average annual temperature, number of extreme temperature days per year, humidity, and annual precipitation. Data from this source could only be collected for 106 cities, and, therefore, are not included in this analysis.

S.1.3. *Input–Output Linkages Data*

We use the input–output table “The Use of Commodities by Industries before Redefinition” for detailed industries in the 1997 benchmark year to create the distance measure d_{ij}. Creating this measure required several steps. First, we had to convert North American Industry Classification (NAIC) 1997 codes into standard industrial criterion (SIC) classification using the concordances provided by http://www.macalester.edu/research/economics/page/haveman/trade.

resources/tradeconcordances.html#FromNAICS. We then convert the SIC codes to Census 1980 industrial codes using concordances available from the same web page. The 1980 Census codes are then aggregated into our industrial classification described above. Once this is done, we sum the value of inputs used by industry i and create d_{ij} as the value of industry j's inputs used as a fraction of all input used by i.

S.1.4. *Net Export Data*

We obtain data on net exports from http://www.som.yale.edu/faculty/pks4/sub_international.htm and use data file *xm_sic*87_72_105_20100504.*dta* from that page. These data are described by Peter K. Schott in http://www.som.yale.edu/faculty/pks4/files/research/data/sic_naics_trade_20100504.pdf. We convert the industry codes in SIC format to Census 1980 format using the concordances described above. These are again aggregated into our industrial classification. We use the variables `cif` and `x` to create our net export variable.

APPENDIX S.2: IMPLEMENTING THE SELECTION ESTIMATOR

As described in the paper, our main approach to addressing the issue of selection on unobservables of workers across cities follows Dahl (2002). To understand the nature of Dahl's approach, consider a model in which each worker has a (latent) wage value that he would earn if he lived in each possible city and chooses to live in the city in which his wage net of moving costs is highest. If we explicitly introduce individual heterogeneity, this implies that we should write the regression corresponding to observed wages as

$$E(\log w_{kict}|d_{kct}=1) = \alpha_{0t} + \beta_{1t}x_{kct} + \alpha_1 \mathrm{ER}_{ct} + \alpha_2 R_{ct} + \nu_i + \nu_c + E(e_{kct}|d_{kct}=1), \tag{S1}$$

where k indexes individuals and d_{kct} is a dummy variable equal to 1 if worker k is observed in city c at time t. The last error mean term is nonzero if worker city selection is not independent of the unobserved component of wages. If one were to estimate equation (16) not taking into account this error mean term, then the estimated regression coefficients will suffer from well known consistency problems.

Dahl argued that the error mean term in equation (S1) for person j can be expressed as a function of the full set of probabilities that a person born in j's state of birth would choose to live in each possible city in the Census year. Further, he presented a sufficiency assumption under which the error mean term is a function only of the probability of the choice actually made by j. That sufficiency condition essentially says that two people with the same probability of choosing to live in a given city have the same error mean term in their regression: knowing the differences in their probabilities of choosing other options is not relevant for the size of the selection effect in the process that determines the wage where they actually live. Dahl, in fact, presented evidence that this assumption is overly restrictive and settled on a specification in which the error mean term is written as a function of the probability of making the migration choice actually observed and the probability that the person stayed in his birth state.

Implementing Dahl's selection correction approach requires two further decisions: how to estimate the relevant migration probabilities and what function of those probabilities to use as the error mean term. For the first, Dahl proposed a nonparametric estimator in which he divides individuals into cells defined by discrete categories for education, age, gender, race, and family status. He then used the proportion of people within the cell that is relevant for person j who actually made the move from j's birth state to his destination and the proportion who stayed in his birth state as the estimates of the two relevant probabilities. This is a flexible estimator which does not impose any assumptions about the distribution of the errors in the processes that determine the migration choice. For the second decision, Dahl used a series estimator to provide a nonparametric estimate of the error mean term as a function of these probabilities.

We essentially implement Dahl's approach in the same manner, apart from several small changes. First, we examine the set of people who live in cities in the various Census years, but we only know the state, not the city of birth. We form probabilities of choosing each city for people from each state of birth. People who live in a city in their state of birth are classified as "stayers" and those observed in a city not in their state of birth are classified as "movers."[5] We estimate the error mean term as a function of the probability that a person born in j's state of birth moved to j's city of residence and the probability that a person born in j's state of birth still resided in that same state. Stayers have an error mean term which is a function only of the probability that the person stayed in their state of birth (since the probability of their actual choice and the probability of staying are one and the same).

As in Dahl (2002), we estimate the relevant probabilities using the proportion of people within cells defined by observable characteristics who made the same move or who stayed in their birth state. Similar to Dahl (2002), we define the cells using four education categories, eight age categories, gender, and a black race dummy. For stayers, we also use extra dimensions based on family status.[6] This is possible because of the larger number of stayers than movers. The full interaction of these various characteristics defines 80 possible person types for movers and 240 for stayers. For movers in a particular city (i.e., for the set of people born outside the state in which that city is situated), the probabilities also differ based on where the person was born. Thus, identification of

[5]For cities that span more than one state, we call a person who is observed in a city that is at least partly in their birth state a stayer.

[6]Specifically, we use single, married without children, and married with at least one child under age 5.

the error mean term comes from the assumption that where a person was born does not affect the determination of their wage, apart from through the error mean term. Intuitively, a person born in Pennsylvania has a lower probability of being observed in Seattle than a person born in Oregon. If both are observed living in Seattle, then we assume that the person from Pennsylvania must have a larger Seattle-specific "ability" (a stronger earnings related reason for being there), and this is what is being captured by the sample correction. Identification in this approach is based on the exclusion of state of birth by current city of residence interactions from the wage regression. That is, we assume that being born in a state close to your city of residence (or, more generally, a state with a high associated probability of moving to that city) does not directly determine the wage a worker receives.[7] For stayers, we do not have this form of variation and, hence, identification arises from the restriction that family status affects the decision to stay in one's state of birth but not (directly) the wage.

Our main difference relative to Dahl (2002) is that he drops immigrants, while we keep them in our sample. We essentially treat them as if they are born in a different state from the city of residence except that we do not include a probability of their remaining in their place of birth. We divide the rest of the world into 11 regions (or "states" of birth). As with other movers, we divide them into cells based on the same education, age, gender, and race variables and assign them a probability of choosing their city of residence. Contrary to other movers, however, we do not assign them the probability that immigrants from their region of birth are observed in their own city in the current Census year. Instead, we assign them the probability that a person with their same education was observed in their city in the previous Census. This follows the type of ethnic enclave assumption used in several recent papers on immigration, that is, essentially using variation based on the observation that immigrants from a particular region tend to migrate to cities where there are already communities of people with their background.

Having obtained the estimated probabilities of following observed migration paths and of staying in state of birth, we need to introduce flexible functions of them into our regressions. We introduce these functions in our first estimation stage. The specific functions we use are quadratics in the estimated probabilities. For movers born in the United States, we introduce a quadratic in the probability of moving to the actual city from the state of birth and a quadratic in the probability of remaining in the state of birth. For stayers, we introduce a quadratic in the probability of remaining in the state in general. For immigrants, we introduce a quadratic in the probability that people from the same region and with the same education chose the observed city. This represents a restriction on Dahl (2002), who allowed for separate functions for

[7]Note that this is different from assuming that state of birth does not affect current wages, since even if we include a set of state of birth dummy variables in our first-stage estimation, our approach remains identified off interactions between city of employment and state of birth.

each destination state. We, instead, assume that parameters in the functions representing the error mean term are the same across all cities. Thus, we estimate individual-level regressions of log wages on the same complete set of education and experience variables, indicators for race, immigrant status, and gender, as well as a full set of city-by-industry dummies, but now we also add our proxies for the error mean term. We again retain the coefficients on the city–industry dummy variables and then proceed with the second-stage regressions as before. The coefficients on the error mean proxy variables are jointly highly significant in the first-stage regressions, implying that there are significant sample selection issues being addressed with this estimator.

Finally, to provide perspective on the effects of the selection correction in Table S.VIII, we replicate Table I from the paper but without the selection corrections. A comparison of the two tables makes it evident that the selection corrections change our estimates to only a minor degree. For example, in Table S.VIII the coefficient on ΔR_{ct} is 2.45 when estimated by OLS and is 2.85 when we use instrument set IV1–IV2–IV3, compared to 2.47 and 2.91, respectively, in Table I.

APPENDIX S.3: RELAXING HOMOGENEITY ASSUMPTIONS

The model presented in the paper assumes a large degree of between-industry homogeneity, which allows for a cleaner presentation. In this appendix, we examine robustness to relaxing some of these assumptions.

S.3.1. *Allowing μ or κ to Vary by Industry*

The model presented in the paper assumes that industry retention, parameterized by μ, and worker bargaining power, parameterized by κ, are the same across all industries. These are obviously simplifying assumptions that we wish to relax here. As can be verified, allowing either of these parameters to vary by industry causes the coefficients in estimating equation (25) to also vary by industry. Accordingly, Figure S.I displays the coefficient on R_{ct} estimated by running equation (25) separately by industry. This gives 143 estimates of α_{2i}, where the i subscript now denotes the fact that α_2 varies by industry.

As can be seen from the figure, the estimates of α_{2i} are concentrated around our estimate of α_2 reported in the paper: α_{2i} has a median of 2.50 and an interquartile range from 2.97 to 1.92. Letting η_i represent the industry share at the national level, from years 1970–2007, we calculate the average by $\sum_i \eta_i \cdot \alpha_{2i} = 2.48$. This is quite close to the estimate obtained in the paper with the homogeneity assumptions imposed. Finally, Figure S.II shows industry-level IV results using instrument set IV1–IV2–IV3 that have a mean of 2.97.

An alternative to allowing coefficients to vary by a three-digit industrial classification is to aggregate industries and allow the coefficients to vary only between the aggregated groups. This relaxes some of the homogeneity assumptions while still allowing for a relatively concise presentation of the results.

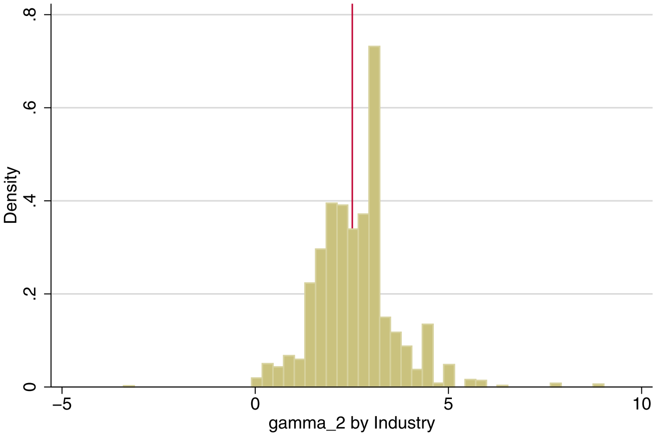

FIGURE S.I.—Estimates of α_{2i} from OLS.

Table S.I reports results from estimates of equation (25) by 16 industry aggregates. Column 1 reports OLS estimates and columns 2–4 report IV estimates using instrument sets IV1–IV3, IV2–IV3, and IV1–IV2–IV3, respectively. As can be seen from columns 1–4, estimates of α_{2i} are economically and statistically significant for nearly all industrial groups. In the last column of the table, we report average aggregated industry employment shares and we use these to calculate the average $\sum_{i=1}^{16} \eta_i \cdot \alpha_{2i}$ over the 16 aggregate industry groups. As can be seen from these rows, these averages are quite close to those reported in Table I of the paper with the homogeneity assumptions imposed.

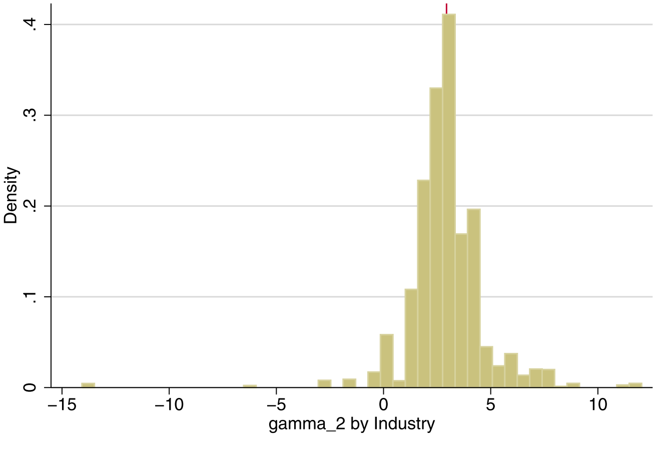

FIGURE S.II.—Estimates of α_{2i} from IV.

TABLE S.I

BASIC RESULTS BY INDUSTRY AGGREGATES[a]

	OLS	IV			η
	(1)	(2)	(3)	(4)	(5)
Agriculture	2.80*	3.13*	2.85*	2.97*	0.015*
	(0.60)	(1.15)	(1.09)	(0.82)	(0.0016)
Mining	2.10*	1.16	1.00	0.90	0.014*
	(0.53)	(3.51)	(2.15)	(1.75)	(0.0024)
Construction	2.97*	3.77*	2.82*	3.15*	0.059*
	(0.25)	(0.47)	(0.40)	(0.35)	(0.0014)
Durable man	2.68*	2.50*	3.07*	2.90*	0.13*
	(0.21)	(0.39)	(0.31)	(0.30)	(0.00050)
Nondurable	2.61*	2.38*	2.57*	2.48*	0.079*
	(0.26)	(0.54)	(0.52)	(0.50)	(0.00056)
Transport	2.23*	2.44*	2.51*	2.49*	0.042*
	(0.27)	(0.55)	(0.47)	(0.46)	(0.00078)
Communications	1.52*	1.67*	2.47*	2.23*	0.014*
	(0.38)	(0.61)	(0.49)	(0.45)	(0.0015)
Utilities	2.15*	2.30*	2.09*	2.16*	0.015*
	(0.25)	(0.40)	(0.39)	(0.36)	(0.0011)
Wholesale	2.74*	2.54*	2.68*	2.63*	0.043*
	(0.27)	(0.49)	(0.46)	(0.44)	(0.00070)
Retail	2.90*	3.47*	3.34*	3.38*	0.15*
	(0.26)	(0.47)	(0.44)	(0.41)	(0.00047)
F.I.R.E.[b]	2.10*	3.35*	3.58*	3.50*	0.071*
	(0.28)	(0.58)	(0.55)	(0.51)	(0.00081)
Business	3.36*	3.22*	4.09*	3.82*	0.064*
	(0.35)	(0.63)	(0.53)	(0.49)	(0.00082)
Personal	3.08*	3.81*	3.48*	3.62*	0.025*
	(0.38)	(0.64)	(0.65)	(0.56)	(0.00091)
Entertainment	2.15*	2.47*	3.12*	2.91*	0.018*
	(0.46)	(0.96)	(0.86)	(0.82)	(0.0012)
Professional	1.97*	2.57*	2.42*	2.47*	0.24*
	(0.19)	(0.34)	(0.30)	(0.27)	(0.00059)
Public admin.	1.44*	1.90*	1.96*	1.94*	0.074*
	(0.21)	(0.48)	(0.44)	(0.42)	(0.00088)
Observations	33,984	33,984	33,984	33,984	33,984
R^2	0.52				.
IV set		IV1–IV3	IV2–IV3	IV1–IV2–IV3	
Over-id. p-value		.	.	0.13	
1st col. ave.					2.56
2nd col. ave.					2.92
3rd col. ave.					2.99
4th col. ave.					2.97

[a]Standard errors, in parentheses, are clustered at the city–year level. The asterisk (*) denotes significance at the 5% level. All models are estimated on a sample of 152 U.S. cities using Census and ACS data for 1970–2007. The dependent variable is the decadal change in regression-adjusted city–industry wages.

[b]Finance, Insurance and Real Estate.

S.3.2. *Allowing δ to Vary by Industry*

The model presented in the paper assumes that the rate of job destruction, parameterized by δ, is the same across industries. When this assumption holds, the shares of vacant jobs across industries is proportional to the share of employment across industries. When this assumption does not hold, workers are more likely to meet industries with higher job destruction rates (i.e., greater turnover).

Denote δ_i as the job destruction rate for industry i. When the model is modified in this way, it can be shown that the relevant outside options for workers, captured by our rent variable, R_{ct}, have to be recalculated to account for the fact that workers now meet industries at different rates compared to the case where δ_i does not vary by industry. This new rent variable becomes $\tilde{R}_{ct} = \frac{\sum_i \eta_{ict} \cdot \delta_i \cdot \nu_{it}}{\sum_i \eta_{ict} \cdot \delta_i}$ and is derived using the same steps used to derive our estimating equation in the paper.

To explore the relevance to this extension, we require estimates of the job destruction rate by industry, δ_i. To obtain these, we use data from the Current Population Survey's February 1998 Job Tenure Supplement. We calculate $\delta_i = \frac{1}{T_i}$, where T_i is the average tenure in industry i.[8] Due to sample size considerations, we use 16 aggregated industry groups and calculate $\tilde{R}_{ct}$ as described above.

Since differences in δ_i across industries imply that the coefficients in (25) again depend on industry, Table S.II displays results from estimates of equation (25) using the new measure $\tilde{R}_{ct}$ for the 16 aggregated industry groups. Similar to tables reported in the paper, Table S.II column 1 shows OLS estimates, while columns 2–4 use the IV sets IV1–IV3, IV2–IV3, and IV1–IV2–IV3. Column 5 shows the estimated δ_i from the CPS data. Finally, the last column shows the share of employment in the aggregate industry group at the national level. As can be seen in columns 1–4 of this table, estimates of the effect of $\tilde{R}_{ct}$ on within-industry wage changes remain important and are similar to those reported in our basic specification and to those reported in Table S.I. In the bottom rows of the table, we report average estimates of α_{2i} using the shares reported in column 6 of the table, which are again similar in magnitude to α_2 reported in Table I of the main paper.

S.3.3. *Allowing the Returns to Education to Vary by Industry*

In our first-stage regression to purge industry–city wages of individual characteristics, such as education and potential experience, we do not allow the returns to these attributes to vary across industries. If the return to education varies across industries, it will bias our estimates of city–industry wages and this may ultimately bias our results.

[8]We use the BLS recode variable `prst1tn`, which identifies "tenure with current employer."

TABLE S.II

ALLOWING δ TO VARY BY INDUSTRY[a]

	OLS	IV			δ_i	η_i
	(1)	(2)	(3)	(4)	(5)	(6)
Agriculture	3.01*	2.71*	2.75*	2.73*	0.18*	0.015*
	(0.58)	(1.14)	(1.04)	(0.87)	(5.3e−19)	(0.0016)
Mining	1.87*	2.44	1.64	0.84	0.11*	0.014*
	(0.54)	(4.28)	(2.38)	(1.61)	(7.9e−19)	(0.0024)
Construction	3.02*	3.46*	2.91*	3.21*	0.17*	0.059*
	(0.24)	(0.39)	(0.48)	(0.36)	(4.9e−19)	(0.0014)
Durable man.	2.70*	2.17*	3.19*	2.71*	0.11*	0.13*
	(0.22)	(0.39)	(0.37)	(0.34)	(1.7e−19)	(0.00050)
Nondurable	2.63*	1.97*	2.33*	1.95*	0.12*	0.079*
	(0.26)	(0.60)	(0.68)	(0.67)	(1.9e−19)	(0.00056)
Transport	2.27*	2.28*	2.58*	2.43*	0.14*	0.042*
	(0.28)	(0.53)	(0.57)	(0.51)	(2.6e−19)	(0.00078)
Communications	1.53*	1.55*	2.69*	2.16*	0.089*	0.014*
	(0.36)	(0.56)	(0.52)	(0.46)	(5.2e−19)	(0.0015)
Utilities	2.18*	2.20*	2.12*	2.16*	0.080*	0.015*
	(0.27)	(0.41)	(0.44)	(0.40)	(3.7e−19)	(0.0011)
Wholesale	2.72*	2.33*	2.74*	2.50*	0.15*	0.043*
	(0.28)	(0.48)	(0.54)	(0.47)	(2.3e−19)	(0.00070)
Retail	3.01*	3.30*	3.55*	3.42*	0.22*	0.15*
	(0.25)	(0.42)	(0.51)	(0.43)	(1.6e−19)	(0.00047)
F.I.R.E.	2.06*	3.15*	4.11*	3.63*	0.16*	0.071*
	(0.30)	(0.60)	(0.63)	(0.56)	(2.7e−19)	(0.00081)
Business	3.51*	3.15*	4.10*	3.69*	0.24*	0.064*
	(0.34)	(0.58)	(0.58)	(0.54)	(2.8e−19)	(0.00082)
Personal	3.13*	3.52*	3.64*	3.56*	0.23*	0.025*
	(0.37)	(0.60)	(0.78)	(0.59)	(3.1e−19)	(0.00091)
Entertainment	2.16*	2.92*	3.00*	2.96*	0.22*	0.018*
	(0.48)	(0.90)	(1.01)	(0.91)	(4.0e−19)	(0.0012)
Professional	2.02*	2.52*	2.55*	2.53*	0.14*	0.24*
	(0.18)	(0.29)	(0.32)	(0.27)	(2.0e−19)	(0.00059)
Public admin.	1.41*	1.87*	2.30*	2.06*	0.097*	0.074*
	(0.21)	(0.47)	(0.50)	(0.44)	(3.0e−19)	(0.00088)
Observations	33,984	33,984	33,984	33,984	33,984	33,984
R^2	0.52				.	.
IV set		IV1–IV3	IV2–IV3	IV1–IV2–IV3		
Over-id. p-value		.	.	0.017		
1st col. ave.						2.60
2nd col. ave.						2.77
3rd col. ave.						3.14
4th col. ave.						2.93

[a]Standard errors, in parentheses, are clustered at the city–year level. The asterisk (*) denotes significance at the 5% level. All models are estimated on a sample of 152 U.S. cities using Census and ACS data for 1970–2007. The dependent variable is the decadal change in regression-adjusted city–industry wages.

We address this issue in two ways. First, toward the end of the paper in Table VI, we report our estimate of our main equation separately by experience/education groups. In doing so, we estimate our first-stage equation separately by each group and, accordingly, the returns to various observable characteristics are allowed to vary by experience/education groups. These first-stage regressions include a full set of city–industry dummies, effectively allowing the return to education to vary across industry.

Second, as an additional robustness check, we reestimate our baseline empirical model on the pooled data, but allow the return to education to vary across 16 aggregated industry groups. The results from this exercise are reported in Table S.III, which is akin to Table I. As can be seen, the results are very similar whether or not this restriction is imposed.

TABLE S.III
ROBUSTNESS: ALLOWING THE RETURNS TO EDUCATION TO VARY BY INDUSTRY[a]

	OLS		IV			
	(1)	(2)	(3)	(4)	(5)	(6)
ΔR_{ct}	2.25*			2.75*	2.65*	2.68*
	(0.17)			(0.35)	(0.29)	(0.28)
$\sum_i \nu_{it-1}(\eta_{ict} - \eta_{ict-1})$		1.94*	2.79*			
		(0.19)	(0.41)			
$\sum_i \eta_{ict}(\nu_{it} - \nu_{it-1})$		2.70*	2.57*			
		(0.37)	(0.37)			
$\Delta\,\mathrm{ER}_{ct}$	0.38*	0.43*	0.62	0.64	0.74	0.70
	(0.079)	(0.078)	(0.47)	(0.46)	(0.48)	(0.45)
Year × ind.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33,984	33,984	33,984	33,984	33,984	33,984
R^2	0.52	0.52				
Instrument set			IV1–IV2–IV3	IV1–IV3	IV2–IV3	IV1–IV2–IV3
F-statistics						
ΔR^W_{ct}			73.94			
ΔR^B_{ct}			638.04			
ΔR_{ct}				65.74	170.47	231.55
$\Delta\,\mathrm{ER}_{ct}$			9.73	9.92	14.06	9.73
AP p-values						
ΔR^W_{ct}			0.00			
ΔR^B_{ct}			0.00			
ΔR_{ct}				0.00	0.00	0.00
$\Delta\,\mathrm{ER}_{ct}$			0.00	0.00	0.00	0.00
Over-id. p-value			.	.	.	0.70

[a]Standard errors, in parentheses, are clustered at the city–year level. The asterisk (*) denotes significance at the 5% level. All models are estimated on a sample of 152 U.S. cities using Census and ACS data for 1970–2007. The dependent variable is the decadal change in regression-adjusted city–industry wages.

TABLE S.IV
BASIC RESULTS BY TIME PERIOD[a]

	1970–1990		1980–2000		1990–2007	
	(1)	(2)	(3)	(4)	(5)	(6)
ΔR_{ct}	2.67*	3.26*	2.63*	2.94*	2.73*	2.98*
	(0.23)	(0.39)	(0.20)	(0.41)	(0.24)	(0.43)
ΔER_{ct}	0.46*	1.08	0.41*	0.73	0.34*	−2.19
	(0.15)	(0.58)	(0.096)	(0.55)	(0.088)	(1.87)
Observations	17,110	17,110	28,109	28,109	27,763	27,763
R^2	0.35		0.57		0.46	
Instrument set		IV1–IV2–IV3		IV1–IV2–IV3		IV1–IV2–IV3
F-statistics						
ΔR_{ct}		192.3		213.0		235.8
ΔER_{ct}		10.7		14.2		1.19
AP p-value						
ΔR_{ct}		0		0		0
ΔER_{ct}		0		0		0.17
Over-id. p-value		0.33		0.74		0.071

[a]Standard errors, in parentheses, are clustered at the city–year level. The asterisk (*) denotes significance at the 5% level. All models are estimated on a sample of 152 U.S. cities using Census and ACS data by indicated time period. The dependent variable is the decadal change in regression-adjusted city–industry wages.

S.3.4. *Allowing Parameters to Vary Over Time*

In the paper, we make the simplifying assumption that the parameters do not vary over time. Allowing, for example, workers' bargaining power given by κ to vary by year implies that the estimated coefficients α_2 and α_3 in estimating equation (25) also varies by year. To evaluate the importance of this restriction, we reestimate our basic equation (25) for different time periods. The results are contained in Table S.IV. We estimate the basic specification for the years 1970–1990 (columns 1 and 2), 1980–2000 (columns 3 and 4), and 1990–2007 (columns 5 and 6). For each set of years, the first column corresponds to the OLS estimates and the second to the IV results, where we use the instrument set IV1–IV2–IV3. For both the OLS and IV results, the estimated coefficient α_{2t} on ΔR_{ct} is reasonably stable. We also estimate equation (25) separately by decade in Table S.V. When we do this, the results are again quite stable with the exception of the IV estimate for the 1990s, which is very imprecise and of the wrong sign. This imprecision comes from the fact that the first-stage regression for the 1990s is very poor.

APPENDIX S.4: ON-THE-JOB SEARCH

Extending the model presented in the paper to include on-the-job search is not straight forward. One difficulty is that there are many models of on-the-job search with different implications for how outside options map into the

TABLE S.V

BASIC RESULTS BY DECADE: 1970–2007[a]

	1970–1980		1980–1990		1990–2000		2000–2007	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔR_{ct}	2.03*	2.81*	3.26*	3.17*	2.33*	-1.37	1.38*	2.26*
	(0.25)	(0.39)	(0.37)	(0.71)	(0.40)	(2.31)	(0.41)	(0.92)
ΔER_{ct}	0.63*	1.55*	0.18	−4.31	0.35*	−0.058	0.53*	0.52
	(0.15)	(0.34)	(0.25)	(6.61)	(0.11)	(0.82)	(0.15)	(0.61)
Observations	6221	6221	10,889	10,889	10,999	10,999	5875	5875
R^2	0.34		0.29		0.16		0.21	
Instrument set		IV1–IV2–IV3		IV1–IV2–IV3		IV1–IV2–IV3		IV1–IV2–IV3
F-statistics								
ΔR_{ct}		87.9		196.5		10.3		64.8
ΔER_{ct}		15.8		0.40		9.31		9.35
AP p-value								
ΔR_{ct}		0		0		0.012		0
ΔER_{ct}		0		0.69		0.044		0.0079
Over-id. p-value		0.80		0.072		0.079		0.014

[a]Standard errors are given in parentheses. The asterisk (*) denotes significance at the 5% level. All models are estimated on a sample of 152 U.S. cities using Census and ACS data by decade. The dependent variable is the decadal change in regression-adjusted city–industry wages.

wage determination process. With on-the-job search, the bargaining process is extremely important. While getting into these details is outside the scope of the current paper, we wish to examine an implication of on-the-job search that is generally found in most setups; that is, most on-the-job search models imply that the expected wage for workers does not simply depend on the first moment of the distribution of outside offers—as is the case with our simple model with no on-the-job search. Instead, the expected wage in an industry depends on several higher moments of the outside options. Given that this implication is common among several on-the-job search frameworks, as a first pass we examine its relevance here by including higher order moments of the distribution of rents. In particular, instead of including only the first moment of the distribution, which is given by $\sum_i \eta_{ict} \cdot \nu_{it}$, we also include the terms $\sum_i \eta_{ict} \cdot (\nu_{it})^j$, where j is the order of the moment. Table S.VI reports results from this exercise for different moments from 2 to 4. The instruments we use to mitigate the endogeneity of these moments parallel our building of IV1. We examined results for centered and uncentered moments. The results in the table are for centered moments; uncentered moments provide a similar picture. As can be seen in the table, there is not much evidence in our data that moments other than the first matter significantly for wage determination at the city level. This does not imply that on-the-job search is not present or important, but it does imply that it effects are likely much more subtle than the effects emphasized in this paper. While this simple approach represents

TABLE S.VI

EXAMINING THE RELEVANCE OF HIGHER ORDER TERMS[a]

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta \sum_i \eta_{ict} \cdot \nu_{it}$	2.47*	2.82*	2.76*	2.96*	2.91*	2.98*	2.87*	2.79*
	(0.18)	(0.35)	(0.49)	(0.74)	(0.31)	(0.43)	(0.53)	(0.80)
$\Delta \sum_i \eta_{ict} \cdot (\nu_{it})^2$		0.94				0.27		
		(0.67)				(1.40)		
$\Delta \sum_i \eta_{ict} \cdot (\nu_{it})^3$			0.50				−0.086	
			(0.71)				(0.99)	
$\Delta \sum_i \eta_{ict} \cdot (\nu_{it})^4$				0.65				−0.17
				(0.90)				(1.16)
ΔER_{ct}	0.42*	0.43*	0.43*	0.43*	0.63	0.62	0.63	0.62
	(0.078)	(0.078)	(0.078)	(0.078)	(0.44)	(0.45)	(0.43)	(0.43)
Year × ind.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33,984	33,984	33,984	33,984	33,984	33,984	33,984	33,984
R^2	0.51	0.51	0.51	0.51				
F-statistics								
$\Delta \sum_i \eta_{ict} \cdot \nu_{it}$					222.8	171.0	167.2	167.5
$\Delta \sum_i \eta_{ict} \cdot (\nu_{it})^2$						51.4		
$\Delta \sum_i \eta_{ict} \cdot (\nu_{it})^3$							88.6	
$\Delta \sum_i \eta_{ict} \cdot (\nu_{it})^4$								113.4
ΔER_{ct}					11.0	9.57	8.31	8.26
Over-id. p-value					0.65	0.66	0.65	0.65

[a]Standard errors, in parentheses, are clustered at the city–year level. The asterisk (*) denotes significance at the 5% level. All models are estimated on a sample of 152 U.S. cities using Census and ACS data for 1970–2007. The dependent variable is the decadal change in regression-adjusted city–industry wages.

only a first pass at the issue, we leave more detailed explorations of particular on-the-job search models for future work.

APPENDIX S.5: EVALUATING LINEAR APPROXIMATION BY SIMULATION

As discussed in the paper, one potential concern with our approach is that higher order terms left in the error term after our linearizations are correlated with our regressors and/or our instruments. To investigate this concern, we implemented a Monte Carlo exercise in which we constructed data on wages using our nonlinear model and then estimated our main regressions using that data. We describe this exercise in this appendix.

It is possible to generate data for an entire Monte Carlo economy using our model and initial values for a, ε, and Ω. However, we are less interested in this than in generating data that both reflect the nonlinearities in the model and have levels and variation that match with the actual data we use in our estimation. For that reason, we use our actual data on city-level employment rates (ER_c), city-by-industry-level employment shares (η_{ic}), and national-level wage premia (ν_i) to generate wage data. More specifically, we generate a wage observation for each industry–city cell in a given Census year using equation (15). Since we are using actual values for ν_i and η_{ic}, we can use our actual measure for R_c,[9] but everything else on the right-hand side of (15) needs to be generated, which we do in the following steps.

Step 1. Assume values for the job destruction rate, δ, and the elasticity parameter in the matching function, σ, and use these together with data on ER_c in equations (9) to generate city-specific values for the probability an unemployed worker meets a job (ψ_c) and the probability a vacancy meets a worker (ϕ_c). In our actual implementation, we follow Cahuc and Zylberberg (2004) and introduce a parameter, θ, which multiplies that matching function and corresponds to the efficiency of matching (regardless of the level of tightness of the market).

Step 2. Assume values for the discount rate, ρ, and the bargaining parameter, κ, and use them together with the values for ψ_c and ϕ_c generated in Step 1 to generate values for γ_{c0}, γ_{c1}, and γ_{c2} (see the formulas below equation (12) in the paper).

Step 3. The industry-specific intercepts in equation (15) depend on the national industry-level prices. To get values for these that are consistent with the rest of our data, we average equation (12) across cities for a given industry and then rearrange to obtain p_i. More specifically, we make use of the γ_{c0}, γ_{c1}, and γ_{c2} values generated in Step 2, national-level average industry wages for the year, and average city wages. The εs average to zero across cities within an industry and so can be ignored in this step. Throughout this exercise, we drop nine industries where a majority of industry–city cells are smaller than 20 observations. We also drop other cells where the number of observations is smaller than 20 observations in either year for a decade (e.g., for either 1980 or 1990 when we simulate data for that decade). We are left with 10,915 usable city–industry cells. We normalize the industry prices (in thousands of dollars) so that the price of good 1 is 100.

Step 4. Finally, we obtain values for the cost shocks, ε_{ic}, as independent draws from a standard normal which we then adjust in two ways. First, we adjust them to average to zero across cities within an industry. Second, we multiply them by a standard deviation parameter chosen so that the final generated wages are close to actual wages in terms of their means and standard deviations.

Given all of these generated and actual values, we can generate values for the city–industry cell average wages for each Census year. These wages reflect the nonlinearities, especially with respect to ER_c, that are inherent in the model. Because the εs are independent draws, there is no reason for our identification conditions to be violated, and both OLS and IV estimates of our main regressions should provide consistent estimates.

[9]The wages are actually created in levels and so we use a version of R_c using level differences in wages by industry rather than proportion differences at this stage.

Having generated the industry–city cell mean wages, we use them as the dependent variable in an OLS estimation of the linearized regression equation from the paper. Because we used actual values for ER_c, η_{ic}, and ν_c in our construction, the values for ER_c and R_c are the same in our constructed world as they are in the actual data and so we use the actual values for these variables as our regressors. We are interested in whether the coefficient on R_c in our regression is close to the correct value as implied by the model and whether its accuracy varies with the model parameter values.

To understand the standard of comparison for the estimated R_c coefficient, recall that we perform our linearization around a point such that the employment rate is the same across all cities. The parameter of interest is then equal to $\gamma_2/(1-\gamma_2)$, where γ_2 is constructed using the common employment rate. To reflect this, for each city we obtain a value for γ_2 as described in Step 2 in our simulation exercise and calculate $\gamma_2/(1-\gamma_2)$. We then obtain the average value of $\gamma_2/(1-\gamma_2)$ across cities and use it as our target. This value will vary with the key parameters in the model.

In Table S.VII, we present the proportionate difference between the estimated regression coefficient and the target value for $\gamma_2/(1-\gamma_2)$ for each of a set of different values for the model parameters.[10] We start by using the values for key parameters recommended in Cahuc and Zylberberg (2004): $\delta = 0.15$, $\sigma = 0.5$, and $\rho = 0.05$. We also set $\kappa = 1$ to represent equal bargaining power

TABLE S.VII

PROPORTIONATE DIFFERENCES BETWEEN ESTIMATED AND MEAN RENT EFFECTS FOR VARIOUS PARAMETER VALUES

δ	σ	ρ	κ	θ	Difference
0.15	0.50	0.05	1.0	20	0.0028
0.05	0.50	0.05	1.0	20	0.32
0.25	0.50	0.05	1.0	20	0.020
0.15	0.25	0.05	1.0	20	0.30
0.15	0.75	0.05	1.0	20	−1.15
0.15	0.50	0.03	1.0	20	0.017
0.15	0.50	0.07	1.0	20	0.0039
0.15	0.50	0.05	0.5	20	−0.019
0.15	0.50	0.05	1.5	20	0.0090
0.15	0.50	0.05	1.0	1	−0.93
0.15	0.50	0.05	1.0	10	−0.033
0.15	0.50	0.05	1.0	40	0.020

[10]The values of the estimated coefficient vary to a small extent between samples, likely because of the way the generated εs appear in interacted terms with the key coefficients in the wage equation. Because of this, we actually run our wage regression 50 times for each set of parameter values. We then use the average of the estimated coefficients from these 50 replications to construct the proportionate difference reported here.

between workers and firms. Finally, we choose the value for θ—the matching efficiency parameter—to provide a close fit between our estimated coefficient and the target parameter value. We do this because we do not have a clear way to map from our data to a value for θ. The fact that we get close agreement between our coefficient and the target value is, therefore, potentially not surprising in this case. It does, however, indicate that at reasonable parameter values, the wage generation function implied in the model is not so nonlinear that we cannot get agreement between these values when we estimate using a linearized version of the model.

In the ensuing rows, we vary each of the key parameter values in turn. In almost all cases, we continue to obtain close agreement between the coefficient and target parameter values. There are, however, a few notable exceptions. When σ, the elasticity coefficient in the matching function, takes either quite high or low values, we tend to get disagreement between the values. Altering this parameter value seems to us to be likely to affect the amount of nonlinearity in the wage function and so this seem reasonable. However, in the range for this parameter deemed reasonable by Cahuc and Zylberberg (2004) in their discussion of earlier studies, the linear approximation seems good. A very low value for the job destruction rate also implies a 30% difference between the estimated and target values, but quite high values do not imply large differences. Finally, if the efficiency parameter takes values near 1, we observe substantial differences, but for values above about 10 (and including quite high values), the differences are small. As mentioned earlier, we do not have a means at present to determine what is a reasonable value of θ (we believe it would be related to the average amount of time a vacancy goes unfilled), so we do not know which specific values to trust. However, we find it encouraging that we see small differences between the estimated and target values over a very large range of values for θ. Overall, we conclude that the wage generation process in this model is not so inherently nonlinear that our linearization approach causes problems given what we see as reasonable ranges for parameter values.

APPENDIX S.6: DERIVATION OF MAIN LINEAR APPROXIMATION

In this appendix, we present the derivation of our main linear approximation for completeness. Recall wage equation (24) from the text:

$$w_{ic} = D_{ic} + \Gamma_{c2}\frac{\gamma_{c1}}{\gamma_1}R_c + \gamma_{c1}\varepsilon_{ic} + \gamma_{c1}\Gamma_{c2}\sum_j \eta_{jc}\varepsilon_{jc},$$

where $\Gamma_{c2} = (\frac{\gamma_{c2}}{1-\gamma_{c2}})$.

Since the coefficients in the wage equation are nonlinear functions of the employment rates, ER_c, we take a linear approximation. Let the vector $\mathbf{e} = [p_1, p_i, R_c, \mathrm{ER}_c, \varepsilon_{ic}, \varepsilon_{jc}]$ denote the variables that affect wages in this equation and with respect to which we take the approximation. Writing out the wage equation to make this relationship explicit yields

$$w_{ic}(\mathbf{e}) = D_{ic}(\mathbf{e}) + \Gamma_{c2}(\mathbf{e})\frac{\gamma_{c1}(\mathbf{e})}{\gamma_1}R_c + \gamma_{c1}(\mathbf{e})\varepsilon_{ic} + \gamma_{c1}(\mathbf{e})\Gamma_{c2}(\mathbf{e})\sum_j \eta_{jc}(\mathbf{e})\varepsilon_{jc}.$$

We expand around a point where the employment rate does not vary across cities, which will occur if cities have a common industrial structure. This occurs at the point $\mathbf{e}_0 = (p_1, p_1, 0, \mathrm{ER}, 0, \mathbf{0})$. This approximation is

$$w_{ic}(\mathbf{e}) \approx w_{ic}(\mathbf{e}_0) + \nabla w_{ic}(\mathbf{e}_0) \cdot (\mathbf{e} - \mathbf{e}_0). \tag{S2}$$

S.6.1. $w_{ic}(\mathbf{e}_0)$

Dealing first with the first term on the right-hand side of S2 gives:

$$\begin{aligned} w_{ic}(\mathbf{e}_0) &= D_{ic}(\mathbf{e}_0) + \Gamma_{c2}(\mathbf{e}_0)\frac{\gamma_{c1}(\mathbf{e}_0)}{\gamma_1}\cdot 0 \\ &\quad + \gamma_{c1}(\mathbf{e}_0)\cdot 0 + \gamma_{c1}(\mathbf{e}_0)\Gamma_{c2}(\mathbf{e}_0)\sum_j \eta_{jc}(\mathbf{e}_0)\cdot 0 \\ &= D_{ic}(\mathbf{e}_0) \\ &= \gamma_0\cdot(1+\Gamma_2) + \gamma_1\Gamma_2 p_1 + \gamma_1 p_1 \\ &= \beta_0, \end{aligned} \tag{S3}$$

where β_0 is a constant that does not vary by city or industry.

S.6.2. $\nabla w_{ic}(\mathbf{e}_0)\cdot(\mathbf{e} - \mathbf{e}_0)$

Now dealing with the second term on the right-hand side of equation (S2) gives:

$$\begin{aligned} &\nabla w_{ic}(\mathbf{e}_0)\cdot(\mathbf{e}-\mathbf{e}_0) \\ &= \left.\begin{pmatrix} \gamma_{c1}\Gamma_{c2} \\ \gamma_{c1} \\ \Gamma_{c2} \\ \dfrac{\partial D_{ic}}{\partial \mathrm{ER}_c} + \dfrac{\partial \gamma_{c1}/\gamma_1 \Gamma_{c2} R_c}{\partial \mathrm{ER}_c} + \dfrac{\partial \gamma_{c1}}{\partial \mathrm{ER}_c}\cdot\varepsilon_{ic} + \sum_i \dfrac{\partial(\gamma_{c1}\Gamma_{c2}\eta_{ic})}{\partial \mathrm{ER}_c}\cdot\varepsilon_{ic} \\ \gamma_{c1} + \gamma_{c1}\Gamma_{c2}\eta_{ic} \\ \gamma_{c1}\Gamma_{c2}\eta_{jc} \end{pmatrix}'\right|_{(\mathbf{e}_0)} \\ &\quad \cdot(\mathbf{e}-\mathbf{e}_0) \end{aligned} \tag{S4}$$

$$= \begin{pmatrix} \gamma_1 \Gamma_2 \\ \gamma_1 \\ \Gamma_2 \\ \gamma_3 \\ \gamma_1 + \gamma_1 \Gamma_2 \frac{1}{I} \\ \gamma_1 \Gamma_2 \frac{1}{I} \end{pmatrix}' \cdot \begin{pmatrix} p_1 - p_1 \\ p_i - p_1 \\ R_c - 0 \\ \mathrm{ER}_c - \mathrm{ER} \\ \varepsilon_{ic} - 0 \\ \boldsymbol{\varepsilon}_{\mathbf{jc}} - \mathbf{0} \end{pmatrix}$$

$$= \gamma_1 \cdot (p_i - p_1) + \Gamma_2 \cdot R_c + \gamma_3 (\mathrm{ER}_c - \mathrm{ER}) + \gamma_1 \cdot \varepsilon_{ic} + \gamma_1 \Gamma_2 \sum_j \frac{1}{I} \varepsilon_{jc}.$$

S.6.3. *Arriving at Equation* (16)

Adding (S3) and (S4), and differencing gives

$$\Delta w_{ic} = \Delta d_i + \Gamma_2 \cdot \Delta R_c + \gamma_3 \Delta\, \mathrm{ER}_c + \gamma_1 \cdot \Delta\varepsilon_{ic} + \gamma_1 \Gamma_2 \sum_j \frac{1}{I} \Delta\varepsilon_{jc}, \tag{S5}$$

where $\Delta d_i = \gamma_1 \Gamma_2 \Delta p_1 + \gamma_1 \Delta p_i$.

S.6.4. *Equation for* η_{ic}

Using equation (6), the expression for η_{ic} is obtained:

$$\eta_{ic} = \left((\Upsilon_i + \Omega_{ic}) \left[p_i - d_{ic} + (1 - \gamma_{c1}) \varepsilon_{ic} + \left(\frac{\gamma_{c2}}{1 - \gamma_{c2}} \right) \gamma_{c1} \left(\sum_j \eta_{jc} (p_j - p_1) - \sum_j \eta_{jc} \varepsilon_{jc} \right) \right] \right)^{1/q} \Bigg/ \sum_i \left((\Upsilon_i + \Omega_{ic}) \left[p_i - d_{ic} + (1 - \gamma_{c1}) \varepsilon_{ic} + \left(\frac{\gamma_{c2}}{1 - \gamma_{c2}} \right) \gamma_{c1} \left(\sum_j \eta_{jc} (p_j - p_1) - \sum_j \eta_{jc} \varepsilon_{jc} \right) \right] \right)^{1/q}. \tag{S6}$$

APPENDIX S.7: DERIVING THE HOUSING PRICE SPECIFICATION

This appendix outlines the derivation of equation (29) in the main text, and proves the claim that the direct and indirect effects of industrial composition changes are not separately identified under the assumption of perfect mobility.

To begin, we add to Worker's Bellman equations the option to move and the cost of housing (as in the paper):

$$\rho U_c^u = b + \tau_c - s \cdot p_c^h + \psi_c \cdot \left(\sum_i \eta_{ic} \cdot U_{ic}^e - U_c^u \right) + \mu_1 \cdot (U_{\max}^u - U_c^u) \tag{S7}$$

and

$$\rho U_{ic}^e = w_{ic} + \tau_c - s \cdot p_c^h - \delta \cdot (U_c^u - U_{ic}^e). \tag{S8}$$

Taking the difference between (S8) and (S7) yields

$$(\rho + \delta) \cdot (U_{ic}^e - U_c^u) = w_{ic} - b - \psi_c \cdot \left(\sum_i \eta_{ic} U_{ic}^e - U_c^u \right) - \mu_1 \cdot (U_{\max}^u - U_c^u). \tag{S9}$$

Next, we solve for $(\sum_i \eta_{ic} U_{ic}^e - U_c^u)$ and substitute back into (S9) to get

$$(\rho + \delta) \cdot (U_{ic}^e - U_c^u) = w_{ic} - b(1 - \Pi_1) - \mu_1 (1 - \Pi_1) \cdot (U_{\max}^u - U_c^u) - \Pi_1 \left(\sum_i \eta_{ic} w_{ic} \right), \tag{S10}$$

where $\Pi_1 = \frac{\psi_c}{\rho + \delta + \psi_c}$. Using $(\sum_i \eta_{ic} U_{ic}^e - U_c^u)$, we can solve for U_c^u in equation (S7):

$$(\rho + \Pi_2) \cdot U_c^u = b(1 - \Pi_1) + \tau_c - s \cdot p_c^h + \Pi_1 \cdot \left(\sum_i \eta_{ic} w_{ic} \right) + \Pi_2 \cdot U_{\max}^u, \tag{S11}$$

where $\Pi_2 = \mu_1 (1 - \Pi_1)$. Substituting this into (S10) yields

$$(\rho + \delta) \cdot (U_{ic}^e - U_c^u) = w_{ic} - b(1 - \Pi_1) \cdot \Pi_3 - \Pi_2 \cdot \Pi_3 \cdot U_{\max}^u - \Pi_1 \cdot \Pi_3 \left(\sum_i \eta_{ic} w_{ic} \right) + \frac{\Pi_2}{(\rho + \Pi_2)} (\tau_c - s \cdot p_c^h), \tag{S12}$$

where $\Pi_3 = \frac{\rho}{\rho + \Pi_2}$. Note that if $\mu_1 = 0$ (the case with no mobility), then this reduces to $\Pi_3 = 1$ and $\Pi_2 = 0$, and

$$(\rho + \delta) \cdot (U_{ic}^e - U_c^u) = w_{ic} - b(1 - \Pi_1) - \Pi_1 \left(\sum_i \eta_{ic} w_{ic} \right),$$

which is what we get in equation (10) if $\mu = 0$.

S.7.1. *Wage Equation*

From the Nash condition,

$$\frac{\kappa}{\rho+\delta}\Bigg[w_{ic} - b(1-\Pi_1)\cdot\Pi_3 - \Pi_2\cdot\Pi_3\cdot U^u_{\max} - \Pi_1\cdot\Pi_3\Bigg(\sum_i \eta_{ic}w_{ic}\Bigg) + \frac{\Pi_2}{(\rho+\Pi_2)}(\tau_c - s\cdot p^h_c)\Bigg] = \frac{p_i - w_{ic} + \varepsilon_{ic}}{\rho+\delta+\phi_c}.$$

Rearranging gives

$$w_{ic} = \gamma_{c0} + \gamma_{c1}\cdot p_i + \gamma_{c2}\sum_i \eta_{ic}w_{ic} + \gamma_{c3}\cdot sp^h_c - \gamma_{c3}\cdot\tau_c + \gamma_{c1}\varepsilon_{ic}, \tag{S13}$$

where

$$\gamma_{c0} = \left[\frac{\kappa(\rho+\delta+\phi_c)}{\kappa(\rho+\delta+\phi_c)+(\rho+\delta)}\right]\cdot[(1-\Pi_1)\Pi_3\cdot b + \Pi_2\cdot\Pi_3U^u_{\max}],$$

$$\gamma_{c1} = \left[\frac{\rho+\delta}{\kappa(\rho+\delta+\phi_c)+(\rho+\delta)}\right],$$

$$\gamma_{c2} = \left[\frac{\kappa(\rho+\delta+\phi_c)}{\kappa(\rho+\delta+\phi_c)+(\rho+\delta)}\right]\cdot\Pi_1\cdot\Pi_3,$$

$$\gamma_{c3} = \left[\frac{\kappa(\rho+\delta+\phi_c)}{\kappa(\rho+\delta+\phi_c)+(\rho+\delta)}\right]\cdot\frac{\Pi_2}{(\rho+\Pi_2)}.$$

Note that plugging $\mu = 0$ into equation (12) of the paper and comparing that result to (S13) with $\mu_1 = 0$ shows that they are equivalent. Also, the claim that the effect of $\sum_i \eta_{ict}w_{ict}$ is less in the case where we control for housing costs can be seen easily, since the coefficient is Π_3 times the coefficient in the paper and $\Pi_3 < 1$. Their difference is increasing in μ_1.

Using the same procedure as outlined in the text of the main paper, we can derive the wage equation

$$w_{ic} = d_{it} + \left(\frac{\gamma_{c2}}{1-\gamma_{c2}}\frac{\gamma_{c1}}{\gamma_1}\right)\cdot\sum_i \eta_{ict}\nu_{it} + s\gamma_{c3}\left(1+\frac{\gamma_{c2}}{1-\gamma_{c2}}\right)p^h_{ct} + \zeta_{ict}, \tag{S14}$$

where

$$d_{it} = \left[\gamma_{c0} + \frac{\gamma_{c2}\gamma_{c0}}{1-\gamma_{c2}} + \gamma_{c1}\cdot p_{it} - \frac{\gamma_{c2}\gamma_{c1}}{1-\gamma_{c2}}\cdot p_{1t}\right],$$

$$\zeta_{ict} = -\gamma_{c3}\left(1 + \frac{\gamma_{c2}}{1-\gamma_{c2}}\right) \cdot \tau_{ct} + \left(\frac{\gamma_{c2}}{1-\gamma_{c2}}\gamma_{c1}\right) \cdot \sum_i \eta_{ict}\varepsilon_{ict} + \gamma_{c1}\varepsilon_{ict}.$$

Again, taking a linear approximation under the same conditions outlined in the main text gives

$$\Delta w_{ic} = \tilde{\alpha}_{it} + \tilde{\alpha}_2 \cdot \Delta R_{ct} + \tilde{\alpha}_3 \Delta\,\mathrm{ER}_{ct} + \tilde{\alpha}_4 \Delta p^h_{ct} + \Delta\tilde{\xi}_{ict}, \tag{S15}$$

where

$$\tilde{\alpha}_{it} = \left[\gamma_1 \cdot \Delta p_{it} - \frac{\gamma_2\gamma_1}{1-\gamma_2} \cdot \Delta p_{1t}\right],$$

$$\tilde{\alpha}_2 = \left(\frac{\gamma_{c2}}{1-\gamma_{c2}}\right),$$

$$\tilde{\alpha}_4 = s\gamma_3\left(1 + \frac{\gamma_2}{1-\gamma_2}\right),$$

$$\tilde{\xi}_{ict} = -\gamma_3\left(1 + \frac{\gamma_2}{1-\gamma_2}\right) \cdot \Delta\tau_{ct} + \left(\frac{\gamma_2}{1-\gamma_2}\gamma_1\right) \cdot \Delta \sum_i \frac{1}{I}\varepsilon_{ict} + \gamma_1 \Delta\varepsilon_{ict}.$$

S.7.2. *Equation for Housing Prices*

Letting $U^u_{\max} = U^u_c$, we can write the unemployment Bellman equation as

$$\rho U^u_{\max} = b + \tau_c - s \cdot p^h_c + \psi_c \cdot \left(\sum_i \eta_{ic} \cdot U^e_{ic} - U^u_c\right), \tag{S16}$$

and substituting in for $(\sum_i \eta_{ic} U^e_{ic} - U^u_c)$ gives

$$s \cdot p^h_c = b(1 - \Pi_1) + \tau_c + \Pi_1 \cdot \sum_i \eta_{ic} w_{ic} - \rho U^u_{\max}. \tag{S17}$$

Housing costs increase in τ_c, ψ_c, and average wages.

We can substitute this directly into the wage equation obtained above (S13):

$$\begin{aligned} w_{ic} = {} & \gamma_{c0} + \gamma_{c1} \cdot p_i + \gamma_{c2} \sum_i \eta_{ic} w_{ic} \\ & + \gamma_{c3} \cdot \left(b(1 - \Pi_1) + \tau_c + \Pi_1 \cdot \sum_i \eta_{ic} w_{ic} - \rho U^u_{\max}\right) \\ & - \gamma_{c3} \cdot \tau_c + \gamma_{c1}\varepsilon_{ic}. \end{aligned}$$

This gives back the exact same wage equation from the paper (equation (12)). For example, collecting the terms for the average wages gives the coefficient

$$\begin{aligned}\gamma_{c2} + \gamma_{c3}\Pi_1 &= \Lambda \cdot \Pi_1 \Pi_3 + \Lambda \cdot \Pi_1 \frac{\Pi_2}{\rho + \Pi_2} \\ &= \Lambda \cdot \Pi_1 \left(\frac{\rho}{\rho + \Pi_2} + \frac{\Pi_2}{\rho + \Pi_2} \right) \\ &= \Lambda \cdot \Pi_1 (1),\end{aligned}$$

where $\Lambda = [\frac{\kappa(\rho+\delta+\phi_c)}{\kappa(\rho+\delta+\phi_c)+(\rho+\delta)}]$. This result shows that under perfect mobility (i.e. $U^u_{\max} = U^u_c$), the direct and indirect mechanisms cannot be separately identified.

APPENDIX S.8: SELF-SELECTION INTO CITIES

As discussed in Appendix S.2, worker self-selection across cities is a potential concern for our identification strategy if changes in ΔR_{ct} are correlated with unobserved characteristics of workers. We deal with this issue in the main paper by implementing a selection correction procedure along the lines of Dahl (2002). This procedure requires the inclusion of variables in our first stage that capture the probability that a worker with a given set of characteristics chooses to live in his/her observed city. These terms turn out to be statistically significant in our first stage, which is a necessary condition for removal of any possible self-selection bias. However, this correction does not turn out to greatly impact the estimate of the coefficient on ΔR_{ct}. In Table S.VIII we replicate Table I without correcting for self-selection. As can be seen, the results are very similar with or without the self-selection correction.

APPENDIX S.9: FIRST-STAGE REGRESSIONS

In Table S.IX, we present the results of the first stage of the IV estimates in columns 4–6 from Table I. The estimates indicate that both IV1 and IV2 are strongly statistically significant predictors of ΔR_{ct}, but IV3 is not. The latter occurs because ΔR_{ct} is constructed as a pure composition measure. In the $\Delta\,\mathrm{ER}_{ct}$ equation, IV3 serves as a strong positive predictor (as expected), while IV1 and IV2 enter with significant and negative coefficients. The latter outcome fits with the bargaining model, since increases in wages generated by increases in R imply declining job creation. We explore the implications of the model for job creation in another paper.

APPENDIX S.10: THE REFLECTION SPECIFICATION

We note in Section 2.1 that wage determination takes the form of a classic reflection or social interaction problem. In particular, equation (10) makes it

TABLE S.VIII

BASIC RESULTS: WITHOUT SELECTION CORRECTION[a]

	OLS		IV			
	(1)	(2)	(3)	(4)	(5)	(6)
ΔR_{ct}	2.45*			2.80*	2.87*	2.85*
	(0.18)			(0.34)	(0.30)	(0.29)
ΔR^W_{ct}		2.10*	2.77*			
		(0.20)	(0.40)			
ΔR^B_{ct}		2.94*	2.93*			
		(0.40)	(0.42)			
$\Delta \mathrm{ER}_{ct}$	0.43*	0.48*	0.70	0.68	0.61	0.64
	(0.076)	(0.075)	(0.44)	(0.42)	(0.45)	(0.42)
Year × ind.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	34,375	34,375	34,375	34,375	34,375	34,375
R^2	0.49	0.49				
Instrument set			IV1–IV2–IV3	IV1–IV3	IV2–IV3	IV1–IV2–IV3
F-statistics						
ΔR^W_{ct}			82.63			
ΔR^B_{ct}			588.97			
ΔR_{ct}				70.56	160.74	224.38
$\Delta \mathrm{ER}_{ct}$			11.22	10.70	16.12	11.22
AP p-value						
ΔR^W_{ct}			0.00			
ΔR^B_{ct}			0.00			
ΔR_{ct}				0.00	0.00	0.00
$\Delta \mathrm{ER}_{ct}$			0.00	0.00	0.00	0.00
Over-id. p-value			.	.	.	0.79

[a]Standard errors, in parentheses, are clustered at the city–year level. The asterisk (*) denotes significance at the 5% level. All models are estimated on a sample of 152 U.S. cities using Census and ACS data for 1970–2007. The dependent variable is the decadal change in regression adjusted city–industry wages.

clear that wages in one sector of a city depend on average wages in that city across all sectors. As emphasized in the main text, equation (10) implies that a change in industrial composition that initially increases average wages in a city by 1%, would lead to a cumulative increase in average wages by a factor of $(\frac{1}{1-\gamma_2})\%$ due to the strategic complementarity of wages. Given that α_2 in equation (16) relates to γ_2 according to $\alpha_2 = \frac{\gamma_2}{1-\gamma_2}$, our estimates of α_2 that are around 2.9 in Table I suggest that γ_2 should be in the range of $0.74 = (\frac{2.9}{3.9})$. This implication of the model can also be examined directly by estimating equation (10) by instrumental variables. Since the coefficients in equation (10) depend on the employment rate, we again take a linear approximation of (10) around the point where the εs and Ωs are zero, and then take first differences to get a

TABLE S.IX

BASIC RESULTS: FIRST STAGE[a]

	Col4		Col5		Col6	
	(1) ΔR_{ct}	(2) $\Delta \mathrm{ER}_{ct}$	(3) ΔR_{ct}	(4) $\Delta \mathrm{ER}_{ct}$	(5) ΔR_{ct}	(6) $\Delta \mathrm{ER}_{ct}$
IV1	1.376*	−0.678*			1.042*	−0.528*
	(0.118)	(0.238)			(0.0946)	(0.237)
IV2			1.137*	−0.516*	0.998*	−0.445*
			(0.0721)	(0.161)	(0.0685)	(0.164)
IV3	−0.0332	0.160*	0.0810*	0.103*	−0.0114	0.150*
	(0.0196)	(0.0350)	(0.01000)	(0.0233)	(0.0115)	(0.0336)
Observations	33,984	33,984	33,984	33,984	33,984	33,984
R^2						

[a]Standard errors are given in parentheses. The asterisk (*) denotes $p < 0.05$.

linear equation of the form

$$\Delta \ln w_{ict} = \psi_1 d_{it} + \psi_2 \Delta \sum_j \eta_{jct} \ln w_{jct} + \psi_{3i} \Delta \mathrm{ER}_{ct} + \tilde{U}_{ict}, \tag{S18}$$

where ψ_2 corresponds to the γ_2 in the model and the error term corresponds to $\gamma_1 \Delta \varepsilon_{ict}$. In the case of equation (S18), estimation by OLS would definitely be expected to give upward biased estimates of $\psi_2 = \gamma_2$, since the relationship suffers from the reflection problem. However, it can be verified that instrumental variable estimation of equation (S18) using our previous set of instruments should give consistent estimates under the same assumption as before; that is, under the assumption that the common component of the εs (a city's absolute advantage) is independent of the past. It is worth emphasizing that the difference between (16) and (S18) pertains only to the main variable of interest. In (S18), this variable is the average city wage, while in (25) it is a city-level average of national wage premia.

Estimates of equation (S18) are present in Table S.X. The first thing to note in this table is that, as should be expected, there is now a large and significant difference between estimates of ψ_2 (denoted by Δw_{ct}) obtained by OLS and IV. The OLS estimate is 0.86, which if translated to compare with α_2 would imply $\alpha_2 = 6.14 = \frac{0.852}{1-0.852}$. However, in this case there are no conditions for which we should expect OLS to give consistent estimates. In contrast, when we estimate by IV, we get an estimate of ψ_2 equal approximately to 0.72, which implies a value of $\alpha_2 = 2.57 = \frac{0.72}{1-0.72}$, which is very close to that obtained in the main text using a different approach. In particular, recall that the estimation of (25) by OLS provides one means to overcome the reflection problem by focusing on national-level wage premia, while the IV estimation of (S18) provides a conceptually quite different approach.

TABLE S.X

REFLECTION SPECIFICATION[a]

	OLS	IV		
	(1)	(2)	(3)	(4)
Δw_{ct}	0.86*	0.69*	0.74*	0.72*
	(0.013)	(0.033)	(0.030)	(0.028)
$\Delta \mathrm{ER}_{ct}$	0.069*	0.11	−0.11	−0.022
	(0.029)	(0.16)	(0.20)	(0.16)
Year × ind.	Yes	Yes	Yes	Yes
Observations	33,984	33,984	33,984	33,984
R^2	0.58			
Instrument set				
F-statistics				
Δw_{ct}		45.18	74.63	69.42
$\Delta \mathrm{ER}_{ct}$		10.41	15.78	10.99
AP p-value				
ΔW_{ct}		0.00	0.00	0.00
$\Delta \mathrm{ER}_{ct}$		0.00	0.00	0.00
Over-id. p-value		.	.	0.10

[a]Standard errors, in parentheses, are clustered at the city–year level. The asterisk (*) denotes significance at the 5% level. All models are estimated on a sample of 152 U.S. cities using Census and ACS data for 1970–2007. The dependent variable is the decadal change in regression-adjusted city–industry wages.

REFERENCES

CAHUC, P., AND A. ZYLBERBERG (2004): *Labor Economics*, Vol. 1. Cambridge: MIT Press. [16-18]

DAHL, G. B. (2002): "Mobility and the Return to Education: Testing a Roy Model With Multiple Markets," *Econometrica*, 6, 2367–2420. [4-6,24]

DEATON, A., AND D. LUBOTSKY (2001): "Mortality, Inequality and Race in American Cities and States," Working Papers 8370, National Bureau of Economic Research, Inc. [2]

MORETTI, E. (2010): "Local Multipliers," *American Economic Review*, 100 (2), 373–377. [2]

PARK, J. H. (1994): "Estimation of Sheepskin Effects and Returns to Schooling Using the Old and the New CPS Measures of Educational Attainment," Papers 338, Princeton, Department of Economics—Industrial Relations Sections. [1]

RUGGLES, S., M. SOBEK, T. ALEXANDER, C. A. FITCH, R. GOEKEN, P. K. HALL, M. KING, AND C. RONNANDER (2004): *Integrated Public Use Microdata Series: Version 3.0 [Machine-readable database]*. Minneapolis, MN: Minnesota Population Center. [1]

Dept. of Economics, University of British Columbia, 997-1873 East Mall, Vancouver, B.C., Canada, V6T 1Z1 and NBER; paulbe@interchange.ubc.ca,

Dept. of Economics, University of British Columbia, 997-1873 East Mall, Vancouver, B.C., Canada, V6T 1Z1 and IFS, London; green@econ.ubc.ca,

and

Dept. of Economics, York University, 4700 Keele Street, Toronto, ON, Canada, M3J 1P3; bmsand@econ.yorku.ca.

Manuscript received June, 2009; final revision received July, 2011.