Labor Market Power*

David Berger Kyle Herkenhoff Simon Mongey

October 5, 2019

Abstract

We develop a tractable quantitative, general equilibrium, oligopsony model of the labor market that we use to measure the macroeconomic implications of labor market power. Strategic interaction complicates inference of parameters that are key to this exercise. To address this challenge, we contribute estimates of market share dependent wage and employment responses to state corporate tax changes in U.S. Census data, which we combine with the structure of the model. We validate against the distribution of local labor market concentration and quasi-experimental evidence on productivity-wage pass-through. Relative to a counterfactual competitive economy, and accounting for transition dynamics, we measure welfare losses from labor market power to be roughly 5 percent of lifetime consumption. Minimum wage and merger experiments caution that concentration and welfare may not negatively comove.

JEL codes: E2, J2, J42

Keywords: Labor markets, Market structure, Oligopsony, Strategic interaction

*Berger: Duke University. Herkenhoff: Federal Reserve Bank of New York and University of Minnesota. Mongey: Kenneth C. Griffin Department of Economics, University of Chicago. We thank Costas Arkolakis, Chris Edmond, Oleg Itskohki, Matt Notowidigdo, Richard Rogerson, Jim Schmitz, and Chris Tonetti, for helpful comments. We thank seminar participants at the SED 2018, SITE, Briq (Bonn), FRB Minneapolis, Princeton, Stanford, UC Berkley, Rochester, BFI Firms and Workers Workshop, Paris School of Economics, Toulouse School of Economics, EIEF, Barcelona GSE Summer Forum. This research was supported by the National Science Foundation (Award No. SES-1824422). Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. The views expressed in this study are those of the author and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

A recent empirical literature has documented important deviations from textbook perfectly competitive labor markets, leading to a growing concern that "labor market power" may generate large welfare losses.[1] One intuitive source of market power is that there may be few firms in a local labor market and these firms understand that their hiring and wage setting decisions affect the local labor market's overall wage and employment levels. Firms that have a significant impact on local labor market conditions and internalize this fact, maximize profits by hiring fewer workers in order to pay lower wages.

In this paper, we contribute a tractable, quantitative, general equilibrium model of oligopsony in the labor market that delivers a structurally consistent formulation of this notion of labor market power. The model then guides how we *measure* labor market power. Strategic interaction complicates inference of parameters that are key to this exercise. To address this challenge, our identification strategy integrates new reduced form estimates of market share dependent wage and employment responses to state corporate tax changes—obtained by extending Giroud and Rauh (2019) in U.S. Census data—into a structural estimation of the model. We then use the model to measure the welfare and macroeconomic implications of labor market power.

We have three main results. First, the model implies substantial welfare losses from labor market power, both across steady states and along the transition path to a competitive economy, which range from 3 to 8 percent of lifetime consumption. Second, despite these large losses, we find that labor market power has not contributed to the declining labor share. Despite the backdrop of an overall increase in *national* concentration, we find that the model-consistent measure of *local* labor market concentration has actually declined over the last 35 years, indicating that most local labor markets are actually more competitive than they were in the 1970s. Third, we draw lessons from two counterfactuals that are important to discussions of labor market power: a minimum wage and mergers. Our key finding is that both counterfactuals yield a non-monotonic relationship between concentration and welfare. We conclude that an observed increase in concentration in a market cannot be used to make statements about welfare without understanding driving forces.

Our model accommodates two natural features of the data. Local labor markets are concentrated. We therefore depart from standard models of atomistic monopsony (Burdett and Mortensen, 1998; Manning, 2003; Card, Cardoso, Heining, and Kline, 2018; Lamadon, Mogstad, and Setzler, 2019) by incorporating strategic interaction between a finite set of employers in each labor market (Atkeson and Burstein, 2008). The macroeconomy consists of many local labor markets. We therefore provide a general equilibrium structure that aggregates these markets.

Our benchmark oligopsony model features two sources of market power. The first is classical *monopsony*: firms face upward sloping labor supply curves, which they internalize. Optimal wages are a markdown relative to competitive wages, i.e. the marginal revenue product of labor. Second, and the focus of this paper, is *oligopsony*: firms are non-atomistic and compete strategically for workers. Equilibrium markdowns are wider at the firms with the most labor market power. Therefore, understanding the welfare consequences of labor market power requires understanding how these markdowns vary across

[1] For example: Azar, Marinescu, and Steinbaum (2017), Benmelech, Bergman, and Kim (2018), Card, Cardoso, Heining, and Kline (2018), and Lamadon, Mogstad, and Setzler (2019).

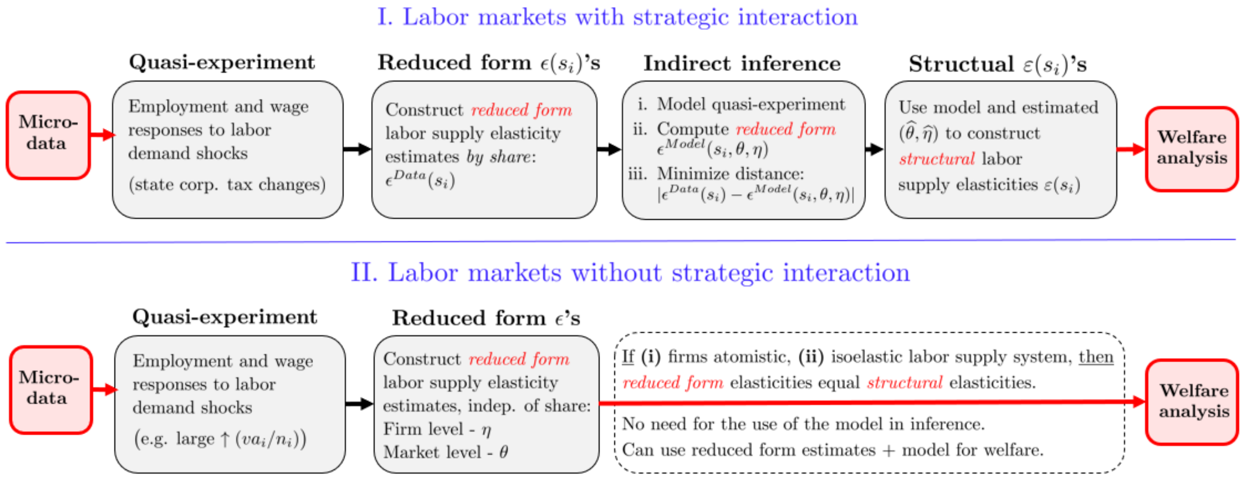

Figure 1: Estimation strategy

firms. In our model the markdown is an exact function of the *structural labor supply elasticity* that a firm faces in equilibrium which—via a closed form—depends on the firm's observable labor market share and parameters that determine substitutability of labor across- (θ) and within- (η) markets.

We estimate these key parameters using U.S. Census Longitudinal Business Database (LBD) micro data (see Figure 1). Given a quasi-experiment that yields an identified shock to labor demand, a researcher can estimate *reduced form labor supply elasticities* off of relative employment and wage responses. The literature so far has assumed a special case of our model: firms do not behave strategically, rationalized by infinitely many firms in each labor market.[2] We show that in abstracting from competitor best responses this assumption implies that empirically estimated *reduced form elasticities* are equal to *structural elasticities*, so one can move directly from empirical analysis to welfare analysis. However, in the more general case of finitely many firms, there is no closed form mapping between (observed) reduced form elasticities and (unobserved) structural elasticities.[3] A model is needed to account for the equilibrium best responses that determine the mapping between underlying structural parameters and the reduced form elasticities we observe.

Our approach is therefore indirect inference. Our quasi-experiment is an extension of Giroud and Rauh (2019). We exploit corporate tax rate changes to estimate *reduced form elasticities*, extending current methodology to characterize how they relate to a firm's share of the local labor market. We then simulate this quasi-experiment in our model and find the structural parameters that minimize the distance between the profile of reduced form elasticities by market share in model and data. We also match concentration in the data, allowing for the possibility that the assumption of unconcentrated markets is valid. The estimated model is then used to compute structural elasticities, markdowns, and conduct

[2]Papers in the literature that study strategic behavior have been theoretical, which we discuss below.

[3]The finitely many firms case is indeed more general. That is, a 'competitive' monopsony model is indeed a special case of our model. Taking the number of firms in all markets in our model toward infinity smoothly yields the 'competitive' economy in which there is no strategic interaction. We let the data tell us where we are on this spectrum between one and infinitely many firms per market.

welfare counterfactuals.[4]

This departure from the literature contributes three additional results. First, in the data, we identify systematically different responses of firms to labor demand shocks: at firms with smaller market shares, we estimate larger reduced form labor supply elasticities. Second, in our particular experiment *reduced form elasticities* at small firms are around 2, yet welfare-relevant *structural elasticities* are around 5. Interpreting the data through the model is necessary for quantifying the high labor supply elasticities faced by small firms. Third, we explore bias in more common empirical settings that exploit purely idiosyncratic variation. Here results are different; when we account for competitors' best responses, the *structural elasticities* faced by firms are always less than implied by empirically estimated *reduced form elasticities*, often by a large amount. Since reduced form estimates overstate structural elasticities, a researcher using the former for welfare analysis would infer flatter labor supply curves that *understate* the degree of labor market power.

We validate the estimated model against two sets of non-targeted moments that enter the discussion of labor market power: (1) the weighted and unweighted distribution of concentration across markets, and (2) wage-pass through. First, in both model and data, the *payroll weighted* average of market wage-bill Herfindahls (HHI) is significantly lower than the *unweighted* average. The ratio, which we do not target, is approximately 3.2 in the data and 2.4 in the model: concentrated markets are small markets.[5] Second, we carefully replicate the quasi-experiment that identifies reduced-form estimates of pass-through in Kline, Petkova, Williams, and Zidar (2019). Pass-through from value added per worker to wages is 79.5 percent, which is larger but conceptually comparable to the 47.0 percent estimate of Kline, Petkova, Williams, and Zidar (2019). In classical monopsony models that ignore strategic complementarities, this measure of pass-through would be 100 percent.[6] That the model reproduces non-degenerate pass-through rates from value added per worker to wages is evidence that strategic interactions play an important role in wage setting.

With our model calibrated to aggregates and local labor markets, we define the *welfare loss* due to labor market power as the consumption subsidy required to make households indifferent between the oligopsonistic economy and a competitive equilibrium. Comparing steady states at an aggregate Frisch elasticity of labor supply of 0.5, we measure a welfare loss of 5.4 percent.[7] Losses are slightly lower (4.1 percent) when accounting for macroeconomic transition dynamics between these two labor market structures. Competitive equilibrium wages, output and labor supply are significantly greater.

The competitive equilibrium also features higher concentration. Roughly half of the output gains in

[4]This procedure has a direct counterpart in the estimation of linearized macroeconomic state-space systems: $A\mathbf{X}_t = B\mathbb{E}[\mathbf{X}_{t+1}] + C\mathbf{X}_{t-1} + D\varepsilon_t$. The *structural model* implies a *reduced form* VAR representation: $\mathbf{X}_{t+1} = H\mathbf{X}_t + F\mathbf{e}_{t+1}$ The researcher then estimates the reduced form on the data to obtain *reduced form* shocks $\{\mathbf{e}_t\}_{t=0}^T$. The researcher then simulates *structural shocks* $\{\varepsilon_t\}_{t=0}^T$ in the model and jointly estimates *structural parameters* $\{A, B, C, D\}$ and *structural shocks* $\{\varepsilon_t\}_{t=0}^T$ such that the model implied reduced form shocks match those obtained from the data.

[5]In the data, markets with only one firm—and so an HHI of one—account for 15 percent of market but only 0.4 percent of employment, so are uninformative of labor market conditions faced by most workers. Our theory consistent measure of concentration that reflects this. In the model, employment in these regions is small as monopsonists pay low wages and hire few workers. Figure 7 provides the distribution of weighted and unweighted market concentration in both model and data.

[6]See papers described in Manning (2003) and Card, Cardoso, Heining, and Kline (2018).

[7]Under an aggregate Frisch of 0.2 (0.8), welfare losses are 3 (8).

competitive equilibrium are driven by a reallocation of workers from smaller, less productive firms to larger, more productive firms. In the oligopsonistic economy, large firms have more labor market power, so are inefficiently small. A by-product of this efficient reallocation is a sharp *increase* in concentration, along with significant increases in welfare, consumption, and output.

Despite large welfare losses from labor market power, we find that declining labor market concentration between 1976 and 2014 *increased* labor's share of income. First, letting our model guide measurement, we show that the distribution of market-level wage-bill Herfindahls is a sufficient statistic for labor's share of income.[8] Second, the model implies that when aggregating these micro measures, market-level payroll shares should be used as weights. We construct this model relevant concentration measure directly from the Census LBD and find it to have declined from 0.20 to 0.14 between 1976 and 2014.[9] Combined with our estimates of model parameters (θ, η), and exploiting a closed form link between our sufficient statistic and the labor share, this decline in concentration would have implied a counterfactual 3.13 percentage point increase in labor's share of income. Changing labor market concentration is not behind the declining labor share.

Given large welfare losses from labor market power in the U.S., we study two applications—a national minimum wage and mergers. The former is often proposed as an antidote to labor market power; the latter a potential culprit for changes in labor market power.

First, minimum wages have been studied extensively in applied labor economics (e.g. Card and Krueger, 1994) and historically motivated the development of monopsonistic models (Boal and Ransom, 1997; Manning, 2003). We provide a new theoretical characterization of how minimum wages affect firm-level and worker-level behavior in an environment with decreasing returns to scale and strategic complementarities. Workers reallocate away from unproductive fims as the minimum wage increases. Efficacy of the minimum wage trades off reallocating workers away from small, unproductive firms, against the additional market power this delivers to larger firms. These opposing forces lead welfare to be hump-shaped in the minimum wage, despite monotonically increasing concentration.

In the second application of our model, we consider mergers. On the one hand, merging the two largest firms in each market *decreases* welfare and increases concentration. On the other hand, merging the two smallest firms in each market *increases* welfare, but still increases concentration. Merged firms have more market power, so lower wages and employment. If the merging firms are productive (unproductive) this increases (decreases) misallocation, decreasing (increasing) output. Our estimates imply that mergers between productive firms could generate maximum welfare losses of up to 2 percent. We further explore merger empirics, welfare and optimal policy in Berger, Herkenhoff, and Mongey (2019). These applications, and our competitive equilibrium counterfactual caution strongly against making inferences about changes in welfare from changes in concentration.

We next review the literature and then proceed as follows. Section 1 provides new statistics characterizing the evolution of U.S. labor market concentration. Sections 2 and 3 lay out the model and

[8]The market-level wage-bill Herfindahl is the sum of the squared payroll shares of all firms within the labor market

[9]The effective number of firms in a typical labor market was equivalent to 5.0 equally sized firms per market in 1976, and 7.1 equally sized firms per market in 2014.

characterize the equilibrium. In Section 4 we provide empirical estimates of the relationship between reduced form labor supply elasticities and market share, then combine this relationship and our new concentration statistics to parameterize the model. Section 5 validates the model against cross-sectional concentration statistics and pass-through estimates. Section 6 presents our two measurement exercises: welfare and the labor share. Section 7 applies the model to study minimum wage and merger policy.

Literature. Our work is related to a growing literature that explores the implications of market power. In the product market, Gutiérrez and Philippon (2016); Autor, Dorn, Katz, Patterson, and Van Reenen (2017) all document an increase in national sales concentration and a fall in the labor share across many industries, while De Loecker and Eeckhout (2017) document an increase in product market power more directly by measuring firm markups. Consistent with our findings, concurrent work by Rossi-Hansberg, Sarte, and Trachter (2018) document declining regional employment concentration, despite rising national concentration. In the labor market, several concurrent studies have documented cross-sectional and time-series patterns of U.S. Herfindahls in employment (Benmelech, Bergman, and Kim, 2018; Rinz, 2018; Hershbein and Macaluso, 2018) and vacancies (Azar, Marinescu, Steinbaum, and Taska, 2019; Azar, Marinescu, and Steinbaum, 2017). Brooks, Kaboski, Li, and Qian (2019) combine theory and Chinese and Indian data to study monopsony by a firm or collusive group of firms and find adverse effects of monopsony power on labor's share of income. Our contributions to this literature are (i) a new, model consistent, measure of U.S. labor market concentration, which we use to (ii) quantitatively measure the welfare losses associated with labor market power. In general the exercises in our paper issue a warning against qualitatively mapping changes in concentration into a change in welfare.

Our work is also related to a large literature which has measured reduced form labor supply elasticities of individual firms (Staiger, Spetz, and Phibbs, 2010; Card, Cardoso, Heining, and Kline, 2018; Suárez Serrato and Zidar, 2016; Dube, Jacobs, Naidu, and Suri, 2019). We provide new estimates of measured labor supply elasticities by building on the approach of Giroud and Rauh (2019), who find significant effects of state corporate taxes on firm-state employment.[10] Our contributions to this empirical literature are (i) estimates of the share-dependency of measured elasticities which highlight that large firms have more market power (ii) to demonstrate that if markets have firms that interact strategically, there can be a large disconnect between measured labor supply elasticities and the structural elasticities that are relevant for welfare. This is a substantive point: the empirical literature cited above typically measures labor supply elasticities that are small. If structural elasticities were equal to these reduced form elasticities, then labor market power would be extremely high.[11] We describe empirical designs under which (i) reduced form estimates of labor supply elasticities may be biased downwards relative to structural elasticities, and even then, (ii) that structural elasticities vary systematically with the firm's labor market share, reconciling the range and level of empirical estimates.

[10]Conceptually, our approach is related to papers that estimate exchange rate pass-through (Amiti, Itskhoki, and Konings, 2014, 2016). The main difference is this literature This literature focuses exclusively on prices whereas we look at both price and quantity responses.

[11]Consider Manning (2011) discussing the widely cited natural experiments of Staiger, Spetz, and Phibbs (2010) and others: *"Looking at these studies, one clearly comes away with the impression not that it is hard to find evidence of monopsony power but that the estimates are so enormous to be an embarrassment even for those who believe this is the right approach to labour markets."*

Finally, our work is related to the large literature that models monopsony in labor markets. We depart from benchmark models of monopsony described in (Burdett and Mortensen, 1998; Manning, 2003; Card, Cardoso, Heining, and Kline, 2018; Lamadon, Mogstad, and Setzler, 2019) by explicitly modeling a finite set of employers that compete strategically for workers. We demonstrate that this addition is crucial for identification: strategic interaction and finiteness of firms jointly imply that reduced form empirical estimates of labor supply elasticities from *any* shock cannot be used to infer the (structural) labor supply elasticities firm face—and hence identify preference parameters—except in the limiting case of monopsonistic competition between infinitesimally sized firms. Additionally, our assumptions allow us to (i) interpret granular measures of concentration, such as Herfindahl indexes, and (ii) accommodate a Walrasian equilibrium as a meaningful counterfactual, in which strategic behavior is shut down.

Our main quantitative contribution is to build a general equilibrium model of oligopsony and measure the welfare costs of current levels of U.S. labor market power.[12] Our framework extends the general tools developed in Atkeson and Burstein (2008) to the labor market, adding multiple non-trivial features: capital, corporate taxes, decreasing returns to scale, set the model in general equilibrium, and study transition dynamics between steady states. Recent related work by Jarosch, Nimcsik, and Sorkin (2019) considers non-atomistic firms, but adapts a random search model to construct a search-theoretic measure of labor market power. We view our papers as complementary.

Our model features firm-specific upward sloping labor supply curves. This is supported by numerous recent studies using (quasi-)experimental approaches.[13] Belot, Kircher, and Muller (2017) randomly assign higher wages to observationally equivalent vacancies on an actual job-board and find that higher wage vacancies attract more applicants. Dube, Jacobs, Naidu, and Suri (2019) and Banfi and Villena-Roldan (2018) also find job-specific upward sloping labor supply curves in well-identified contexts.[14]

Finally, our quantitative model features strategic complementarity between oligopsonists. Strategic complementarity in labor markets is not new to the theoretical literature. The earliest models used to motivate monopsony power were Robinson (1933) and the spatial economies of Hotelling (1990) and Salop (1979).[15] Our contribution, relative to these stylized single-market models, is a quantitative general equilibrium framework. We incorporate firm heterogeneity, decreasing returns to scale, and general equilibrium across multiple markets, making it rich enough to be estimated on U.S. Census data. Moreover, by modeling a finite set of employers, our model may be used in the future to understand the wage and welfare effects of mergers, firm exit, and other shocks to local labor market competition.

[12]Our work is therefore related to a literature measuring the welfare consequences of misallocation. There the focus has been on the product market (Baqaee and Farhi, 2017; Edmond, Midrigan, and Xu, 2018), and measures misallocation via heterogeneous markups. Our paper measures misallocation from heterogeneous mark-downs.

[13]See Ashenfelter, Farber, and Ransom (2010) for a summary of prior papers.

[14]We are unaware of experimental evidence regarding the market-share dependence of the elasticity of labor supply.

[15]Boal and Ransom (1997) and Bhaskar, Manning, and To (2002) provide excellent summaries of strategic complementarity in spatial models of the labor market.

1 Labor market concentration: 1976 and 2014

We provide new statistics summarizing labor market concentration in 1976 and 2014 using the Census Longitudinal Business Database (LBD).[16] The LBD provides high quality measures of employment, location, and industry with nearly universal coverage of the non-farm business sector. Data are carefully linked over time at the establishment and firm level.

Market. In order to compute concentration we must define a market. In our model, a market will have two features: (i) a worker drawn at random from the economy will have a greater attachment to one market than others on the basis of idiosyncratic preferences, but will be able to move across markets nonetheless, and (ii) firms within a market compete strategically.

With these assumptions in mind and given what we can observe in the LBD, we define a *local labor market* as a 3-digit NAICS (NAICS3) industry within a Commuting Zone (CZ).[17] Examples of adjacent 3-digit NAICS codes are subsectors 323-325: *'Printing and Related Support Activities'*, *'Petroleum and Coal Products Manufacturing'* and *'Chemical Manufacturing'* which we regard as suitably different. Examples of commuting zones include the collection of counties surrounding downtown Minneapolis and those surrounding Duluth.[18]

Industries. A key step in our analysis is to restrict our attention to tradeable goods industries. Our aim is to cleanly study labor market power without the potential confounds of product market power. Our assumption is that the spot market for tradeable goods is outside the local labor market, which we make explicit in our model. We keep the industries specialized in tradeable goods as identified by Delgado, Bryden, and Zyontz (2014).[19] Appendix D verifies that the trends we report for these sectors are also true for the economy as a whole.

Firm. Finally, we define a firm in a local labor market as the collection of establishments operated by that firm. We aggregate employment and annual payroll of all establishments owned by the same firm within the same NAICS3-CZ market.[20] For each resulting firm-market-year observation, we observe employment, payroll, and herein define the *wage* as payroll per worker. Appendix C provides more details on the sample restriction and data definitions.

[16] 1976 and 2014 are the first and last years of data availability in the LBD 'snapshot' for, which our project had access. At this point in the review process for this paper, we have only disclosed these two years, but in future revisions are able to release and plot the full time series between these points. For additional information regarding the data sources in this paper see Appendix C.

[17] Using BLS Occupational Employment Statistics microdata, Handwerker and Dey (2018) show that when it comes to concentration there is little practical difference in defining a market at the occupation-city level rather than the industry-city level as these two measures are highly correlated. In particular the across-city correlation of Herfindahl-Hirschman Indices at the CBSA-occupation and CBSA-industry level is 0.97.

[18] Many more examples are provided in Tables C1 and C2 in Appendix C.

[19] See their Table 2, in which they rank 2-digit NAICS sectors by the fraction of employment in tradeable sub-industries. The 2 digit industries we use have more than 70 percent of their employment in tradeable sub-industries. These include 11 (Agriculture), 21 (Mining), 31-33 (Manufacturing), 42 (Wholesale trade) and 55 (Management). When identifying industries throughout the paper, we use the time consistent 2007 NAICS codes provided by Fort and Klimek (2016).

[20] Firms are identified by the LBD variable *firmid*.

A. Firm-market-level averages	1976	2014
Total firm pay ($1,000's)	470.90	1839.00
Total firm employment	37.09	27.96
Pay per employee	$ 12,696	$ 65,773
Firm-market level observations	660,000	810,000

B. Market-level averages	1976	2014
Wage-bill HHI (Unweighted)	0.45	0.45
Employment HHI (Unweighted)	0.43	0.42
Wage-bill HHI (Weighted by market's share of total payroll)	0.20	0.14
Employment HHI (Weighted by market's share of total payroll)	0.17	0.11
Wage-bill HHI (Weighted by market's share of total employment)	0.19	0.14
Employment HHI (Weighted by market's share of total employment)	0.18	0.12
Firms per market	42.6	51.6
Percent of markets with 1 firm	14.6%	14.7%
National employment share of markets with one firm	0.63%	0.36%
Market level observations	15,000	16,000

C. Across market correlations with wage-bill HHI	1976	2014
Number of firms	-0.22	-0.21
Market employment	-0.20	-0.21
Employment Herfindahl	0.98	0.98
Standard deviation of relative wages	-0.49	-0.51
Market level observations	15,000	16,000

Table 1: Summary Statistics, U.S. Census Longitudinal Employer Database 1976 and 2014

Notes: Tradeable NAICS2 codes (11,21,31,32,33,42,55). Market defined to be NAICS3 within Commuting Zone. Observations rounded to nearest thousand and numbers rounded to 4 significant digits according to Census disclosure rules. Firm-market-level refers to a 'firmid by Commuting Zone by 3-digit NAICs by Year' observation. Market-level refers to a 'Commuting Zone by 3-digit NAICs by Year' aggregation of observations.

Summary statistics. Table 1A describes characteristics of the firm-market observations. Average nominal payroll was $470,900 in 1976 and $1,839,000 in 2014. Average firm-market employment was 37 workers in 1976 and 28 workers in 2014. Average wage was $12,696 in 1976 and $65,773 in 2014.[21]

Concentration. In Appendix D, we show that non-tradeable industries also display the same patterns we now describe for tradeable industries. First, we consider two measures: Herfindahl indexes for payroll and employment. Let i denote a firm and j denote a market. Let w_{ij} and n_{ij} denote the firm's wage and employment in market j, respectively. Equation (1) defines the wage-bill Herfindahl: the sum of the squared wage-bill shares. Our model will reveal this to be the welfare relevant measure of market concentration.

$$HHI_j^{wn} := \sum_{i \in j} \left(s_{ij}^{wn} \right)^2 \quad , \quad s_{ij}^{wn} = \frac{w_{ij} n_{ij}}{\sum_{i \in j} w_{ij} n_{ij}} \tag{1}$$

[21] Real wage growth in these industries was relatively low at 0.6 percent per annum. By our statistics, nominal wage growth was 4.42 percent p.a., while the CPI increased at a rate of 3.82 percent p.a. over the same period.

Equation (2) defines the employment Herfindahl.

$$HHI_j^n := \sum_{i\in j}\left(s_{ij}^n\right)^2 \qquad , \qquad s_{ij}^n = \frac{n_{ij}}{\sum_{i\in j} n_{ij}} \tag{2}$$

This measure does not capture the positive empirical covariance between wages and employment and so tends to be less than the wage-bill Herfindahl.

Different weighting schemes of across-market averages imply different levels and trends. The unweighted average Herfindahls for wages and employment are between two and three times larger than their counterparts weighted by either employment or payroll. Little employment or payroll is located in highly concentrated markets: in both periods, 14 percent of markets have only one employer and so HHIs equal to one. This is a distinct statistical property of labor market concentration that our model will reproduce. In terms of the time-series, unweighted average payroll and employment Herfindahls are approximately unchanged between 1976 and 2014. When weighted by market payroll, the payroll (employment) wage-bill Herfindahl declines from 0.20 to 0.14 (0.17 to 0.11). Aggregation in our model will imply that welfare relevant concentration measures should be the payroll weighted measures, which declined. In Appendix D, we show that these patterns are consistent in non-tradeable industries.

Table 1C confirms that the number of firms and total market employment are negatively correlated with concentration. This is important for understanding why weighted and unweighted Herfindahls are so different and will be used as an over-identifying test of the estimated model. Moreover, employment and wage-bill Herfindahls are highly correlated. We also document that more concentrated markets have less dispersed wages: the correlation between concentration and within market dispersion of relative wages is strongly negative.[22]

When estimating our model we will target a single concentration measure. We then use many of these other moments as over-identifying tests of the quantitative relevance of our theory.

Summary. Figure 2 summarizes changes in concentration. To interpret the payroll weighted Herfindahls in Panel A, Panel B plots the *inverse* of the wage-bill and employment Herfindahls. The Inverse Herfindahl Index (IHI) can be interpreted as the effective number of firms competing in a market.[23] The payroll weighted wage-bill Herfindahl implies that the effective number of firms in tradeable U.S. labor markets increased from 5.01 in 1976 to 7.09 in 2014: concentration fell. By comparison, in the raw data, we observe a 20 percent increase in the average number of firms per market.[24]

To map these measures of labor market concentration to welfare, we require a model. Our theoretical framework is specifically designed to accommodate these commonly used statistics. In fact, within our framework wage-bill Herfindahls and knowledge of key structural parameters are sufficient to compute

[22] A firm's relative wage is defined by $w_{ij}^{rel} := w_{ij}/(\sum_{i\in j} w_{ij}/M_j)$, where M_j is the number of firms in market j. We then compute the standard deviation of this term within each market j.

[23] If three firms operate in a market and have equal shares, then the Herfindahl is $1/3 = \sum_{i\in j}(1/3)^2$. So a market with M_j firms of different sizes and a Herfindahl of $1/x$ has the same level of concentration as a market with x firms of equal size.

[24] Rinz (2018) describes employment concentration in a number of non-tradeable sectors using a NAICS4×Commuting zone definition of a labor market. Tradeable and non-tradeable sectors do not have systematically different *levels* of concentration (his Figure 11). Rinz (2018) does not aggregate establishments within firms when computing employment shares. When averaged at the 2-digit level, he finds similar trends in tradeable and non-tradeable sectors.

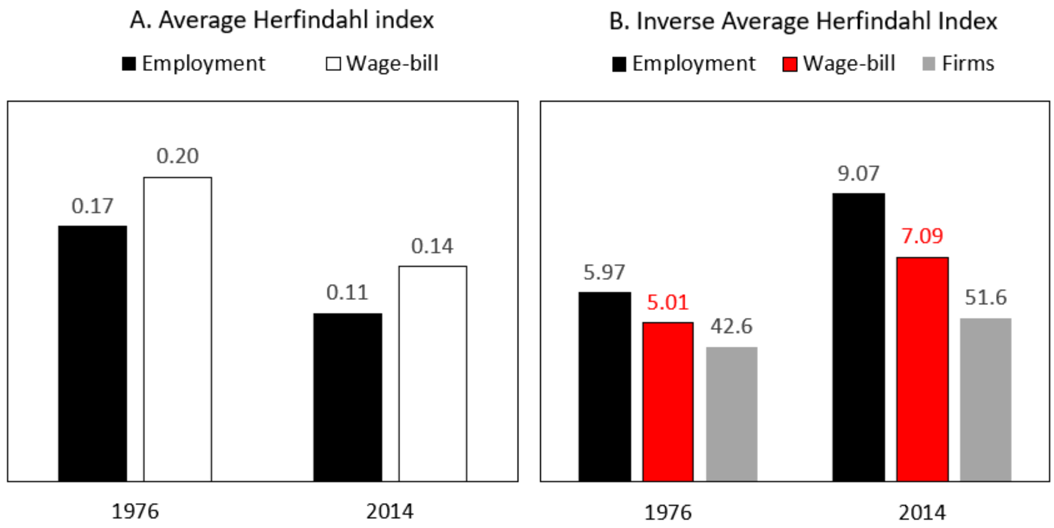

Figure 2: Labor Market Concentration, 1976 and 2014

Notes: Panel A plots employment and payroll Herfindhals in the Census LBD in 1976 and 2014 for tradeable firms (see notes to Table 1). Panel B plots the inverse of these Herfindhals along with firms per CZ×NAICS3 market. See text for details.

the share of aggregate income paid to workers. From the point of view of measurement, we can also quantify how other measures of concentration that are not welfare relevant may bias inferences regarding welfare.

2 Model

2.1 Environment

Agents. The economy consists of a representative household and a continuum of firms. Firms are heterogeneous in two dimensions. First, firms inhabit a continuum of local labor markets $j \in [0,1]$, within which there exists an exogenously given finite number of firms indexed $i \in \{1,2,...M_j\}$. Second, firms' productivity $z_{ijt} \in (0,\infty)$ are drawn from a location invariant distribution $F(z)$. The *only ex-ante difference* between markets is the number of firms $M_j \in \{1,\ldots,\infty\}$. Time subscripts are necessary in that we study welfare counterfactuals on transition paths between steady-states, but at the firm and market level productivity and number of firms are constant.

Goods and technology. The continuum of firms produce goods that are perfect substitutes, and so trade in a perfectly competitive national market at a price P_t that we normalize to one. Firms operate a *value-added* production function that uses inputs of capital k_{ijt} and labor n_{ijt}.[25] Let $\overline{Z}$ be a common component of productivity across firms. A firm produces y_{ijt} units of net-output (value-added) according

[25]Since aggregating firm-level value-added yields aggregate output (GDP), we abuse terminology and refer to the output of this production function interchangeably in terms of goods and value-added. We carefully distinguish the two when comparing our results to empirical studies.

to the production function:

$$y_{ijt} = \overline{Z} z_{ijt} \left(k_{ijt}^{1-\gamma} n_{ijt}^{\gamma} \right)^{\alpha} \quad , \quad \gamma \in (0,1) \quad , \quad \alpha > 0.$$

The degree of returns to scale α is unrestricted and later estimated. These goods may be used for consumption or investment. Investment augments the capital stock K_t which is owned by the representative household, rented to firms in a competitive market at price R_t and depreciates at rate δ. To model imperfect labor market competition with decreasing returns to scale production, we extend tools developed in the trade literature (Atkeson and Burstein, 2008).

2.2 Household

Preferences and problem. The household chooses the measure of workers to supply to each firm, n_{ijt}, investment in next period capital K_{t+1}, and consumption of each good c_{ijt} to maximize their net present value of utility. Given an initial capital stock K_0, the household solves

$$\mathcal{W}_0 = \max_{\{n_{ijt}, c_{ijt}, K_{t+1}\}} \sum_{t=0}^{\infty} \beta^t u \left(\mathbf{C}_t - \frac{1}{\overline{\varphi}^{\frac{1}{\varphi}}} \frac{\mathbf{N}_t^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}} \right) \quad , \quad \beta \in (0,1) \quad , \quad \varphi > 0 \tag{3}$$

where the aggregate disutility of labor supply is given by,

$$\mathbf{N}_t := \left[\int_0^1 \mathbf{N}_{jt}^{\frac{\theta+1}{\theta}} \, dj \right]^{\frac{\theta}{\theta+1}} \quad , \quad \theta > 0$$

$$\mathbf{N}_{jt} := \left[n_{1jt}^{\frac{\eta+1}{\eta}} + \cdots + n_{M_j jt}^{\frac{\eta+1}{\eta}} \right]^{\frac{\eta}{\eta+1}} \quad , \quad \eta > \theta$$

and maximization is subject to the household's budget constraint:

$$\mathbf{C}_t + \left[K_{t+1} - (1-\delta) K_t \right] = \int_0^1 \left[w_{1jt} n_{1jt} + \cdots + w_{M_j jt} n_{M_j jt} \right] dj + R_t K_t + \Pi_t, \tag{4}$$

$$\mathbf{C}_t = \int_0^1 \left[c_{1jt} + \cdots + c_{M_j jt} \right] dj. \tag{5}$$

Firm profits, Π_t, are rebated lump sum to the household. The function u is twice continuously differentiable with $u' > 0$, $u'' < 0$ and satisfies the Inada conditions. The consumption index captures perfect substitutability of consumption goods, such that our assumption of a single market price $P_t = 1$ is valid.[26]

Notation. Indexes, in bold font, are book-keeping devices and not directly observable in the raw data but can be constructed from observables. For example, the disutility of labor supply $\mathbf{N}_t$ does not corre-

[26] Observe that since we are solving the model with decreasing returns to scale in production, we are arbitrarily able to introduce monopolistic competition in the national market for goods. Let $\mathbf{C}_t = [\int \sum_{i \in j} c_{ijt}^{(\sigma-1)/\sigma} dj]^{\sigma/(\sigma-1)}$, then given household's optimal demand schedules, a firm would optimize a decreasing returns to scale *revenue function* as opposed to the decreasing returns to scale *production function* used here. Firms would charge identical, constant *markups*, and profits from product market power would be rebated to the household. To keep our analysis clean, we ignore this case.

spond to any aggregates reported by the Bureau of Labor Statistics. However, given parameters, $\mathbf{N}_t$ can be constructed from the universe of firm-level employment $\{n_{ijt}\}$. We denote aggregate labor computed by adding bodies as unbolded: $N_t = \sum_{ij} n_{ijt}$.

Elasticities. The elasticities of substitution at the firm and market levels, $\eta > 0$ and $\theta > 0$, jointly affect the labor market power of firms. Both across and within markets, the lower the degree of substitutability, the greater the market power of firms. Below we discuss a micro-foundation of the representative agent problem—presented in full in Appendix B—that exactly maps these parameters into the relative net costs to individuals of relocating between firms within the same market versus across markets.

Across-market substitutability θ stands in for mobility costs across markets, which are often estimated to be significant (Kennan and Walker, 2011). As such costs increase ($\theta \to 0$), the household minimizes labor disutility $\mathbf{N}_t$ by choosing an equal division of workers across markets: $\mathbf{N}_{jt} = \mathbf{N}_{j't}$, $\forall j, j' \in [0,1]$. This limit imparts the largest degree of local labor market power for firms as market employment is completely inelastic market by market and unresponsive to across-market wage differences. As substitutability approaches infinity, the representative household optimally sends all workers to the market with the highest wage, eroding market power of firms in other markets.

Within-market substitutability η stands in for within-market, across-firm mobility costs such as the job search process (Burdett and Mortensen, 1998), some degree of non-generality of accumulated human capital (Becker, 1962), or heterogeneity in worker-firm specific amenities or commuting costs. As such costs increase ($\eta \to 0$), the household minimizes within-market disutility $\mathbf{N}_{jt}$ by choosing an equal division of workers across firms: $n_{ijt} = n_{i'jt}$, $\forall i, i' \in \{1, 2, ...M_j\}$. This generates the largest degree of monopsony power for firms. Regardless of its wage, firm ij will employ the same number of workers, allowing it to pay less while maintaining its workforce. As substitutability increases, the representative household reallocates workers toward firms with higher wages. In the limit as η increases, the local labor market tends to perfect competition and within each market wages and marginal revenue products are equalized.

Labor supply. Given the distribution of wages $\{w_{ijt}\}$, the necessary conditions for household optimality consist of first order conditions for labor at each firm $\{n_{ijt}\}$. Combining these yields the following system of firm specific, upward sloping, labor supply curves:

$$n_{ijt} = \overline{\varphi} \left(\frac{w_{ijt}}{\mathbf{W}_{jt}} \right)^{\eta} \left(\frac{\mathbf{W}_{jt}}{\mathbf{W}_t} \right)^{\theta} \mathbf{W}_t^{\varphi} \quad , \quad \text{for all} \quad ij \tag{6}$$

The supply curve includes more book-keeping terms: the *market wage index* $\mathbf{W}_{jt}$ and *aggregate wage index* $\mathbf{W}_t$ are defined as the numbers that satisfy

$$\mathbf{W}_{jt}\mathbf{N}_{jt} := \sum_{i \in j} w_{ijt} n_{ijt} \quad , \quad \mathbf{W}_t \mathbf{N}_t := \int_0^1 \mathbf{W}_{jt}\mathbf{N}_{jt}\, dj.$$

Together with (6) these definitions imply the following indexes:

$$\mathbf{W}_{jt} = \left[\sum_{i \in j} w_{ijt}^{1+\eta} \right]^{\frac{1}{1+\eta}} \quad , \quad \mathbf{W}_t = \left[\int_0^1 \mathbf{W}_{jt}^{1+\theta} dj \right]^{\frac{1}{1+\theta}} .$$

Since we focus on Cournot competition, we work with the inverse labor supply function:

$$w_{ijt} = \overline{\varphi}^{-\frac{1}{\varphi}} \left(\frac{n_{ijt}}{\mathbf{N}_{jt}}\right)^{\frac{1}{\eta}} \left(\frac{\mathbf{N}_{jt}}{\mathbf{N}_t}\right)^{\frac{1}{\theta}} \mathbf{N}_t^{\frac{1}{\varphi}} \tag{7}$$

The remaining optimality conditions include the household Euler equation for consumption. We derive this and the labor supply system in detail in Appendix E.

Micro-foundation. What is our representative household representative of? In Appendix B we micro-found our preference specification. In the model presented, labor supply curves to firms are determined by a representative agent with nested-CES preferences. We show that the exact same supply system described by equations (6) and (2.2) can be obtained in an environment with heterogeneous workers making independent decisions.

The environment is as follows. Each worker decides which firm to work for and how many units of labor to supply. In making this decision, each worker minimizes the total disutility of attaining some random level of income. Total disutility is the sum of the logarithm of hours supplied and a worker specific disutility of supplying labor to each firm, ξ_{ij}. The worker specific disutility of supplying labor to each firm is *iid* across individuals and time, and drawn from a correlated Gumbel distribution in which θ governs the overall variance of ξ_{ij}, and η governs the within-market conditional correlation of ξ_{ij}.

Similar, non-nested, formulations of individual decisions have been used to model the total supply of labor to a firm in competitive markets by Card, Cardoso, Heining, and Kline (2018) and Borovickova and Shimer (2017). Our contribution is to adapt existing results in the discrete choice literature to demonstrate a supply-system equivalence with our 'nested-CES' representative household specification, and to set the problem in oligopsonistic markets.[27,28]

We also establish that, under constant returns to scale, the same supply system obtains in the steady-state of a *dynamic* discrete-choice setting in which workers are paid constant individual-firm specific wages. Workers separate from their firm with probability δ and when separating draw a new firm. Firms then compete in a dynamic oligopsony for these workers.

Beyond unifying alternative approaches, this micro-foundation delivers an intuitive interpretation of our key parameters θ and η.[29] In the discrete choice setting, increasing θ decreases workers' overall variance of net disutility ξ_{ij}. If θ is high, a worker has a high likelihood that their lowest draws of non-wage utility ξ_{ij} are close together, increasing overall competition on wages between firms. Increasing η increases the covariance of ξ_{ij} within markets. If $\eta > \theta$, then the smallest realizations of a worker's

[27] We adapt arguments from the product market case due to Verboven (1996). In that paper the author establishes the equivalence of nested-logit and nested-CES, extending the results of Anderson, De Palma, and Thisse (1987) who established an equivalence between single sector CES and single sector logit.

[28] We are primarily interested in discussing θ and η and within- and across- market preferences conditional on employment. To get to the overall representative household formulation requires (i) adding an additional nest to the logit that induces a decision between work and not-work, where the dispersion in idiosyncratic tastes over these choices maps into φ, (ii) complete markets such that the evolution of the aggregate capital stock and aggregate consumption due to the savings decisions of agents can be represented by the consumption and savings decisions of a representative household.

[29] This framework also clarifies the economics of the wage indexes $\mathbf{W}_t$ ($\mathbf{W}_{jt}$). These relate the *ex-ante expected* utility of one unit of labor supply in the economy (sector j)

disutilities are more likely to be *bunched* within a particular j, so facing similar non-pecuniary utility the worker closely compares wages within j. If $\eta \approx \theta$, then the smallest realizations of a worker's disutilities are more likely to be *spread* across markets, so the worker compares wages across j's. In the former case, a productive firm in market j is shielded from competing with the continuum of firms outside of its market. This directly maps the model into the proposed sources of monopsony power of Robinson (1933).[30]

An important feature of the model is that workers are not confined to particular markets. The micro-foundation makes clear that workers are able to move across markets. The limitation that markets impose is on the boundary of the strategic behavior of firms. Within markets firms are strategic, but with respect to firms in the continuum of other markets, firms are price takers. We now describe the behavior of the firm.

2.3 Firms

Firms draw idiosyncratic productivities z_{ijt} from a distribution $F(z)$. Within a market, we assume that M_j firms engage in Cournot or Bertrand competition. Firms take the aggregate wage $\mathbf{W}_t$ and labor supply $\mathbf{N}_t$ as given. In order to maximize profits, firms choose how much capital to rent, k_{ijt}, and either the number of workers to hire n_{ijt} (i.e. Cournot) or wages w_{ijt} to post (i.e. Bertrand). Our baseline assumes Cournot and Appendix E explores Bertrand.

The firm maximizes profits:

$$\pi_{ijt} = \max_{n_{ijt},k_{ijt}} \overline{Z} z_{ijt} \left(k_{ijt}^{1-\gamma} n_{ijt}^{\gamma}\right)^{\alpha} - R_t k_{ijt} - w_{ijt} n_{ijt} \quad , \quad \text{subject to (7).}$$

Given capital demand, we can rewrite firm profits using three auxiliary parameters:

$$\widetilde{\alpha} := \frac{\gamma\alpha}{1-(1-\gamma)\,\alpha} \quad , \quad \widetilde{z}_{ijt} := [1-(1-\gamma)\,\alpha]\left(\frac{(1-\gamma)\,\alpha}{R_t}\right)^{\frac{(1-\gamma)\alpha}{1-(1-\gamma)\alpha}} z_{ijt}^{\frac{1}{1-(1-\gamma)\alpha}} \quad , \quad \widetilde{Z} := \overline{Z}^{\frac{1}{1-\alpha(1-\gamma)}}$$

With this notation, the firm's labor demand problem can be expressed as follows:

$$\pi_{ijt} = \max_{n_{ijt}} \widetilde{Z}\widetilde{z}_{ijt} n_{ijt}^{\widetilde{\alpha}} - w_{ijt} n_{ijt} \quad , \quad \text{subject to (7).}$$

Define the marginal revenue product of labor: $MRPL_{ijt} = \widetilde{\alpha}\widetilde{Z}\widetilde{z}_{ijt} n_{ijt}^{\widetilde{\alpha}-1}$.[31] Then the first order conditions of this problem yields the solution that the wage is a *markdown* (μ_{ijt}) below the marginal revenue product of labor:

$$w_{ijt} = \mu_{ijt}\, MRPL_{ijt} \qquad , \qquad \mu_{ijt} \in (0,1).$$

[30] To quote in full: *"We have seen in what circumstances the supply of a factor to an industry may be less than perfectly elastic. The supply of labor to an individual firm might be limited ... there may be a certain number of workers in the immediate neighborhood and to attract workers from further afield it may be necessary to pay a way equal to what hey can earn at home plus their fares to and fro; or there may be workers attached to the firm by preference or custom... Or ignorance may prevent workers from moving from one firm to another."* In our micro-foundation of the CES supply structure the heterogeneous ξ_{ij} realizations across workers could reasonably be interpreted in any of these ways. A firm's marginal cost of labor curve lies above its supply curve because to hire more labor it must (i) pay more to hire a new worker away from another firm that workers have a low disutility of working at, (ii) must then pay this wage to all workers.

[31] Here we have abused description slightly since we are using a value-added production function and maximized out optimal capital, so this is really the marginal "revenue net of capital and intermediate input expense" product of labor.

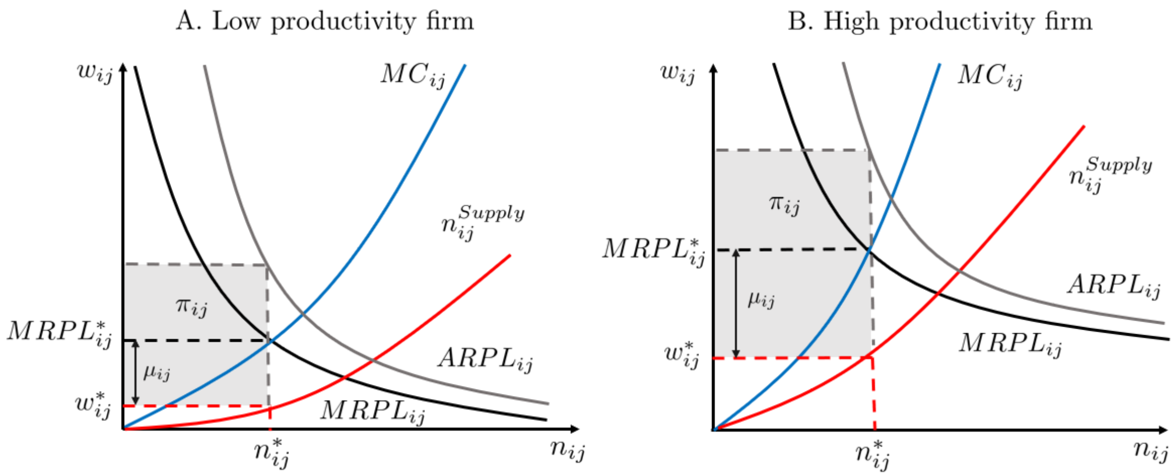

Figure 3: Firm level optimality

Figure 3 characterizes firm optimality. Decreasing returns to scale in production yields a downward sloping marginal revenue product of labor strictly below the average revenue product. Firms internalize the upward sloping labor supply curve, and so their marginal cost of labor is also upward sloping and lies strictly above labor supply (which is equivalent to the average cost of labor). A marginal unit of labor costs more than just the higher wage paid to the marginal worker, since the firm must increase wages paid to all workers. As such, choosing n_{ijt} so that labor's marginal revenue product equals its marginal cost necessarily implies a *markdown* of the wage relative to marginal revenue product. Firms earn profits from difference between average and marginal revenue products of labor due to decreasing returns to scale and labor market power in the form of the markdown.

In the Cournot Nash equilibrium, this markdown is determined by the equilibrium (inverse) labor supply elasticity faced by the firm $(1/\varepsilon_{ijt})$ *at the equilibrium allocation*. We refer to ε_{ijt} as the *structural elasticity of labor supply*. Computing ε_{ijt} requires us to *fix* competitors' labor demand, something which we will not be able to ask of the data under any configuration of shocks in the market. Using (7):

$$\frac{1}{\varepsilon_{ijt}} := \left.\frac{\partial \log w_{ijt}}{\partial \log n_{ijt}}\right|_{\overline{n}_{-ijt}} = \frac{1}{\eta} + \left(\frac{1}{\theta} - \frac{1}{\eta}\right) \left.\frac{\partial \log \mathbf{N}_{jt}}{\partial \log n_{ijt}}\right|_{\overline{n}_{-ijt}}$$

Conveniently, in the nested-CES case, the inverse labor supply elasticity is linear in the market payroll share of the firm, s_{ijt}^{wn}. Markdowns are therefore given by:[32]

$$\mu_{ijt} = \frac{\varepsilon\left(s_{ijt}^{wn}\right)}{\varepsilon\left(s_{ijt}^{wn}\right)+1} \quad , \quad \varepsilon\left(s_{ijt}^{wn}\right) := \left[\frac{1}{\eta}\left(1-s_{ijt}^{wn}\right) + \frac{1}{\theta}s_{ijt}^{wn}\right]^{-1} \quad , \quad s_{ijt}^{wn} := \frac{w_{ijt}n_{ijt}}{\sum_{k\in j} w_{kjt}n_{kjt}}. \tag{8}$$

These markdowns constitute our measure of *firm level labor market power* and depend on a firm's own (observable) market share as well as the degree of within-market (η) and across-market (θ) labor substi-

[32]Appendix E provides the derivations of these expressions.

tutability. In other words, markdowns vary by firm characteristics. This can easily seen by returning to Figure 3. Panel A describes the equilibrium outcomes for a low productivity firm. Relative to the high productivity firm in panel B, the low productivity firm has a lower $MRPL_{ij}$ for any n_{ij}. In equilibrium, it has both lower wages w_{ij}^*, and lower employment n_{ij}^*, so its share of wage payments s_{ij}^{wn*}, is smaller. With a smaller share of the labor market wage payments, its elasticity of labor supply is higher, and its inverse labor supply curve is flatter. A flatter inverse supply curve yields a narrower markdown at its optimal labor demand, n_{ij}^*. The larger firm faces an endogenously steeper supply curve and hires more workers at higher wages but a wider markdown.

Before proceeding, we note that in the tradeable goods market (the focus of our paper) labor market power, μ_{ijt}, is identified distinctly from product market power, and that this is robust. We make the simplest assumption: tradeable goods prices are competitive and so not set by any firm. If tradeable goods firms set output prices in a *monopolistically competitive national market*, invariant markups would enter the marginal *revenue* product, $MRPL_{ijt}$, and remain distinct from our markdown. If tradeable goods firms set output prices in a *oligopolistically competitive national market*, then this price setting decision will not affect our estimation of μ_{ijt} which exploits variation across labor markets within the same state. So long as tradeable good prices set by a firm (e.g. furniture prices) do not differ across local labor markets within a state, our estimate of μ_{ijt} will only capture labor market power.

2.4 Equilibrium

For the rest of this section we will focus on a steady state equilibrium and drop time subscripts. We later return to the time dimension when studying welfare counterfactuals under transitions between steady-states. The economy-wide vector of wage-bill shares, $\mathbf{s}^{wn} = \{\mathbf{s}_j^{wn}\}$ where $\mathbf{s}_j^{wn} = (s_{1j}^{wn}, \ldots, s_{M_j j}^{wn})$, is the only object that needs to be determined in a steady state equilibrium. This is key to our empirical strategy, since in Census data we will be able to measure exactly these shares.

A steady state equilibrium is a vector of wage-bill shares that yields wages and employment consistent with the vector of wage-bill shares. The steady state equilibrium interest rate is determined by the discount factor.

Definition. *A steady state equilibrium is a vector of wage-bill shares $\mathbf{s}^{wn}$ and an interest rate r, that are consistent with firm optimization, and that clear the labor market, capital market, and final good market.*

3 Characterization

We discuss the properties of the equilibrium in two steps. First, we describe the role of labor market power in determining employment and wages at the market level. Second, we describe the role of labor market power in determining employment and wages at the aggregate level.

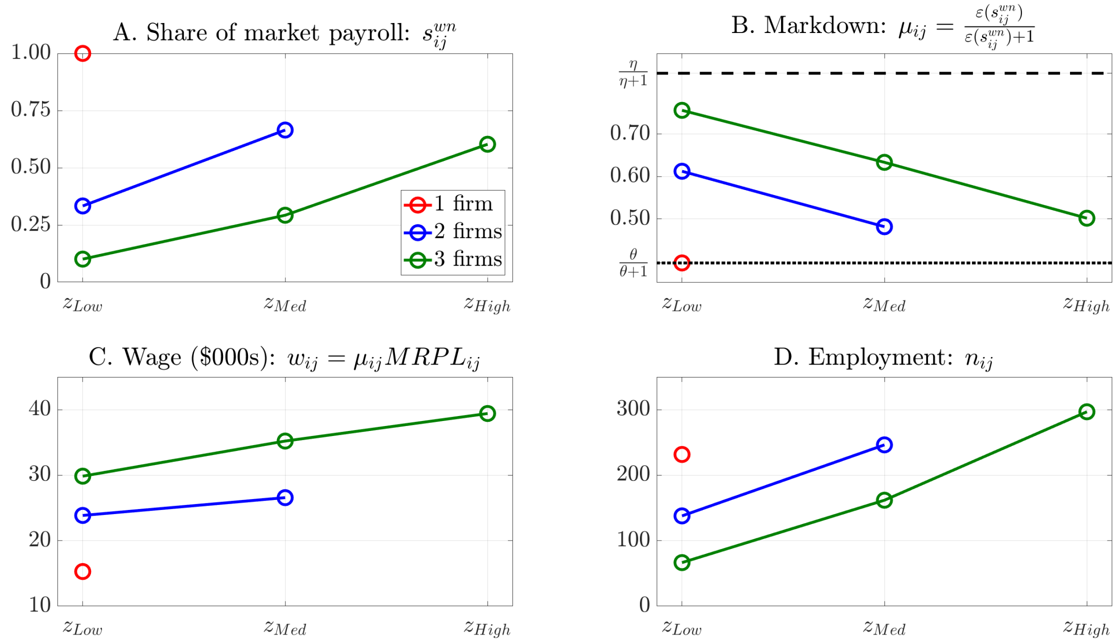

Figure 4: Oligopsonistic equilibrium in three labor markets

Notes: Figure constructed from model under estimated parameters (Table 3). Low, medium and high productivities of the firms correspond to the 10^{th}, 50^{th} and 90^{th} percentiles of the productivity distribution.

3.1 Market equilibrium

Lemma 3.1 summarizes the relationship between wage-bill shares, labor supply elasticities, and mark-downs. If $\mu_1 < \mu_2$ our convention will be to describe μ_1 as having a greater, or wider, mark-down.

Lemma 3.1. *Firms with larger market shares face smaller labor supply elasticities, and pay wages that represent larger mark-downs:*

$$\frac{\partial \varepsilon_{ij}}{\partial s_{ij}^{wn}} < 0 \qquad , \qquad \frac{\partial \mu_{ij}}{\partial s_{ij}^{wn}} < 0.$$

Under $\eta > \theta$—which we maintain here but are agnostic about in our empirics—large firms within a market face lower *equilibrium* labor supply elasticities (if $s_{ij} > s_{kj}$, then $\varepsilon_{ij} < \varepsilon_{kj}$). Single firm monopsonists face a labor supply elasticity of θ, whereas infinitesimally small firms face a labor supply elasticity of η. In Section 4 we will use quasi-natural experiments that shift $MRPL_{ij}$ to estimate how ε_{ij} varies by s_{ij} in the data and use this to infer η and θ.

To further explore how strategic interaction works in the model, Figure 4 plots examples of the equilibrium shares, markdowns, wages, and employment in three markets. The first market has a single low productivity firm (red), the second adds a firm with median productivity (blue), the third an additional high productivity firm (green).[33]

Consider the market with a single firm (red). By construction, the wage bill share is one (Panel A). Panel B shows that the markdown on the marginal product of labor is approximately 39 percent which

[33]Figure **??** is constructed from our benchmark calibration of the model (Section 4).

is equal to $\theta/(\theta+1)$ since the firm faces the lower bound on labor supply elasticities, θ (see Lemma E.2). Panel C shows that wages are low due to low productivity *and* a wide markdown. Despite this, the relatively inelastic labor supply across markets means the firm still employs many workers (panel D).

Consider the addition of a firm with higher productivity, a duopsony (blue). The low-productivity firm's wage bill share drops to around 30 percent and the firm with higher productivity employs the majority of the market, and market employment is higher. As its share falls, the low-productivity firm's markdown narrows to around 60 percent, as more competition increases their equilibrium labor supply elasticity toward η. Panel C shows that with no change to its productivity, but with a narrower markdown, the less productive firm's wage increases. Despite this wage increase, the higher wage at its new competitor bids away labor, causing the low productivity firm's employment to fall. Adding another firm (green), the markdown at the low- and mid-productivity firms decline. The largest firm has the widest markdown (Panel B), but pays more (Panel C) and employs more workers (Panel D).

In equilibrium, strategic interaction naturally occurs when there is local labor market power $(\eta > \theta)$ and finitely many firms. This leads to a negative covariance between markdowns and productivity—visible along the green line in Panel B. This will show up as a wedge in the aggregate conditions that we now turn to. This endogenous covariance is a direct consequences of labor market power that would be ignored in a model of monopsonistic competition.

3.2 General equilibrium

A key measure in the macroeconomic study of labor markets is the share of total output being paid to labor. In this section, we aggregate across markets to characterize the general equilibrium labor share.

We show that labor's share of income is a function of the distribution of market-level wage-bill Herfindahl indexes which we define as $HHI_{jt}^{wn} := \sum_{i\in j}(s_{ijt}^{wn})^2$. The model relevant aggregate measure of the extent of local labor market concentration is the inverse of the payroll weighted wage-bill Herfindahl:

$$\widetilde{IHI}_t^{wn} = \left[\int_0^1 s_{jt}^{wn} HHI_{jt}^{wn} dj\right]^{-1} \quad , \quad s_{jt}^{wn} = \frac{\sum_{i\in j} w_{ijt} n_{ijt}}{\int_0^1 \sum_{i\in j} w_{ijt} n_{ijt}\, dj} \quad ,$$

where s_{jt}^{wn} is market j's share of aggregate income.

Under Cournot competition, we can show that the labor share is determined by this statistic, intermediated by the key parameters of our model, η and θ:

$$LS_t = \underbrace{\alpha\gamma}_{\text{Comp. } LS} \frac{\widetilde{IHI}_t^{wn}}{\left(\frac{\eta+1}{\eta}\right)\widetilde{IHI}_t^{wn} + \left(\frac{\theta+1}{\theta} - \frac{\eta+1}{\eta}\right)} \tag{9}$$

The intuition for the reduction relative to the competitive labor share is as follows. A single firm's labor share is proportional to its equilibrium markdown. The market-level labor share LS_{jt} will put highest weight on firms that pay the largest share of wages in each market, which, in our model, are also firms with the widest markdowns and so lowest labor shares. Comparing two markets, a market with a higher HHI_{jt}^{wn} has more dispersed shares, a measurement that captures the fact that larger firms have both a wider markdown and a greater share of wage payments, leading to a lower market-level labor share. This delivers a closed-form relationship between LS_{jt} and HHI^{wn}_{jt}. That local labor shares LS_{jt} are then aggregated to the economy-wide labor share using payroll weights s^{wn}_{jt}, owes to simple accounting.

Under the assumption of stable preferences—and once θ and η are known—equation (9) implies that the dynamics of the distribution of local wage-bill Herfindahls is sufficient to forecast labor share dynamics. A contribution of this paper is to (i) identify θ and η, and (ii) measure this statistic in the same Census data. Lemma 3.2 shows these expressions have important implications for measurement.

Lemma 3.2.

(i) Under oligopsonistic competition ($\eta > \theta$) the labor share is an increasing function of the wage-bill weighted inverse Herfindahl index, $\frac{\partial LS}{\partial \widetilde{IHI}^{wn}} > 0$. Under monopsonistic competition ($\eta = \theta$), the labor share is independent of the wage-bill weighted inverse Herfindahl index.

(ii) Suppose $cov(w_{ij}, n_{ij}) > 0$, then the wage-bill Herfindahl is strictly larger than the employment Herfindahl, $HHI^{wn}_j \equiv \sum_{i\in j}(s^{wn}_{ij})^2 > HHI^{n}_j \equiv \sum_{i\in j}(s^{n}_{ij})^2$.

Part (i) implies that labor's share of income is determined by the *wage-bill Herfindahl*. Our theory rationalizes why the wage-bill Herfindahl can be used as a proxy for both local and national labor shares.

The model-implied measure of labor market concentration differs from most existing studies. For example, recent work by Benmelech, Bergman, and Kim (2018) and Rinz (2018) use employment Herfindahls and various weighting schemes. Independent of our model framework, employment Herfindahls understate concentration since they ignore the positive relationship between wages and employment, i.e. the positive size-wage premium. Part (ii) states this formally. So long as there is a size-wage premium—a robust feature of the data (Brown and Medoff, 1989; Lallemand, Plasman, and Rycx, 2007; Bloom, Guvenen, Smith, Song, and von Wachter, 2018)—Lemma 3.2 shows that the employment Herfindahl understates concentration relative to the wage-bill Herfindahl.[34]

4 Estimation

Our key parameters to estimate are the degree of across- (θ) and within- (η) market labor substitutability. We first describe our novel approach which integrates (i) new empirical estimates from a quasi-natural experiment in a large cross-section of firms and (ii) moments from Table 1 into (iii) a simulated method of moments routine in which all unknown parameters are estimated jointly.

4.1 Approach - Structural vs. reduced form labor supply elasticities

Structural elasticities. We motivate our approach from the following observation. If a researcher could observe the *structural elasticity of labor supply* that firms perceive *at* their Nash equilibrium level of employment, then they could combine these with data on payroll shares to exactly pin down (θ, η) by

[34]The unconditional firm-level correlation of log employment and log wages is 0.30 in our 2014 tradeable LBD sample.

inverting one of our key model equations:

$$\varepsilon\left(s_{ij}^{wn},\theta,\eta\right) := \left[\frac{\partial \log w_{ijt}}{\partial \log n_{ij}}\left(s_{ij}^{wn}\right)\Bigg|_{\overline{n}_{-ij}}\right]^{-1} = \left[\frac{1}{\eta}\left(1-s_{ij}^{wn}\right)+\frac{1}{\theta}s_{ij}^{wn}\right]^{-1}. \tag{10}$$

The *structural elasticity* of labor supply $\varepsilon(s_{ij}^{wn},\theta,\eta)$ depends only on observable shares and parameters.

Reduced form elasticities. When firms behave strategically, however, the structural elasticity cannot be measured using wage and employment responses even to well identified firm-level shocks. As is clear from the notation above, the structural elasticity is a strictly partial equilibrium concept and answers the question: *How much will firm i's wage have to increase in order to expand employment by 1 percent, holding its competitors' employment fixed?* This is the correct notion for the Nash equilibrium, however given a shock to any firm in the market, an employment change at firm i will lead competitors to best-respond, which will cause i to best respond and so on. What we measure in the data following a shock is therefore a *reduced form elasticity* $\epsilon(s_{ijt},\theta,\eta,\dots)$ that encodes all other firms' employment, wages, productivity, etc.[35]

Our key insight is that, despite this, the reduced form elasticities that we may aspire to measure in the data are still informative of (θ,η). The following expression provides a first order approximation of the reduced form elasticity of labor supply a researcher would measure for firm ij following a shock to it or any other firm(s) in the market:

$$\epsilon\left(s_{ijt}^{wn},\theta,\eta,\dots\right) = \frac{\mathrm{d}\log n_{ijt}}{\mathrm{d}\log w_{ijt}} = \frac{\varepsilon\left(s_{ijt}^{wn},\theta,\eta\right)}{1+\varepsilon\left(s_{ijt}^{wn},\theta,\eta\right)\left(\frac{1}{\theta}-\frac{1}{\eta}\right)\left(\frac{\sum_{k\neq i}s_{kjt}^{wn}d\log n_{kjt}}{d\log n_{ijt}}\right)}. \tag{11}$$

We unpack this below, but immediately note a distinct property of (11). The reduced form and structural elasticities coincide exactly under two special cases: (i) $\theta=\eta$, that is the labor market structure is one of national monopsonistic competition, (ii) $d\log n_{kjt}$ is zero at all other firms in the market. Case (ii) occurs when the shock hits firm i and firm i is either infinitesimal or the only firm in the market. Outside of these two special cases, strategic interaction and finiteness of firms jointly imply that micro-estimates of labor supply elasticities cannot be used to directly infer labor supply elasticities or underlying parameters.

Bias. To facilitate the relationship between structural and reduced form elasticities, we consider two cases: (i) a revenue productivity shock that hits one firm only, and (ii) a shock that hits a handful of firms in the market. The latter will be our empirical case.

If a shock hits only one firm and a researcher computes ϵ_{ij}, then this will *overstate* ε_{ij}. An increase in z_{ij} causes firm i to increase its employment and wages. An increase in n_{ij} when $\theta<\eta$ contracts labor supply at competing firms.[36] Competitors best responses move them back along their demand curves, *decreasing* their employment and increasing wages, such that $\sum_{k\neq i}s_{kj}^{wn}d\log n_{kjt}<0$. This best response from competitors shifts out labor *supply* for firm i, such that reaching its desired employment requires a

[35]We borrow the notation of ϵ for reduced form elasticities and ε for structural elasticities from the estimation of structural macroeconomic models. In this literature *reduced form shocks* which are empirical objects estimated out of VARs are often denoted ϵ, and *structural shocks* that are backed out of an estimated structural model are denoted ε.

[36]This from the point of view of Cournot, from the point of view of Bertrand other firms would see higher market wages.

smaller increase in wages. On net, wages at firm i increase *less* which—since they enter the denominator of the reduced form elasticity—lead to a *larger* estimate of $\epsilon_{ij} > \varepsilon_{ij}$. Since ϵ_{ij} overstates ε_{ij}, the researcher would conclude that the firm faces a *high* labor supply elasticity and markets are *more competitive* than they truly are.

Now consider a shock that hits a small firm and a large firm but misses medium sized firms. For the large firm the above holds, despite employment growth at the small firm, the *share-weighted* response of its competitors to the large firm's increase in employment is negative. For the small firm, the decrease in labor supply due to the expansion of the large firm will lead it to increase wages by more in order to grow, leading to a *smaller* estimate of $\epsilon_{ij} < \varepsilon_{ij}$. Since ϵ_{ij} understates ε_{ij}, the researcher would conclude that the small firm faces a *low* labor supply elasticity and markets are *less competitive* than they truly are.

Indirect inference. The above demonstrates that the full equilibrium structure of the model is necessary to take observed responses of wages and employment following identified shocks and map them into the underlying structural parameters. Our approach recognizes this. We first use a quasi-natural policy experiment to estimate the relationship between payroll shares and average *reduced form* labor supply elasticities:

$$\widehat{\epsilon}^{Model}\left(s,\theta,\eta\right) := \mathbb{E}\left[\epsilon^{Model}\left(s,\theta,\eta,\dots\right)\right].$$

The expectation being taken with respect to the distribution of all relevant labor market variables and shocks. We then replicate this experiment in our model and choose (θ,η)—along with other parameters—to replicate the empirical relationship between average reduced form elasticities and payroll shares. Formally, we proceed by indirect inference, where we minimize distance between average reduced forms estimates by share in the model and in the data: $\left|\widehat{\epsilon}^{Data}(s) - \widehat{\epsilon}^{Model}(s,\theta,\eta)\right|$. We now describe how we construct each of these terms.

4.2 Estimating reduced form labor supply elasticities in the data: $\widehat{\epsilon}^{Data}(s)$

Regression framework. To estimate $\widehat{\epsilon}^{Data}(s)$—the relationship between labor market payroll-shares and average reduced form labor supply elasticities—we compare how plants owned by the same firm, within the same state, but in different markets and with different market shares s_{ij}^{wn}, change wages and employment differently following a change in state corporate taxes.

We employ a regression framework. Let i denote the firm identifier (firmid), j industry (3-digit NAICS), s state, k commuting zone, and t year. Let y_{ijkt} denote the outcome of interest at the firm-i, market-jk, year-t level, such as employment or the wage. We control for payroll shares s_{ijkt}^{wn} and are interested in coefficients on state corporate taxes $\tau_{s(k)t}$ and their interaction with payroll shares.[37] To isolate the variation described above, we introduce fixed effects at the firm-state level $\alpha_{is(k)}$ as well as,

[37]State-level corporate taxes are proportional flat-taxes on firms' accounting profits. Our data for state-level corporate taxes comes from the data made publicly available by Giroud and Rauh (2019): (https://web.stanford.edu/ rauh/).

		A. Employment		B. Wages	
		$\log n_{ijkt}$	$\log n_{ijt}$	$\log w_{ijt}$	$\log w_{ijt}$
		(1)	(2)	(3)	(4)
State corporate tax	$\tau_{s(k)t}$	-0.00357	-0.00368	-0.00181	-0.00187
		(0.000644)	(0.000757)	(0.000584)	(0.000588)
Payroll share	s^{wn}_{ijkt}		2.085		0.214
			(0.0467)		(0.0072)
Interaction	$\tau_{s(k)t} \times s^{wn}_{ijkt}$		0.0158		0.0031
			(0.00495)		(0.00075)
R-squared		0.872	0.877	0.819	0.821
Observations		4,425,000	4,425,000	4,425,000	4,425,000

Table 2: Estimation results for equation (12)

Notes: All specifications include fixed effects for: (i) year, (ii) commuting zone, (iii) NAICS3 industry (iv) Firm×State. According to Census requirements, the number of observations is rounded to the nearest 1,000. Standard errors are clustered at State × Year level. Sample includes tradeable C-Corps from 2002 to 2012.

separately, industry and commuting zone.[38] Our regression specification is as follows:

$$\log n_{ijkt} = \alpha_{is(k)} + \delta_j + \phi_k + \mu_t + \psi s^{wn}_{ijkt} + \beta^n \tau_{s(k)t} + \gamma^n \left(\tau_{s(k)t} \times s^{wn}_{ijkt}\right) + \nu_{ijkt}. \tag{12}$$

The coefficients β and γ capture the average effect of state corporate tax rate changes and their differential effect by market share. We first estimate (12) separately for log employment and log wages (total payroll per worker). We then show how coefficient estimates from (12) can be used to construct $\widehat{\varepsilon}^{Data}(s)$.

Sample. To abstract from changes in product market power we restrict our sample to tradeable industries identified by Delgado, Bryden, and Zyontz (2014) and listed in Appendix C.[39] As in Section 1, a *market* is a 3-digit NAICS industry within a commuting zone. Plants owned by the same firm are aggregated within a market, such that an observation is a firm-market-year. As we rely on state-level corporate tax variation to isolate changes in labor demand, we restrict our sample to C-Corporation firms (C-Corps) in the LBD from 2002 to 2012.[40] Table A2 includes summary statistics of our 4.5 million observations at the firm-market-year level.

Estimates. Table 2 presents our empirical estimates of (12). We start with (log) employment as a dependent variable. Column (1) projects firm-market-year employment on corporate taxes $\tau_{s(k)t}$ alone. Since units of $\tau_{s(k)t}$ are percents, the coefficient on $\tau_{s(k)t}$ is an elasticity: a one percent increase in corporate taxes results in a 0.36 percent reduction in employment at the firm-market-year level. Column (3) presents the same specification for wages, which fall following an increase in taxes. Positively correlated employ-

[38] In this exercise only, we exclude commuting zones that straddle multiple states since defining a market gives rise to conceptual issues.

[39] See additional discussion in Section 1

[40] Firms are identified in Census by the variable *firmid*. The tax series ends in 2012.

ment and wage changes verify our interpretation of the tax change as shifting firm labor demand, which we formalize in the model below.

Columns (2) and (4) present the full set of interaction terms between payroll shares and corporate taxes. Both interaction terms are positive and significant, combined with negative direct effects of taxes, these indicate that the average responses of larger firms are smaller. Consider the mean effect of a 1 ppt increase in $\tau_{s(j)t}$ on a small firm at the mean payroll share of 0.03 and a large firm with a one standard deviation higher share of 0.14. Employment declines by -0.32 percent at the small firm and -0.15 percent at the large firm. Consistent with Giroud and Rauh (2019), corporate tax rate increases reduce employment, but this reduction is around half as large at larger firms.

Share-dependent reduced form labor supply elasticities. These regression results may be used to read off the relationship between the average reduced form labor supply elasticity and payroll shares which, as we previously argued, is informative regarding θ and η. Differentiate (12) with respect to $\tau_{s(j)t}$ to obtain share-dependent reduced form wage and employment elasticities:

$$\frac{\mathrm{d}\log n_{ijkt}}{\mathrm{d}\tau_{s(k)t}} = \beta^n + \gamma^n s_{ijkt}^{wn} \quad , \quad \frac{\mathrm{d}\log w_{ijkt}}{\mathrm{d}\tau_{s(k)t}} = \beta^w + \gamma^w s_{ijkt}^{wn}. \tag{13}$$

The ratio of the expressions in (13) yields the average reduced form labor supply elasticity by share:

$$\widehat{\epsilon}^{Data}(s) = \frac{\mathrm{d}\,\widehat{\log n_{ijkt}}}{\mathrm{d}\,\widehat{\log w_{ijkt}}} = \frac{\widehat{\beta}^n + \widehat{\gamma}^n s}{\widehat{\beta}^w + \widehat{\gamma}^w s} = \frac{-0.00368 + 0.0158 \times s}{-0.00187 + 0.0031 \times s}. \tag{14}$$

One might be concerned that this differs from the true expected reduced form labor supply elasticity because we divide the expected labor response by the expected wage response, however we compute the object in the model in the same way.

Figure 5 plots $\widehat{\epsilon}^{Data}(s)$ over $s_{ij}^{wn} \in [0, 0.15]$ which covers the bulk of our data. Nothing in our empirical exercise forces the relationship to be negative, but the blue dashed line is indeed downward sloping. The average reduced form labor supply elasticity at the smallest firms is around 2, while around 1 for the largest firms. Equipped with our estimate of the relationship between the average reduced form labor supply elasticity and payroll shares *for this particular quasi-experiment*, we now replicate the shock in the model and choose θ and η to minimize the distance between $\widehat{\epsilon}^{Model}(s,\theta,\eta)$ and $\widehat{\epsilon}^{Data}(s)$.

4.3 Simulating reduced form labor supply elasticities in the model: $\widehat{\epsilon}^{Model}(s,\theta,\eta)$

In order to construct $\widehat{\epsilon}^{Model}(s,\theta,\eta)$, we add corporate taxes to the environment and show how they shift marginal revenue products of labor.

We make several modifications to our theory. Corporate taxes are a tax on profits, net of interest payments on debt. The firm therefore maximizes post-tax accounting profits:

$$\pi_{ijt} = \Big(1-\tau_C\Big)\overline{Z}z_{ijt}\Big(k_{ijt}^{1-\gamma}n_{ijt}^{\gamma}\Big)^{\alpha} - \Big(1-\tau_C\lambda_K\Big)Rk_{ijt} - \Big(1-\tau_C\Big)w_{ijt}n_{ijt},$$

where we have assumed that all firms finance a fraction $\lambda_K \in [0,1]$ of their capital using debt. Only a

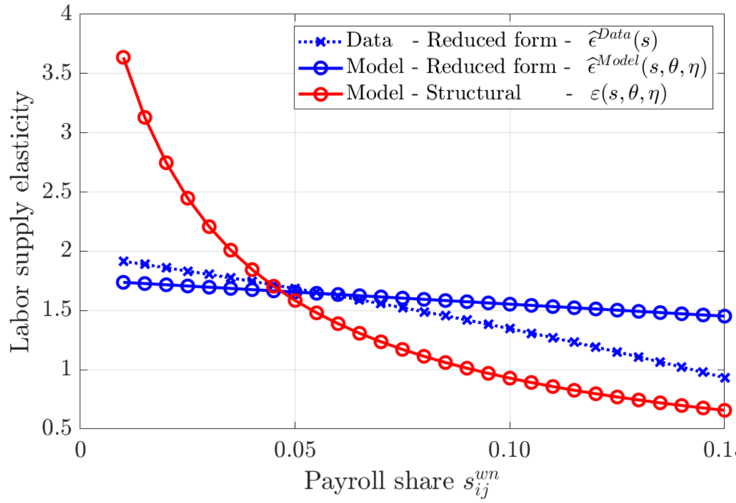

Figure 5: Labor supply elasticity by firm market wage share

Notes: This graph compares reduced form and structural labor supply elasticities by firm payroll share in response to a corporate tax shock of 1 percentage point. The line labeled 'Data - Reduced form' is the labor supply elasticity given by equation (14). The line labeled 'Model - Reduced form' plots reduced form labor supply elasticity estimates, estimated on simulated model data as described in Appendix H.2. The line labeled 'Model - Structural' plots $\varepsilon(\cdot)$ from equation (10).

random fraction $\omega_C \in [0,1]$ of firms in each market are C-corps and so subject to τ_C. For all other firms $\tau_C = 0$. To capture average size differences we assume C-corps are Δ_C more productive:

$$\log(z_{ijt}) \sim \begin{cases} N\left(1, \sigma_z^2\right) & \text{if } i \text{ is not a } C\text{-corp (i.e. w.prob } 1-\omega_C) \\ N\left(1+\log(\Delta_C), \sigma_z^2\right) & \text{if } i \text{ is a } C\text{-corp (i.e. w.prob } \omega_C) \end{cases}$$

To simplify notation, we substitute in the firm's optimal capital choice—which is non-strategic—and write the employment decision of the firm as:

$$\widetilde{\pi}_{ijt} = \max_{n_{ijt}} \widetilde{Z}\widetilde{z}_{ijt} n_{ijt}^{\widetilde{\alpha}} - w_{ijt} n_{ijt} \quad , \quad \text{subject to (7)}$$

where we make the following modifications to our prior notation,

$$\widetilde{\pi}_{ijt} = \frac{\pi_{ijt}}{1-\tau_C} \,,\ \widetilde{\alpha} := \frac{\alpha\gamma}{1-\alpha(1-\gamma)} \,,\ \widetilde{z}_{ijt} = [1-\alpha(1-\gamma)] \left[\frac{\alpha(1-\gamma)(1-\tau_C)}{(1-\tau_C\lambda_K)R}\right]^{\frac{\alpha(1-\gamma)}{1-\alpha(1-\gamma)}} z_{ijt}^{\frac{1}{1-\alpha(1-\gamma)}} \,,\ \widetilde{Z} = \overline{Z}^{\frac{1}{1-\alpha(1-\gamma)}}$$

For C-corps, an increase in τ_C reduces the marginal revenue product of labor via its distortion of the firm's capital decision.

We can now compute the model-based average reduced form labor supply elasticity by payroll share: $\widehat{\epsilon}^{Model}(s,\theta,\eta)$. Given parameters $\{\omega_C, \lambda_K, \tau_C\}$, we solve the initial equilibrium. For the ω_C fraction of firms that are C-corps, we increase τ_C by $\Delta\tau_C$. We solve the model again and then treat our model panel like our Census data: we run regression (12) and apply (14) to compute $\widehat{\epsilon}^{Model}(s,\theta,\eta)$.

4.4 Indirect inference

Average reduced form labor supply elasticities are not available in closed form, so we proceed by indirect inference implemented as simulated method of moments (SMM).

Externally calibrated. To capture the distribution of firms across markets in 2014, $M_j \sim G(M_j)$, we use a mixture of two Pareto distributions. We fit this mixture to the first three moments of the distribution, given in Table H1. Appendix H provides additional details, including parameter estimates. By construction we generate the correct fraction of markets with one firm. Throughout we simulate $5,000$ markets and verify that our results are not sensitive to this choice.

We proxy the fraction of capital that is debt financed by the debt to capital ratio defined by Graham, Leary, and Roberts (2015). We compute this in Compustat, restricting our sample to tradeable industries in 2014 and obtain $\lambda_K = 0.309$. We assume that $\omega_C = 0.428$ of firms are C-corps based on the 2014 County Business Patterns data for our tradeable sectors (CBP).[41]

In each iteration of the simulated method of moments, we compute two steady-state model economies. One with τ_C set to the mean state corporate tax rate of 7.14 percent, and a second with higher taxes $\tau_C' = \tau_C + \Delta_\tau$. We assume Δ_τ is 1 percent, approximately one standard deviation of the distribution of state corporate tax changes observed in our data (e.g. see Table A2 in Appendix C)). We compute these from the data made publicly available by Giroud and Rauh (2019). We then treat the outcomes across these two model economies as panel data and estimate regression (12) with firm fixed effects. Replicating our treatment of the data, we transform these point estimates into average reduced form elasticities by local labor market payroll share using equations (13, 14). Appendix H.2 includes additional details on the simulated tax experiment.

We assume a baseline aggregate Frisch elasticity of labor supply is $\varphi = 0.50$ which lies in the range of estimates obtained in micro-data analyses (Keane and Rogerson, 2012).[42] Our main results also consider $\varphi \in \{0.2, 0.8\}$. On an annual basis, the discount rate is 4 percent $(\beta = 0.9615)$, and the depreciation rate is 10 percent $(\delta = 0.10)$.

Internally estimated. We now estimate $\boldsymbol{\psi} = \{\theta, \eta, \gamma, \alpha, \sigma_{\tilde{z}}, \Delta_C, \overline{Z}, \overline{\varphi}\}$. In any solution of the model $(\overline{Z}, \overline{\varphi})$ can be chosen to match average firm size (27.96) and payroll per worker $(\$65,773)$ (Table 1).[43]

[41] We were unable to release this share from the Census. A higher value of ω_C is conservative in that the shock is more, rather than less, at the market level. We therefore choose an upper bound for ω_C. The fraction of establishments in the U.S. that are C-corps represents such an upper bound in four steps. (i) The CBP has only establishment data. (ii) On average, C-corps have more establishments than other firms. (iii) Our approach throughout the paper has been to aggregate firms' establishments within a market. (iv) Therefore the share of *establishments* that are C-corps in the data is an *upper bound* on the share of our 'market-firms' that are C-corps.

[42] The U.S. Congressional Budget Office uses estimates between 0.27 and 0.53. Reichling and Whalen (2012) discuss.

[43] We provide the closed-form mapping in Appendix F.1. As an alternative approach, we have considered inverting the model's equilibrium conditions to recover the distribution of productivities. Appendix Section F.2 provides the details. Given θ and η, the vector of wage-bill shares s_{ijt}^{wn} within a market determines mark-downs, and ratios of markdowns deliver ratios of productivities (equation F6). Up to a normalization of productivity at one firm, the distribution of relative productivities can be obtained. Absolute productivities can then be obtained by integrating employment data. In practice, however, the logistics of working with U.S. Census data render this approach infeasible.

Parameter	Description	Value	Moment	Model	Data
A. Assigned					
r	Risk free rate	0.04			
δ	Depreciation rate	0.10			
φ	Aggregate Frisch elasticity	0.50			
J	Number of markets	5,000			
$G(M_j)$	Mix two paretos		Mean, Std. Dev., Skewness of distribution		
			15 percent of markets have only 1 firm		
ω_C	Share of firms that are C-corps	0.43	Share of estabs. that are C-corps (CBP, 2014)		
τ_C	State corporate tax rate	0.071	Mean of state corp. tax rate $\tau_{C,st}$ (Giroud Rauh, 2019)		
Δ_τ	State corporate tax rate increase	0.010	Std. dev. of annual $\tau_{C,st}$ (Giroud Rauh, 2019)		
λ_K	Fraction of capital debt financed	0.309	Tradeable industries (Compustat, 2014)		
B. Estimated					
η	Within market substitutability	5.38	Average $\hat{\epsilon}^{Data}(s^{wn})$ for $s^{wn} \in [0, 0.05]$	1.70	1.80
θ	Across market substitutability	0.66	Average $\hat{\epsilon}^{Data}(s^{wn})$ for $s^{wn} \in [0.05, 0.10]$	1.60	1.52
Δ_C	Relative productivity of C-corps	1.29	Emp. share of C-corps	0.65	0.66
$\sigma_{\tilde{z}}$	Productivity dispersion	0.227	Payroll weighted $\mathbb{E}[HHI^{wn}]$	0.14	0.14
α	DRS parameter	0.985	Labor share	0.58	0.57
γ	Labor exponent	0.811	Capital share	0.18	0.18
$\widetilde{Z}$	Productivity shifter	2.10e+04	Ave. firm size	28.0	28.0
$\overline{\varphi}$	Labor disutility shifter	5.383	Ave. payroll per worker ($000)	65.8	65.8

Table 3: Summary of Parameters

Notes: See Table H1 and Figure H1 for model and data moments of the across-market firm distribution $G(M_j)$

We estimate $\theta = 0.66$ and $\eta = 5.38$ so that average reduced form labor supply elasticities by payroll share in the model coincide with our data estimates (e.g. Figure 5). Rather than targeting the entire function (13), we compute the average reduced form labor supply elasticity of firms with payroll shares between 0 and 5 percent, and 5 and 10 percent. This captures the bulk of variation in our data, spanning roughly ± 1 standard deviation around the mean wage-bill share. The reduced form elasticity of small firms is most informative of η whereas the reduced form elasticity of the larger firms contains more information about θ.

We estimate productivity dispersion $\sigma_{\tilde{z}}$ to match the payroll weighted wage-bill Herfindahl of 0.14 (Table 1). Increasing $\sigma_{\tilde{z}}$ increases the market power of large firms, increasing concentration.

We estimate α and γ to match the capital share and labor share of income. As can be seen from equation (9), α shifts the labor share, which is 57 percent in the data (Giandrea and Sprague, 2017). We then choose γ to match the aggregate capital share, which we take to be 18 percent (Barkai, 2016). Our estimate of α implies a small degree of decreasing returns to scale.

Finally, we choose Δ_C to match the 66 percent employment share of C-corps in tradeables, measured exactly from the CBP. Table 3 summarizes all parameters and the model's fit to the target moments.

4.5 Discussion

We quantify the importance of strategic complementarities in our inference by computing the bias associated with assuming that reduced form labor supply elasticities directly reveal structural labor supply elasticities and therefore the key underlying parameters. We also discuss caveats related to the labor demand shifter used here: state corporate taxes.

Structural v. reduced form labor supply elasticities. Using data on employment and wage changes in response to identified firm-level shocks to infer key structural parameters creates sizeable bias. This shows that strategic complementarities are a quantitatively important feature of labor markets.

Consider a truly idiosyncratic shock to a single randomly selected firm in our economy. Drawing the treated firm at random, compute $\widehat{\varepsilon}_{ij}$ and compare it to ε_{ij}. We repeat this 5,000 times for small (one percent), medium (20 percent) and large (50 percent) productivity shocks. We plot the results in Figure 6. This Monte Carlo exercise reveals a significant difference in reduced form and structural labor supply elasticities for firms with market shares not equal to 0 or 1, even when the identifying variation is *perfectly idiosyncratic*. The bias between reduced form and structural elasticities is nearly 15 percent even for large firms with market shares nearing 10 percent. In fact, for relatively small firms with market shares between 1 and 2 percent, which is below the average in our sample, the bias exceeds 7 percent for any size shock we consider. Accounting for the Nash equilibrium of the local labor market is quantitatively important for recovering fundamental parameters of the model.

This exercise implies that even if a researcher aims to use perfect idiosyncratic variation in productivity to infer structural elasticities and do welfare analysis, they would have to deflate their *reduced form elasticity* estimates substantially in order to recover the true *structural elasticities*. Inferring structural elasticities that are too large one would infer narrower markdowns which would bias *downward* the welfare costs of labor market power. The details of our Monte Carlo exercise are included in Appendix H.3.

Figure 6 shows that two important caveats apply, both summarized in equation (11). If the firm has a share of one, then reduced form and structural elasticity estimates coincide and reveal θ. If the firm has an infinitesimal share, then reduced form and structural elasticity estimates coincide and reveal η. Finally, a *market level* shock will directly reveal θ, so long as the market itself is not large. If the market is very large then a market level shock will also effect the macroeconomic equilibrium of the labor market, and reduced form elasticities will be contaminated by φ.

Corporate taxes. Three aspects of our corporate tax rate estimation are worth discussing: (i) apportionment of state taxes across multi-state production units may mean that state corporate taxes do not affect firms within a state, (ii) anticipation of tax changes, (iii) if corporate taxes affect *all* firms in a region then they can only be used to identify θ. To address the first two issues, we rely on prior analysis by Suárez Serrato and Zidar (2016) and Giroud and Rauh (2019). Our baseline estimation addresses the third issue directly.

First, Suárez Serrato and Zidar (2016) show that the impact of state corporate taxes on local economic activity is extremely similar for both (i) the statutory corporate taxes used here, and (ii) effective corpo-

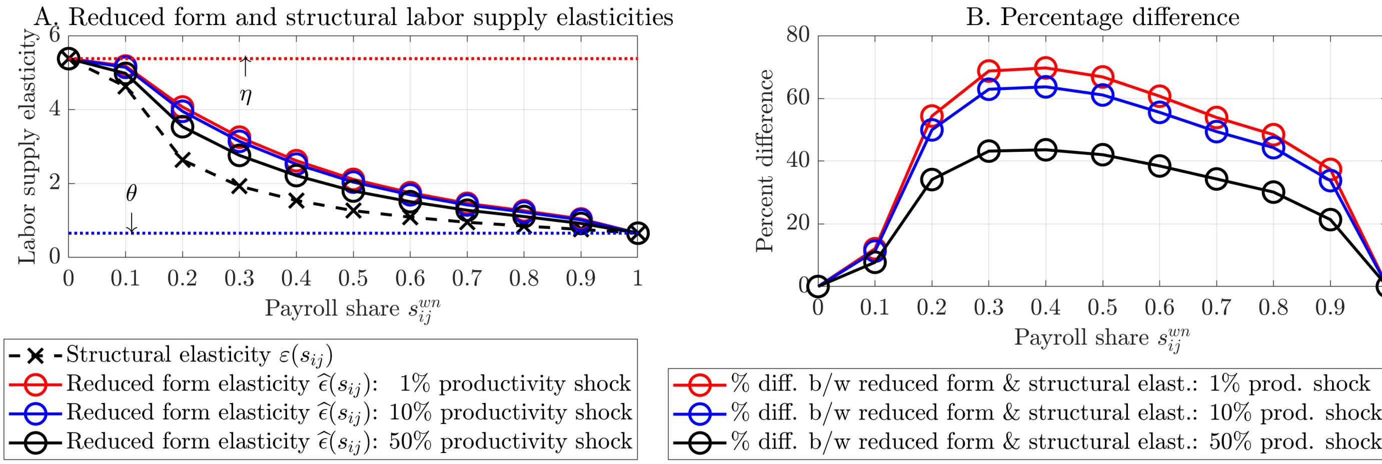

Figure 6: Reduced form and structural elasticities in response to idiosyncratic productivity shocks.

Notes: Panel A plots Monte Carlo results which compare reduced form to structural labor supply elasticities in response to a perfectly idiosyncratic shock to a single firm. The lines labeled 'Reduced form elasticity' plot the average estimated reduced form labor supply elasticity $\widehat{\varepsilon}(s)$ as detailed in Appendix H.3. The dashed line labeled 'Structural elasticity' plots $\varepsilon(s)$ from equation (10). Panel B reports the error of the average reduced form elasticity relative to the structural elasticity: $100 \times (\widehat{\varepsilon}(s) - \varepsilon(s))/\varepsilon(s)$.

rate taxes—i.e. 'business taxes'—carefully adjusted for apportionment weights.[44] Since establishment sales and company property values are not available to us, we cannot construct accurate apportionment weights and thus we focus on statutory tax rates compiled by Giroud and Rauh (2019).

Second, both Suárez Serrato and Zidar (2016) and Giroud and Rauh (2019) establish that including other aspects of changes to fiscal policy around corporate tax adjustments have negligible affects on their measured elasticities of local economic activity to state corporate taxes.[45] We interpret this as indirect evidence that the reforms are not paired with other predictable components of fiscal stimulus, such as unemployment insurance, which follow time-invariant threshold rules and are typically triggered in recessions (e.g. Mitman and Rabinovich (2019)).

Third, only a fraction of firms pay statutory corporate tax rates. According to the County Business Patterns database, the employment share of C-corps is 66 percent. Our model estimation took this into account. We chose an upper bound for the fraction of C-corp firms in our economy, and matched the aggregate share of employment in C-corps. The fact that C-corps are not *all* firms is in fact key to our approach of using one shock and its effect across the distribution of payroll shares to estimate *both* within- and across- market labor substitutability.

[44]See their discussion of Table A21, p.19 (emphasis added): "Column (6) of Table 5 and Appendix Table A21 show that using statutory state corporate tax rates in Equation 21 (instead of business tax rates τ_b) results in similar and significant estimates, indicating that *our measure of business tax rates is not crucial for the results*."

[45]Giroud and Rauh (2019) establish the plausible exogeneity of state-corporate tax changes. From a public finance perspective they study the effects of state corporate tax changes on employment and wages. Their focus is *within firm, across state* responses, and the reallocation of firm employment across states following tax changes. For an exhaustive description of these tax changes we point the interested reader to their paper.

5 Validation

Before conducting our key measurement exercises, we carry out two over-identifying tests of the model. We compare our model to (i) pass-through rates which are often used to infer and measure monopsony power, and (ii) the cross-market distribution of concentration, both of which have been at the forefront of the "labor market power" discussion. We argue that these tests situate the model well for our measurement exercises and future research on labor markets.

Since it does not fit neatly below, we note here that in terms of the size-wage relationship in the model and data, we measure an unconditional 0.30 (0.56) correlation between log employment and log wages across firms in our 2014 sample (the model).

5.1 Pass-through

We check that the model produces rates of pass-through from value added per worker to wages consistent with recent empirical estimates (Kline, Petkova, Williams, and Zidar, 2019; Card, Cardoso, Heining, and Kline, 2018). With strategic complementarities, increases in firm productivity lead to more hiring and higher wages, but with an expanding market share the firm's markdown widens, which dampens the wage increase. As a result, with strategic complementarities, pass through from value added per worker to wages is less than 1, consistent with (Kline, Petkova, Williams, and Zidar, 2019; Card, Cardoso, Heining, and Kline, 2018). Standard models without variable markdowns, such as neoclassical models of monopsony (e.g. Manning (2003) and Card, Cardoso, Heining, and Kline (2018)), imply pass through from value added per worker to wages of 1.

Background. Throughout this section, we define *pass-through* to be the elasticity of wages with respect to value added per worker (VAPW) following a productivity shock. This metric conforms with the literature. Since productivity is often not observed, most empirical studies compare how some identified shock affects wages relative to sales- or value added.[46]

While many papers compute pass-through estimates, we focus on Kline, Petkova, Williams, and Zidar (2019) (henceforth KPWZ). KPWZ provides sufficient summary statistics to replicate their quasi-experiment in our model. The quasi-experiment exploits patent issuance as an instrument, comparing consequent changes in value added per worker (VAPW) and wages. They find that the receipt of a high-value patent increases VAPW by approximately 13 percent, and for every one percent increase in VAPW, wages increase by approximately 0.47 percent.

Replication. We replicate KPWZ as follows. Given the baseline equilibrium, we randomly sample one percent of firms. We match the large average size of the regression sample in KPWZ by drawing from firms with employment greater than $\underline{n}$. We increase the productivity of treated firms by a factor Δ. The values of $\underline{n}$ and Δ are calibrated to match the KPWZ (1) median firm size of 25 employees, (2)

[46] Wages are measured either as labor compensation per worker or as an hourly wage.

increase in post-tax value added per worker of 13 percent.[47] To replicate the partial equilibrium nature of the experiment, we keep aggregates fixed and solve the new market equilibrium. Table A3 compares summary statistics of our regression sample to theirs.

For our estimate of pass-through, we conduct an apples-to-apples comparison adopting the procedure of KPWZ.[48] We treat the untreated and treated observations for each firm as a panel with two observations per firm of wages $\{w_{ij0}, w_{ij1}\}$ and value added per worker, $\left\{\frac{y_{ij0}}{n_{ij0}}, \frac{y_{ij1}}{n_{ij1}}\right\}$.[49] We then regress the wages in levels on VAPW in levels and a firm-specific fixed effect. The regression coefficient is converted into an elasticity using untreated mean wages and mean value added per worker (see their Section 7).

Results. Table 4 reports our estimates. We find a pass-through rate of 79.5 percent, which is larger than the KPWZ estimate of 47.0 percent (Table VIII, panel B, column 2) but significantly less than the 100 percent pass-through rate that standard models of monopsony would imply. Recent work by Card, Cardoso, Heining, and Kline (2018) uses lagged log sales per worker as an instrument for log value added per worker. We have insufficient information to replicate their empirics, but report their pass-through estimate of 32.7 percent.[50] Finally, a structural approach is taken by Friedrich, Laun, Meghir, and Pistaferri (2019), who estimate pass-through of 31 percent from permanent shocks in a model of worker and firm dynamics estimated on Swedish employer-employee data.[51]

We compute the same statistic in two model economies that lack strategic complementarities: (1) the competitive version of our model, defined below (section 6.2), and (2) the *monopsonistically competitive* version of our model in which $\theta = \eta$. From the perspective of pass-through, both market structures are counterfactual, producing pass-through one one in both cases.

Importance of strategic complementarities. Absent endogenous markdown responses to productivity shocks, our model would not match these facts. A large class of monopsony models that abstract from variable markdowns would predict an elasticity of wages with respect to value added per worker of 1 (e.g. Manning, 2003; Card, Cardoso, Heining, and Kline, 2018). Our model nests such models under $\eta = \theta$, in which case firms face a labor supply curve $w_{ijt} \propto n_{ijt}^{1/\eta}$. There are no strategic complementarities between wages within a market. Markdowns are therefore constant and common to all firms: $\mu_{ijt} = \eta/(\eta+1)$. Value added is $va_{ijt} \propto \widetilde{Z}\widetilde{z}_{ijt} n_{ijt}^{\widetilde{\alpha}}$. The wage is therefore proportional to value added per worker. Following any shock, the measured elasticity of wages to value added per worker is one:

$$\frac{va_{ijt}}{n_{ijt}} = \frac{1}{1-\alpha(1-\gamma)} \frac{\widetilde{Z}\widetilde{z}_{ijt} n_{ijt}^{\widetilde{\alpha}}}{n_{ijt}} \quad , \quad w_{ijt} = \frac{\eta}{\eta+1} \widetilde{\alpha}\widetilde{Z}\widetilde{z}_{ijt} n_{ijt}^{\widetilde{\alpha}-1} \quad , \quad \frac{\Delta \log(w_{ijt})}{\Delta \log(va_{ijt}/n_{ijt})} = 1$$

This simple result establishes that a large class of models without strategic complementarities yield pass-through rates of one, which a growing body of evidence from reduced form and structural studies

[47] See KPWZ. We take the *Median firm size* of 25.36 from their Table II, panel A, column 7. The percentage increase in VAPW is 0.13=15.74/120.16, where 15.74 is the mean increase in value added per worker (Table V, column 4), and 120.16 is the mean value added per worker (Table II, panel A, column 5). This is exactly equal to our value-added production function which represents sales minus costs of intermediate inputs.

[48] They describe this procedure in Section VII, and footnotes to Table VIII.

[49] *Value-added* in KPWZ is defined as sales minus costs of goods sold net of labor costs.

[50] See their Table 2, panel A, row IV, column 1.

[51] See their Table 12, column 1.

	(1) Oligopoly	(2) Kline et al (2018)	(3) Card et al (2018)	(4) Competitive	(5) $\theta = \eta$
Pass-through coefficient	0.795	0.470	0.327	1.000	1.000
Dependent Variable	w_{ij}	Wage bill per worker	Hourly wage	w_{ij}	w_{ij}
Independent Variable	$VAPW_{ij}$	$VAPW_{ij}$ (IV: Patent approvals)	$VAPW_{ij}$ (IV: Sales per worker)	$VAPW_{ij}$	$VAPW_{ij}$

Table 4: Wage pass-through, model versus U.S. data

Notes: Description of estimation procedure in text. See Table A3 for a comparison of the model-based summary statistics for the simulated KPWZ experiment and the summary statistics reported in KPZW.

rejects.[52] In summary, we view the model's ability to match this moment as (i) validation of our model's transmission mechanism from productivity to wages, and (ii) evidence that strategic complementarities are important for understanding wage setting.

5.2 Distribution of labor market concentration

In calibrating the model we matched the payroll weighted average of payroll concentration across markets, and by construction the 15% of markets with only one firm. In those markets with one firm, we trivially match payroll and firm concentration. Our second exercise establishes that the model matches the non-targeted spatial variation in concentration.

Data. Figure 7 compares the weighted and unweighted distributions of wage-bill Herfindahl indexes in the model and new data that we have extracted from the Census.[53] Panel A plots the payroll weighted distribution of the wage-bill Herfindahl. Very little of *aggregate wage payments* are accounted for by concentrated markets with a payroll HHI more than 0.25. On the other hand, by plotting the *unweighted* distribution in Panel B, we see that the majority of *markets* are highly concentrated. The weighted mean HHI is 0.14, while the unweighted mean is more than three times larger (0.45).

The stark difference is due to the negative correlation between concentration and total payroll / employment. While 15 percent of markets have one firm, those markets comprise less than half of one percent of total employment. To show that this large difference between the weighted and unweighted Herfindahl distributions is not specific to the tradeable sector, Figure 7 also includes the economy-wide distribution.

Model. The model fits both distributions, and while it matches the weighted average HHI_j^{wn} exactly by construction it also generates an unweighted average payroll Herfindahl of 0.33, which is quite close to the data counterpart. In the model, markets with one firm have high labor market power, restricting quantity (lower employment) and widening markdowns (lowering wages). The model correlation between market size and market concentration is -0.80, whereas in the data the correlation is -0.21.

[52] Note that pass-through in *levels* is less than one, but is not direct evidence of *monopsony* since it conflates the elasticity of labor supply and the output elasticity of labor in value-added.

[53] Table A1 in Appendix C provides point estimates referenced in this section and additional statistics.

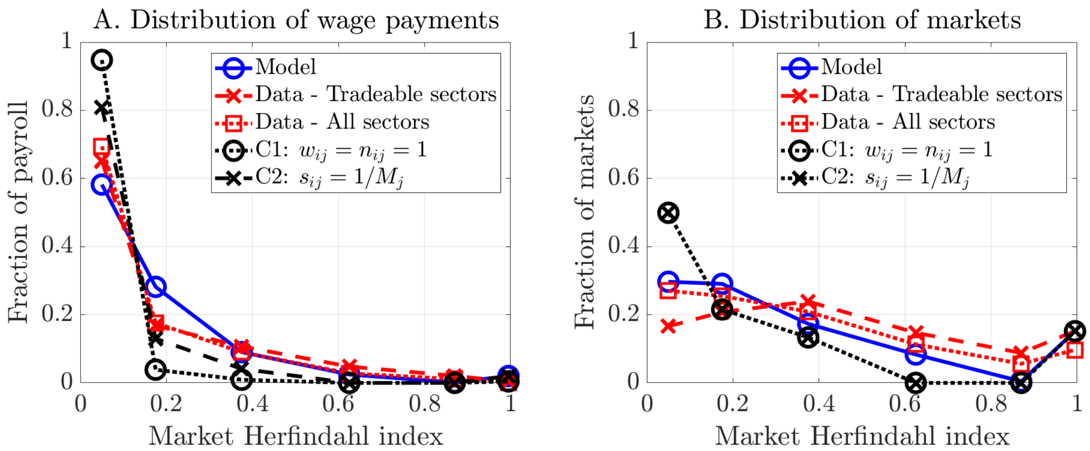

Figure 7: Cross-market distribution of concentration: Model and Data (U.S. Census), 2014.

Notes: Figure plots the across market distribution of the payroll Herindahl index (HHI_j^{wn}). Bins are determined by the following bounds: $\{0, 0.10, 0.25, 0.50, 0.75, 0.99, 1\}$. The horizontal axis gives the center of each bin. Panel A plots the fraction of total payroll in each bin. Panel B plots the fraction of markets in each bin. The former informs the *payroll weighted* index, the latter informs the *unweighted* index referenced in the text. Data is Census LBD. See Appendix C for additional details. Table A4 provides the exact weighted and unweighted means of employment and payroll HHIs. Counterfactuals are described in the text.

Counterfactuals. The model's endogenous firm size heterogeneity is important for matching the data. Two counterfactual distributions are plotted. Under *C1* we arbitrarily set all wages and employment equal to one, so that the only object in the model is $G(M_j)$ and each market's HHI_j^{wn} is simply $(1/M_j)$, and its payroll share is proportional to M_j. Under *C2* we keep this counterfactual distribution of HHI_j^{wn} but construct the weighted distribution using payroll shares from the baseline model $s_{j,baseline}^{wn}$:

$$\text{C1:} \quad \left(HHI_j^{wn}, s_j^{wn}\right) = \left(\frac{1}{M_j}, \frac{M_j}{\sum_j M_j}\right) \quad , \quad \text{C2:} \quad \left(HHI_j^{wn}, s_j^{wn}\right) = \left(\frac{1}{M_j}, s_{j,baseline}^{wn}\right)$$

Relative to *C*1, model and data display a smaller fraction of payroll in competitive markets ($HHI_j^{wn} < 0.10$). With symmetric firms and the benchmark distribution of wage payments across markets, *C*2 also misses the data substantially. We conclude that within market heterogeneity in employment and payroll is key to matching the data.

6 Measurement

We measure labor market power in three exercises. First, at the micro-level we compute the distribution of markdowns and labor supply elasticities in the economy. Second, to compute the macroeconomic affects of labor market power, we compare our benchmark oligopoly model economy to a counterfactual competitive economy. Third, to measure the time-series impact of labor market power on the labor share we combine our closed form expression with our estimates of (θ, η) and data on concentration over time. With all deep parameters of the model estimated we remove all features of the model related to corporate taxes that were introduced for estimation.

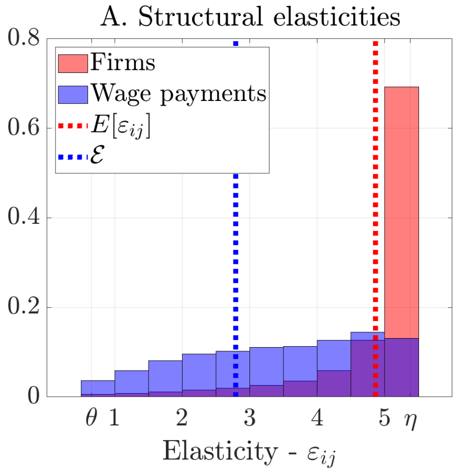

Figure 8: Model implied distribution of structural labor supply elasticities and markdowns

Notes: Panel A plots the distribution of equilibrium structural labor supply elasticities $\varepsilon(\cdot)$ from equation (10), unweighted ('Firms') and weighted ('Wage Payments'). $E[\varepsilon_{ij}]$ is the unweighted mean structural elasticity, and $\mathcal{E}$ is the aggregate structural labor supply elasticity defined in section 6.2. Panel B conducts the same exercise for markdowns. $E[\mu_{ij}]$ is the unweighted mean markdown, and $\mu = \frac{\mathcal{E}}{\mathcal{E}+1}$ is the aggregate markdown defined in section 6.2.

These exercises speak to the two key sources of labor market power in our model. First, at our estimated parameters, firms face upward sloping labor supply curves. Firms internalize this feature of their environment, understanding that hiring an additional worker requires not only a higher wage to the marginal worker, but also all previous workers hired. In a competitive market, firms perceive flat supply curves. Second, in concentrated labor markets firms are non-atomistic and so compete strategically for workers. In this framework, a higher HHI implies that larger firms can *differentially* distort their input choices and wages to an even greater extent.

6.1 Microeconomic measurement

Figure 8 plots the model implied distribution of structural labor supply elasticities ε_{ij} and associated markdowns μ_{ij}. In red, the *average firm* has a labor supply elasticity close to η and a narrow markdown. However, in blue, the distribution of wage payments which determines where the *average dollar of wages* is earned is skewed toward low elasticity, wider markdown firms.

Looking forward, a key step in what follows is the stipulation of a *representative monopsonist* economy that delivers the same allocations as the benchmark *heterogeneous oligopsonist* economy. The elasticity of labor supply that would face these fictitious identical firms would be $\mathcal{E}$, marked in blue. This reflects the distribution of wage payments, and so it is much lower than the average elasticity. Even if a researcher did have the distribution of structural labor supply elasticities, this figure cautions against inferring welfare from simple averages.

The next two exercises use the macroeconomic structure of the model to aggregate these microeconomic distributions into macroeconomic measures of labor market power.

6.2 Macroeconomic measurement I - Counterfactual competitive economy

We define a competitive equilibrium, and then compute the welfare gain associated with a transition to a competitive equilibrium.

Competitive equilibrium. To measure the welfare losses from both sources of market power, we compare our benchmark oligopsonistic equilibrium to a competitive equilibrium. We keep preferences, technology and the distribution of firms-per-market (M_j) fixed, changing only the equilibrium concept. The competitive equilibrium still features upward sloping labor supply curves, but firms do not internalize this. The competitive equilibrium still features finitely many firms in each market, but firms behave as atomistic price takers.[54] Thus, there are no strategic complementarities.

We formally define the competitive equilibrium as follows:

Definition *A **Walrasian equilibrium** is an allocation of employment n_{ijt}, and wages w_{ijt} such that:*

1. *Taking w_{ijt} as given, n_{ijt} solves each firm's optimization problem*
$$n_{ijt} = \arg\max_{n_{ij}} \widetilde{Z}\widetilde{z}_{ijt} n_{ijt}^{\widetilde{\alpha}} - w_{ijt} n_{ijt}$$
2. *Taking w_{ij} as given, n_{ij} solves the household's labor supply problem:*
$$n_{ijt} = \overline{\varphi}\left(\frac{w_{ijt}}{\mathbf{W}_{jt}}\right)^{\eta}\left(\frac{\mathbf{W}_{jt}}{\mathbf{W}_t}\right)^{\theta}\mathbf{W}_t^{\varphi}$$

Figure 9 describes the difference between a firm behaving monopsonistically (Panel A) and competitively (Panel B). The wage is unambiguously higher in the competitive equilibrium, but the net equilibrium effect on employment varies. Since large firms have the widest markdowns in the oligopsonistic equilibrium, their wages increase the most. This reallocates employment away from small firms toward large firms. This undoes the direct effect of small firms' higher wages to the extent that employment may *decline* at small firms. To demonstrate this reallocation, Figure J1 shows how the prices and allocations from our example Figure 4 change under Walrasian competition.

Welfare. Throughout the paper we use the term *welfare gain / loss* to mean the λ_{SS} percent subsidy to consumption in the benchmark oligopsonistic economy that would be required to make the household indifferent with respect to the counterfactual allocation.[55]

First, we compute the welfare gain from competition across steady-states: λ_{SS}. Let $\{\mathbf{C}_o, \mathbf{N}_o\}$ denote consumption and disutility of labor in the benchmark <u>o</u>ligopsonistic equilibrium. Let $\{\mathbf{C}_c, \mathbf{N}_c\}$ denote consumption and disutility of labor in the <u>c</u>ompetitive equilibrium, such that

$$u\left(\left(1+\lambda_{SS}\right)\mathbf{C}_o - \frac{1}{\overline{\varphi}^{\frac{1}{\varphi}}}\frac{\mathbf{N}_o^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}\right) = u\left(\mathbf{C}_c - \frac{1}{\overline{\varphi}^{\frac{1}{\varphi}}}\frac{\mathbf{N}_c^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}\right) \tag{15}$$

[54] Keeping M_j constant in each market purges our exercise of changes in welfare due to 'love of variety' effects.

[55] Note that aggregate consumption incorporates the effect of competition on wages, employment and firm profits. Aggregating firms' profit conditions $\left(\pi_{ij} = y_{ij} - w_{ij}n_{ij} - Rk_{ij}\right)$ under goods market clearing and the definition of $\mathbf{W}$ and $\mathbf{C}$, returns the household budget constraint ($\Pi = \mathbf{C} - \mathbf{WN} - RK$), so $\mathbf{C} = \Pi + \mathbf{WN} + RK$. Recall that $\mathbf{W}$ is defined by $\mathbf{WN} = \int \sum_{i\in j} w_{ij}n_{ij}\,dj$, and $\mathbf{C}$ is defined by $\mathbf{C} = \int \sum_{i\in j} c_{ij}\,dj$.

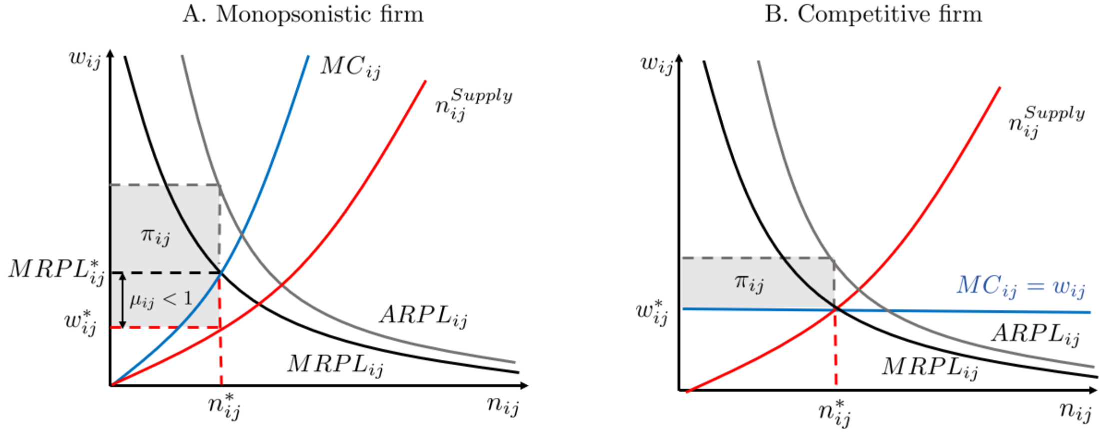

Figure 9: Oligopsonistic vs. Competitive equilibrium

Notes: In a *oligopsonistic equilibrium* (Panel A) the firm understands that its marginal cost MC_{ij} is increasing in its employment. In a *competitive equilibrium* (Panel B) the firm perceives that its marginal cost MC_{ij} is simply equal to its wage, which it takes as given. The true labor supply curve to the firm, however, is still upward sloping, reflecting household preferences.

Second, we compute the welfare gain from competition along the transition path between steady states: λ_{Trans}. We assume that market structure changes, unexpectedly, at date 0:

$$u\left((1+\lambda_{Trans})\,\mathbf{C}_o - \frac{1}{\bar{\varphi}^{\frac{1}{\varphi}}}\frac{\mathbf{N}_o^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}\right) = (1-\beta)\sum_{t=0}^{\infty}\beta^t u\left(\mathbf{C}_t - \frac{1}{\bar{\varphi}^{\frac{1}{\varphi}}}\frac{\mathbf{N}_t^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}\right) \tag{16}$$

As discussed in Section 2, market-level equilibrium payroll shares are independent of aggregates. In Appendix F.2, we show that payroll shares are determined by underlying firm-level productivities, z_{ijt}, and the competitive structure of the economy (e.g. Cournot, Bertrand, Perfect Competition). Since firm-level productivity immediately jumps and then remains constant along the transition path, our model implies that payroll shares $\{s_{ijt}^{wn}\}$, the aggregate wage index $\{\mathbf{W}_t\}$, and the aggregate employment index $\{\mathbf{N}_t\}$ jump immediately to their competitive steady-state values. However, consumption slowly increases to its competitive steady-state level because the representative household's Euler equation dictates a gradual accumulation of capital. As a result, welfare gains along the transition path, λ_{Trans}, are lower than across steady-states, λ_{SS}. Figure A1 and its footnote describe the transition dynamics of the economy in detail and show that utility is higher at all points along the transition.

Table 5 reports welfare gains at our benchmark calibration, which assumes a Frisch elasticity of $\varphi = 0.5$. Across steady states, individuals would require 5.4 percent more consumption under 2014 oligopsonistic labor market conditions in order to compensate them for their losses relative to the competitive benchmark. With higher wages, aggregate employment increases 14.6 percent, producing higher consumption. As expected welfare gains along the transition are smaller, but not drastically so. Welfare gains vary with φ: the larger is φ, the lower the utility cost of additional labor supply so the larger the welfare gains.

Statistical decomposition - Aggregate employment vs. Reallocation. A shift to competitive labor market conditions generates reallocation of labor from lower productivity firms to higher productivity

Frisch elasticity	A. Welfare		B. Labor market		C. Concentration	
φ	Steady state $\lambda_{SS} \times 100$	Transition $\lambda_{Trans} \times 100$	Ave. wage $\mathbf{E}[w_{it}]$	Agg. emp. $\sum_i n_{it}$	Unweighted ΔHHI^{wn}	Weighted ΔHHI^{wn}
0.2	3.3	2.8	41.8	4.1	0.14	0.09
0.5	5.4	4.1	41.6	14.6	0.14	0.09
0.8	7.6	5.6	41.3	26.1	0.14	0.09

Table 5: Welfare gains from competition

Notes: Welfare gain λ_{SS} is given by (15), λ_{Trans} is given by (16). Both correspond to moving from the benchmark oligopsony to competitive equilibrium. Average wage and aggregate employment are expressed in percentage increases from oligopsony to competitive steady-state. Average wage is total wage payments divided by total employment, and aggregate employment is in 'bodies' not disutility.

firms as high productivity firms markdowns narrow disproportionately more. Figure 10 compares low and high productivity firms and their changes in employment across market structures. In the lowest deciles of productivity, firms *decrease* employment. High-productivity, high wage-bill share firms had disproportionately larger markdowns in the oligopsonistic equilibrium. In the competitive equilibrium, they pay disproportionately higher wages and expand, attracting a greater share of employment. The result is a reallocation to bigger, more productive firms. Appendix J characterizes this reallocation by replicating Figure 4 in the competitive equilibrium.

To isolate the contribution of reallocation we decompose output gains into two components through a simple statistical exercise: (1) an aggregate employment effect due to overall higher wages, (2) reallocation toward more productive firms.[56] To isolate scale effects, we distribute the competitive level of aggregate employment among firms according to their employment shares in the benchmark oligopsony equilibrium. Output gains not generated by scale effects are attributable to reallocation. Our main result is that around one-quarter of the 19 percent increase in output is due to aggregate employment effects. This channel is amplified by a high Frisch elasticity, which explains the sensitivity of welfare gain calculations in Table 5.

Despite playing a small role in output, reallocation leads to significant increases in concentration (Table 5). Bigger, more productive firms gain market share in the competitive economy, such that weighted and unweighted wage-bill Herfindahls *increase* by 0.14 and 0.09, respectively. This exercise highlights an important lesson of the paper. Extrapolating from this exercise, pro-competitive labor market reforms may increase concentration and welfare. Concentration itself is an imperfect measure of welfare absent a general equilibrium theory.

Theoretical decomposition - Representative firm and monopsony. Another approach to decomposing welfare gains across steady-states is to identify a representative firm economy that reproduces the

[56] Baqaee and Farhi (2017) advocate a general decomposition of output into technology and reallocation effects. The former are changes in output due to productivity shocks and aggregate changes in factor supply, holding fixed the distribution of input shares across production units. Reallocation effects are changes in output due only to changes in this distribution. Hence, our decomposition into an aggregate employment, or 'scale', effect and a reallocation effect is exactly consistent with theirs. As they document, this has many advantages relative to previously used ad-hoc decompositions.

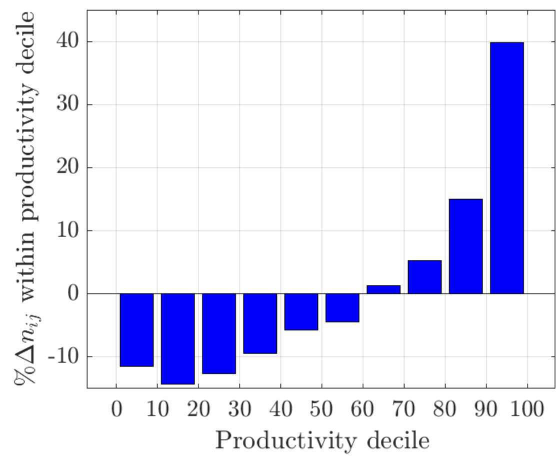

Figure 10: Employment reallocation due to perfect competition.

Notes: Percent change in total employment within productivity decile bin. Change measured between benchmark oligopsony equilibrium and competitive equilibrium. Let n^o_{ij} (N^o) denote firm-level (aggregate) employment in the oligopsonistic equilibrium, and n^c_{ij} (N^c) competitive equilibrium values. Counterfactual firm employment and output under no reallocation $(\widehat{n}_{ij}, \widehat{y}_{ij})$ are computed keeping firms' share of aggregate employment constant: $\widehat{n}_{ij} = (n^o_{ij}/N_o) \times N_c$, $\widehat{y}_{ij} = \frac{1}{1-\alpha(1-\gamma)}\widetilde{Z}\widetilde{z}_{ij}\widehat{n}^{\widetilde{\alpha}}_{ij}$, $\widehat{Y} = \int \sum_i \widehat{y}_{ij}\, dj$. The share of output gains due to aggregate employment effects are $(\widehat{Y} - Y_o)/(Y_c - Y_o)$ where Y_c and Y_o are output in the competitive and oligopsony equilibria.

same macroeconomic aggregates, and then understand what changes in that economy represent wedges between the oligopsony and competitive heterogeneous firm models. We refer to this economy as the *representative firm* economy.

The *representative firm* economy is constructed as follows. A continuum of firms that have identical productivity $\mathbf{Z}$. We assume that these firms face the same labor supply elasticity $\mathcal{E}$ and compete in a national monopsonistic labor market.[57] Each firm solves:

$$\max_{n_{ij},k_{ij}} \quad \mathbf{Z}\left(k_{ij}^{1-\gamma}n_{ij}^{\gamma}\right)^{\alpha} - w_{ij}n_{ij} - Rk_{ij} \tag{17}$$

$$s.t. \quad w_{ij} = \overline{\varphi}^{-\frac{1}{\varphi}}\left(\frac{n_{ij}}{\mathbf{N}}\right)^{\frac{1}{\mathcal{E}}}\mathbf{W}$$

Each firm chooses identical wages, employment and capital, such that aggregate factor demands can be read off of the first order conditions of (17).

Proposition 1. *Under the following two values of* $\mathbf{Z}$ *and* $\mathcal{E}$*, the equilibrium quantities (consumption* $\mathbf{C}$*, output* $\mathbf{Y}$*, capital* K*, labor disutility* $\mathbf{N}$*) and prices (wage index* $\mathbf{W}$*) of the representative firm economy coincide with the*

[57]Another way to view this is that we take our benchmark model, set $z_{ij} = \mathbf{Z}$ for all ij, and then make workers as substitutable across as within markets ($\theta = \eta = \mathcal{E}$) resulting in a national monopsonistic labor market.

equilibrium quantities and prices of the heterogeneous firm benchmark:

$$\frac{\mathcal{E}}{\mathcal{E}+1} = \left[HHI^{wn}_{HA}\left(\frac{\theta}{\theta+1}\right)^{-1} + (1-HHI^{wn}_{HA})\left(\frac{\eta}{\eta+1}\right)^{-1}\right]^{-1}, \; HHI^{wn}_{HA} = \int_0^1 s^{wn}_{j,HA} HHI^{wn}_{j,HA}\, dj \tag{18}$$

$$\mathbf{Z} = \overline{Z}\left[\int \sum_{i=1}^{M_j}\left(z_{ij}\nu_{ij,HA}{}^{\gamma\alpha}\right)^{\frac{1}{1-(1-\gamma)\alpha}} dj\right]^{1-(1-\gamma)\alpha}, \; \nu_{ij,HA} := \frac{n_{ij,HA}}{\mathbf{N}_{HA}}, \tag{19}$$

where $\nu_{ij,HA}$ is a parameter equal to firm ij's share of aggregate labor disutility in the heterogenous firm benchmark equilibrium $\mathbf{N}_{HA}$, and HHI^{wn}_{HA} is the associated concentration measure. The wage index is

$$\mathbf{W} = \frac{\mathcal{E}}{\mathcal{E}+1}\mathbf{MRPL} = \frac{\mathcal{E}}{\mathcal{E}+1}\left(\gamma\alpha\frac{\mathbf{Y}}{\mathbf{N}}\right) \quad \rightarrow \quad \frac{\mathbf{WN}}{\mathbf{Y}} = \frac{\mathcal{E}}{\mathcal{E}+1}\gamma\alpha.$$

Proof *See Appendix E.*

This representation isolates the two sources of inefficiency in the oligopsony economy. First, in a competitive economy, $\mathcal{E} = \infty$ which returns the standard result that equilibrium factor shares are equated to output elasticities, adjusted for the profit share due to decreasing returns. The source of finite $\mathcal{E}$ is clearly due to (i) upward sloping labor supply ($\theta,\ \eta > 0$), and exacerbated by (ii) concentration ($HHI^{wn}_{HA} > 0$). Second, in the competitive equilibrium of the underlying economy we know that employment shares are higher at more productive firms (Figure 10). The higher correlation between z_{ij} and $\nu_{ij,HA}$ in the competitive economy generates higher aggregate productivity. The lower correlation between z_{ij} and $\nu_{ij,HA}$ in the benchmark oligoposony economy implies lower $\mathbf{Z}$ and a misallocation of labor to firms. From capital demand and the aggregate production function, it is clear that this misallocation depresses both capital and output:

$$K = \mathbf{Z}^{\frac{1}{1-(1-\gamma)\alpha}}\left(\frac{(1-\gamma)\alpha}{R}\right)^{\frac{1}{1-(1-\gamma)\alpha}}\mathbf{N}^{\frac{\gamma\alpha}{1-(1-\gamma)\alpha}} \quad , \quad \mathbf{Y} = \mathbf{Z}\left(K^{1-\gamma}\mathbf{N}^{\gamma}\right)^{\alpha}.$$

Decomposition. We can assess the roles of misallocation and labor market power through the lens of the *representative firm* economy. We solve the oligoposony and competitive equilibrium economies as before and compute $(\mathbf{Z}_o, \mathcal{E}_o)$ and $(\mathbf{Z}_c, \mathcal{E}_c)$, respectively. We know that moving from $(\mathbf{Z}_o, \mathcal{E}_o)$ to $(\mathbf{Z}_c, \mathcal{E}_c)$ leads to a welfare gain of 5.4 percent (Table 5, λ_{SS}). Leveraging the representative agent formulation, we can isolate the role of misallocation by leaving $\mathcal{E} = \mathcal{E}_o$, but setting $\mathbf{Z} = \mathbf{Z}_c$. Table 6 shows that this accounts for 56 percent of the increase in welfare in the competitive economy. Shutting down labor market power by setting $\mathcal{E} = \mathcal{E}_c$ while keeping aggregate productivity fixed accounts for 38 percent of the welfare increase.[58] Labor market power and the misallocation it induces are both important sources of welfare losses.

6.3 Macroeconomic measurement II - Labor share

We combine three of the novel contributions of this paper to link the dynamics of labor's share of income to labor market power: (i) the closed-form expression for labor's share of income given by equation (9)

[58] These do not add to 100 since the decomposition is not additive. The interaction of the two forces further reduces welfare.

Counterfactual		Welfare gain (λ_{SS})	Percent of total
Competitive	$(\mathcal{E}_c, \mathbf{Z}_c)$	5.4	—
Efficiency only	$(\mathcal{E}_o, \mathbf{Z}_c)$	3.0	56.0
Labor market only	$(\mathcal{E}_c, \mathbf{Z}_o)$	2.0	37.9

Table 6: Decomposing welfare gains through the lens of the representative agent model

Notes: Welfare gains in the counterfactual exercise consider only steady-state to steady-state, hence λ_{SS}. The *Efficiency only* counterfactual takes the competitive TFP $\mathbf{Z}_c$ but keeps the oligoposony model level of labor market power $\mathcal{E}_o$. The *Labor market only* counterfactual considers the opposite case. The final column gives the fraction of the total welfare gain (λ_{SS}), that is accounted for by the given counterfactual.

of Section 3, (ii) our estimates of θ and η, and (iii) our new measures of wage-bill Herfindahls in Table 1.

We conclude that changes in labor market concentration are unlikely to have contributed to the declining labor share in the United States (e.g. Karabarbounis and Neiman (2013)). The weighted wage-bill Herfindahl fell from 0.20 in 1976 to 0.14 in 2014, which implies that the inverse weighted wage-bill Herfindahl increased from 5.01 to 7.09. Under the assumption of stable preference parameters ($\theta = 0.66, \eta = 5.38$) and technology ($\widetilde{\alpha} = 0.913$) as calibrated in Table 3, equation (9) implies that declining wage-bill Herfindahls between 1976 and 2014 have contributed to an increase in the labor share of 3.13 percentage points.

7 Applications

Having concluded our main measurement exercise we now consider two real world applications of the model. These contribute new theoretical results, demonstrate the applicability of the framework, and further illustrate the complicated relationship between concentration and welfare.

Application I - Minimum wages in strategic labor markets

As an application of the model, we study minimum wages. That monopsony can rationalize small, and positive, employment responses to minimum wages is in part responsible for the theory's historical development (Card and Krueger, 1994; Boal and Ransom, 1997; Manning, 2003). A minimum wage may force some firms paying below their marginal revenue products to compress their markdowns, increase wages, and at the same time expand employment along their labor supply curves. Our model shares this prediction, but due to decreasing returns to scale and strategic complementarities, the mechanics are more complex.

Our main contribution is theoretical. We show that the presence of decreasing returns to scale and strategic complementarities in wages has new implications for the theory of minimum wages, and an elegant solution in our model. We provide a graphical characterization of the theory and then show that, consistent with the theory, our calibrated model features an optimal minimum wage that significantly increases welfare relative to the baseline.

Theory. Firms choose employment subject to the household labor supply constraint and the additional constraint that their wage w_{ij} is greater than a minimum wage $\underline{w}$. At the minimum wage, households may wish to supply more labor than—due to decreasing returns to scale—a firm demands, so we add a firm-by-firm constraint that labor supply is less than labor demand, which the household takes as given. The representative household maximizes utility as before with the additional constraint $(*)$: $n_{ijt} \leq \underline{n}_{ijt}$.

An intuitive shadow wage $\widetilde{w}_{ijt}$ equates labor supply and demand in this context. The marginal utility of sending an extra worker to firm ij when the minimum wage binds and labor supply would otherwise outstrip demand is given by the multiplier ν_{ijt} associated with $(*)$.[59] From the household's perspective, $\underline{n}_{ijt}$ workers are supplied according to the *shadow wage*, $\widetilde{w}_{ijt} = w_{ijt} - \nu_{ijt}$.

The key theoretical result is that the Cournot-Nash equilibrium of labor market j can be solved as before but in terms of *shadow* objects. In fact, the equilibrium is characterized by a vector of markdowns that are functions of *shadow* payroll shares $\mu\left(\widetilde{s}^{wn}_{ijt}\right)$.[60] Strategic complementarities are also now rendered through shadow wages. In the baseline model a higher wage at one firm always leads to a higher wage as a best response from a competitor. Here, despite an increase in the minimum wage pushing one firm's wages up, its unconstrained competitors may cut wages in response to the firm's falling shadow wage. The macroeconomy is then solved in terms of the *shadow* aggregate wage $\widetilde{\mathbf{W}}_t$. We derive these result and provide a solution algorithm in detail in Appendix I.

Characterization. Figure 11 illustrates the economics underlying the impact of a minimum wage on a firm. There are four relevant cases:

$$\text{Range for } \underline{w}\text{:}\quad 0 \underset{\textit{Region I}}{\xrightarrow{\hspace{2cm}}} w^*_{ij} \underset{\textit{Region II}}{\xrightarrow{\hspace{2cm}}} w^{Comp.}_{ij} \underset{\textit{Region III}}{\xrightarrow{\hspace{2cm}}} MC^*_{ij} \underset{\textit{Region IV}}{\xrightarrow{\hspace{2cm}}} \infty$$

Panel A illustrates the impact of a very low minimum wage. In *Region I* the minimum wage has no effect on equilibrium labor supply ($\underline{w} < w^*_{ij}$). In Panel B, the minimum wage now binds, but still below the competitive wage. Here in *Region II* the firm absorbs the effects of minimum wage into its markdown. Employment increases relative to *Region I* and the household remains on its labor supply curve.

The marginal cost curve is now quite different from the benchmark economy. The new marginal cost curve is horizontal and equal to $\underline{w}$ until it reaches the labor supply curve at $\underline{n}_{ij}$ workers. Up to this point workers are paid $\underline{w}$. Marginal cost then jumps, as above the minimum wage, additional hiring requires increasing pay for existing workers. Since marginal cost jumps above the marginal revenue product of labor, profit maximizing employment is $\underline{n}_{ij}$. Firms still generate profits from a non-zero markdown μ_{ij} and the wedge between average and marginal revenue products due to decreasing returns to scale.

Increasing the minimum wage further pushes the firm into *Region III* (Panel C). Here the minimum wage is above the competitive wage so labor supply exceeds demand. Since $w_{ij} = \underline{w} = MRPL_{ij}$, the

[59] We normalize ν_{ijt} by the household budget constraint multiplier for ease of interpretation. We discuss these details in Appendix I.

[60] The shadow wage payment share is $\widetilde{s}^{wn}_{ijt} = \widetilde{w}_{ijt} n_{ijt} / \sum_{k\in j} \widetilde{w}_{kjt} n_{kjt}$, where $\widetilde{w}_{ijt} = w_{ijt} - \nu_{ijt}$.

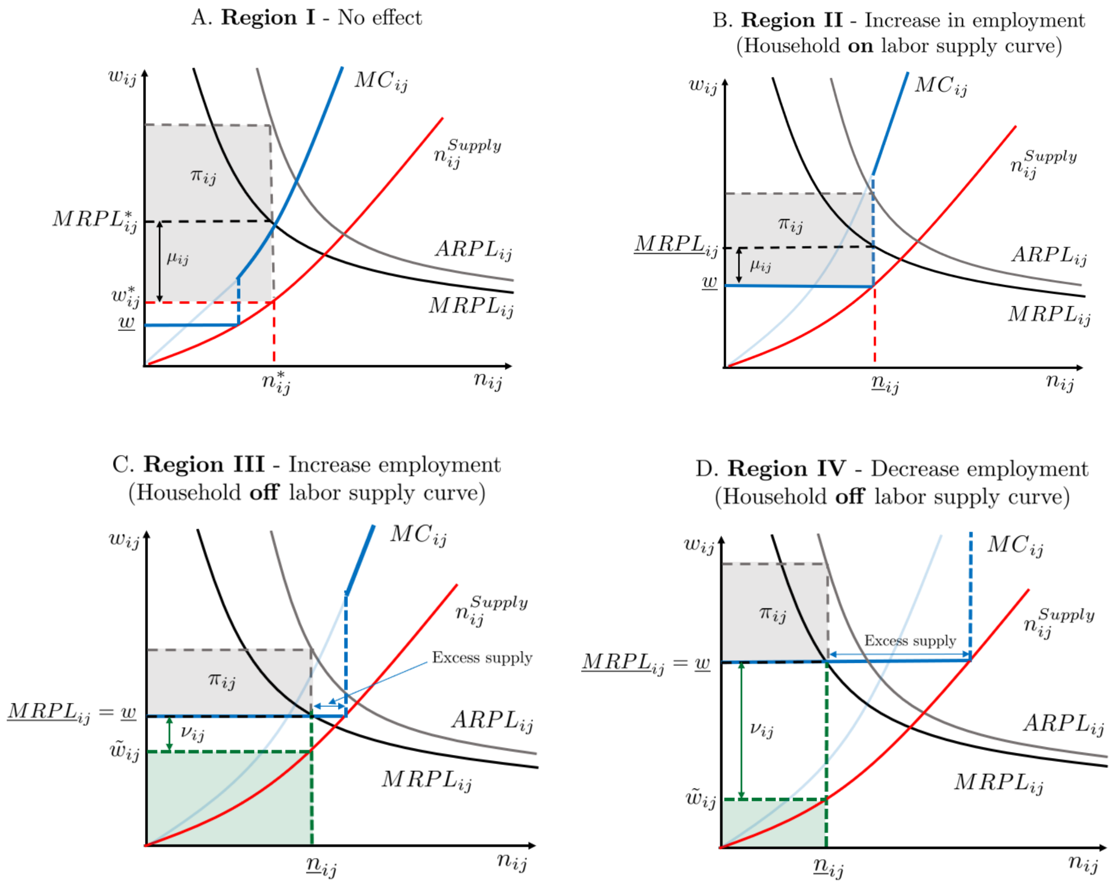

Figure 11: Theory of minimum wage

wage markdown is zero but profits π_{ij} are positive due to decreasing returns to scale.[61] Our theory rationalizes household labor supply of $\underline{n}_{ij} < n_{ij}^{Supply}$ through the shadow wage. At $\widetilde{w}_{ij} = \underline{w} - \nu_{ij}$, the household supplies $\underline{n}_{ij}$ workers to the firm. Our key result was that competitors respond to the *shadow* payroll share of the firm (in green), which is less than the measured share, and falling as $\underline{w}$ increases.

Increasing the minimum wage beyond the equilibrium marginal cost in the unconstrained case causes the firm to enter *Region IV* (Panel D). The same economics apply as in *Region III*, but here equilibrium employment is *less* than what would occur absent a minimum wage. Note also that relative to *Region III* the *shadow* wage and shadow share have declined as ν_{ij} increases as excess supply widens.

Our *shadow wage* result permits a sharp characterization of the equilibrium effect of a minimum wage on unconstrained firms (*Region I*). The shadow wages $\widetilde{w}_{ij}$ of their smaller competitors are lower than their actual wages, and falling as the minimum wage increases. As an unconstrained firm responds to the shadow sectoral wage $\widetilde{\mathbf{W}}_j$, which is falling, they best-respond by cutting their own wages. Despite lower wages, employment at large firms grow as employment is reallocated from small to large firms.

[61] Note that *Region III* does not exist with constant returns to scale. With constant returns to scale the competitive wage is equal to $MRPL_{ij}$ which is a constant. Therefore as $\underline{w}$ increases past the competitive wage, the firm exits.

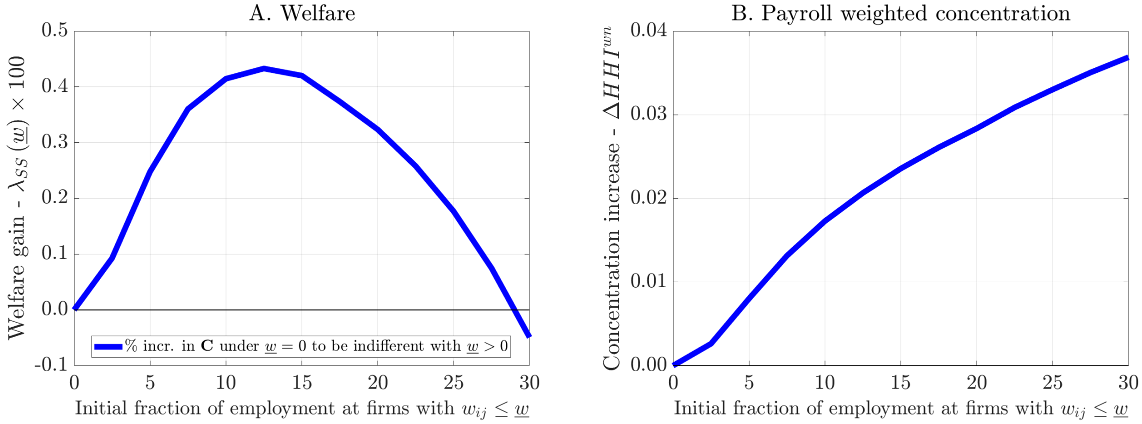

Figure 12: Welfare, concentration and the minimum wage

Notes: On the y-axis the figure plots the $\lambda(\underline{w})$ associated with each level of the minimum wage $\underline{w}$ (see equation (15)). This gives the percent increase in aggregate consumption in the benchmark economy ($\underline{w} = 0$) required to leave the household indifferent between the benchmark allocation and the allocation under a positive minimum wage ($\underline{w} > 0$). On the x-axis is the fraction of employment in the *benchmark economy* ($\underline{w} = 0$) that is employed at firms that pay a wage $w_{ij} < \underline{w}$.

Welfare maximizing minimum wage. In our model, a welfare maximizing minimum wage exists: $\underline{w}^* > 0$. This trades off narrower markups at small firms against more market power for large firms, a mechanism that is unique to our model. Figure 12A shows that the optimal minimum wage welfare gains are roughly .4 percent, delivering around one tenth of the welfare gains associated with the Walrasian allocation (Table 5). The latter could be obtained through minimum wages but would require a menu of $\underline{w}_{ij}$ each equal to the firm specific competitive wage. In the presence of significant heterogeneity, an economy-wide minimum wage is a relatively blunt tool at undoing labor market power. Indeed, this heterogeneity implies that as minimum wages rise too much, concentration (Panel B.) and the market power of unconstrained firms increases undoing the positive welfare gains of a lower minimum wage.

Application II - Mergers in strategic labor markets

Here we conduct a simple quantitative experiment to show the potential welfare effects that might propogate through local labor markets following mergers. We do so, as we have throughout, by abstracting from product market effects and potential productivity gains associated with mergers. For both a comprehensive exposition of the theory and cross-sectional empirical analysis of mergers see Berger, Herkenhoff, and Mongey (2019).

Experiment. Consider a single sector in the baseline oligopsony model and take two single plant firms i and i'. A *merger* gives control of both plants to a single central manager who chooses employment at both plants simultaneously. As usual the central manager takes competitors' employment decisions as given, but now internalizes the spillovers from labor demand at plant i on wages at plant i'. To bound welfare gains / losses we consider three possible cases, merging firms in all markets with 5 or more firms: (i) the two least productive firms in a market merge, (ii) the two most productive merge, (iii) two

Merger firms	A. Welfare	B. Labor market		C. Concentration		D. Counterfactual	
	Steady state $\lambda_{SS} \times 100$	Ave. wage $\mathbf{E}[w_{it}]$	Agg. emp. $\sum_i n_{it}$	Unweighted ΔHHI^{wn}	Weighted ΔHHI^{wn}	Unweighted ΔHHI^{wn}	Weighted ΔHHI^{wn}
Two highest productivity	-1.063	-2.83	-0.38	0.013	0.025	0.044	0.070
Two random	-0.083	-0.23	-0.34	0.003	0.009	0.005	0.013
Two lowest productivity	0.003	-0.01	-0.10	0.001	0.003	0.000	0.002

Table 7: Welfare effects of mergers

Notes: The welfare gain λ_{SS} is given by (15). These correspond to moving from benchmark oligopsony economy to oligopsony economy under mergers. In all three cases mergers occur in all markets with 5 or more firms. The counterfactual change in concentration is obtained by adding the benchmark equilibrium shares of the merging firms, and computing concentration measures with these combined shares and benchmark equilibrium shares of all other firms. Comparisons of concentration are for markets with mergers only. For comparison, recall that in the benchmark equilibrium the unweighted (weighted) HHI^{wn} is 0.33 (0.14).

random firms merge.[62]

Results. Table 7 provides results. Merging the two largest firms in each market can cause sizeable welfare losses, with λ_{SS} around minus 1 percent. The merged firms exert additional market power, cutting wages and restricting quantity (employment). With less employment at the most productive firms, misallocation worsens and total factor productivity falls.

Results vary by the type of merger. With less directed mergers, welfare losses shrink substantially as misallocation effects are muted. Merging low productivity firms is welfare improving. As the merged firms increase their market power their wages fall, which reallocates labor toward more productive firms.

Two important lessons emerge from the small positive effects on concentration. First, the endogenous response of firms mitigates the 'partial equilibrium' increase in concentration that comes from combining firms. Suppose that the shares of non-merging firms remain constant, the increase in concentration given by this counterfactual is given in Panel D.[63] Such back of the envelope calculations overstate the equilibrium increase in concentration by a factor of three. Merging firms increase market power, reducing wages and employment, which reallocates employment to other firms. This dampens the increase in concentration that comes from effectively removing a firm from the market, especially so when the merging firms are the largest firms.

Second, the exercise highlights the deceptive nature of concentration as a measure of welfare. Our main competitive equilibrium counterfactual of Section 6 was a *'pro-competitive'* intervention that drove up concentration and increased welfare. The merger experiment is an *'anti-competitive'* intervention that also drives up concentration but reduces welfare. The relationship is non-trivial.

We view these results as suggestive evidence that mergers between the largest firms in a local labor market may generate significant welfare losses through more labor market power and increased misallocation. We further explore the welfare implications of mergers as well as optimal merger policy in Berger, Herkenhoff, and Mongey (2019).

[62] We focus on markets with 5 or more firms in order to make the distinction between two low vs. two high productivity firms meaningful.

[63] More precisely, counterfactual market j payroll concentration uses benchmark equilibrium payroll shares and is computed $HHI^{wn}_{j,Counterfactual} = \sum_{k \notin \{i,i'\}} = (s^{wn}_{kj})^2 + (s_{ij} + s_{i'j})^2$.

8 Conclusion

In this paper, we develop a general equilibrium model of labor market oligopsony. We use the framework to (1) inform measurement of labor market concentration and map labor market concentration to labor market power, (2) link labor market power to labor's share of income, (3) measure the welfare losses of labor market power, and (4) the effects of minimum wages and mergers in strategic environments.

In our framework, we show that the relevant measure of labor market concentration is the wage-bill Herfindahl and the distribution of wage-bill Herfindahls is a sufficient statistic for the labor share. We apply our measures of labor market concentration to tradeable sector firms in the Longitudinal Business Database (LBD). We show that the payroll weighted wage-bill Herfindahl fell from 0.20 to 0.14 between 1976 and 2014, indicating a significant decrease in labor market concentration. Using our theory's closed-form mapping between labor's share of income and wage-bill Herfindahls, we show that declining labor market concentration has *increased* labor's share of income by 3.13 percent between 1976 and 2014.

To assess the normative implications of our measures of labor market concentration, we estimate our model and conduct several counterfactuals. We use within-state-firm, across-market differences in the response of employment and wages to state corporate tax changes (Giroud and Rauh, 2019) to estimate reduced form size-dependent labor supply elasticities. In conjunction with the model, these reduced form estimates imply parameters and structural elasticities that are key to measuring welfare. To test how sensible our estimated model is, we show that the model successfully replicates two key non-targeted moments: the large discrepancy that we document between weighted and unweighted distributions of market concentration, and the degree of imperfect pass-through between value added per worker and wages. To the best of our knowledge our model is the first to simultaneously replicate the difference between weighted and unweighted Herfindahls, as well as generate a pass-through rate from value added per worker to wages that is less than 100%.

We then use our model to measure the welfare gains associated with a transition to a competitive equilibrium. We find that households in our 2014 benchmark economy would require an additional 5.4 percent more lifetime consumption to be indifferent between the status-quo and transitioning to a competitive economy with no labor market power. Welfare gains associated with lower labor market power come from a reallocation of workers from smaller, less productive firms to larger, more productive firms.

Finally, as applications of the model, we derive a new theoretical characterization of minimum wages under oligopsony and decreasing returns to scale. Our *perceived wage* formulation allows us to characterize the hump-shaped profile of welfare with respect to the minimum wage. Concentration goes up, but welfare is non-monotonic. The same holds for merger counterfactuals. When two productive firms merge or two unproductive firms merge, concentration increases. However, the former generates welfare losses whereas the latter generates welfare gains. Thus, these two applications make clear that an observed increase in concentration in a market cannot be used to make statements about welfare without understanding driving forces.

References

AMITI, M., O. ITSKHOKI, AND J. KONINGS (2014): "Importers, exporters, and exchange rate disconnect," *American Economic Review*, 104(7), 1942–78.

——— (2016): "International shocks and domestic prices: how large are strategic complementarities?," NBER Working Paper 22119, National Bureau of Economic Research.

ANDERSON, S. P., A. DE PALMA, AND J.-F. THISSE (1987): "The CES is a Discrete Choice Model?," *Economics Letters*, 24(2), 139–140.

ASHENFELTER, O. C., H. FARBER, AND M. R. RANSOM (2010): "Labor Market Monopsony," *Journal of Labor Economics*, 28(2), 203–210.

ATKESON, A., AND A. BURSTEIN (2008): "Pricing-to-market, trade costs, and international relative prices," *American Economic Review*, 98(5), 1998–2031.

AUTOR, D., D. DORN, L. F. KATZ, C. PATTERSON, AND J. VAN REENEN (2017): "The fall of the labor share and the rise of superstar firms," NBER Working Paper 23396, National Bureau of Economic Research.

AZAR, J., I. MARINESCU, AND M. I. STEINBAUM (2017): "Labor market concentration," NBER Working Paper 24147, National Bureau of Economic Research.

AZAR, J. A., I. MARINESCU, M. I. STEINBAUM, AND B. TASKA (2019): "Concentration in US labor markets: Evidence from online vacancy data," NBER Working Paper 24395, National Bureau of Economic Research.

BANFI, S., AND B. VILLENA-ROLDAN (2018): "Do high-wage jobs attract more applicants? Directed search evidence from the online labor market," *Journal of Labor Economics*, 37(3).

BAQAEE, D. R., AND E. FARHI (2017): "Productivity and Misallocation in General Equilibrium," NBER Working Paper 24007, National Bureau of Economic Research.

BARKAI, S. (2016): "Declining labor and capital shares," New Working Paper Series 2, Stigler Center for the Study of the Economy and the State.

BECKER, G. S. (1962): "Investment in human capital: A theoretical analysis," *Journal of Political Economy*, 70(5, Part 2), 9–49.

BELOT, M., P. KIRCHER, AND P. MULLER (2017): "How Wage Announcements Affect Job Search Behaviour-A Field Experimental Investigation," Working papers, European University Institute.

BENMELECH, E., N. BERGMAN, AND H. KIM (2018): "Strong employers and weak employees: How does employer concentration affect wages?," NBER Working Paper 24307, National Bureau of Economic Research.

BERGER, D., K. HERKENHOFF, AND S. MONGEY (2019): "Mergers and the Labor Market," Working paper.

BHASKAR, V., A. MANNING, AND T. TO (2002): "Oligopsony and monopsonistic competition in labor markets," *Journal of Economic Perspectives*, 16(2), 155–174.

BLOOM, N., F. GUVENEN, B. S. SMITH, J. SONG, AND T. VON WACHTER (2018): "Inequality and the disappearing large firm wage premium," in *American Economic Review Papers & Proceedings*, vol. 108, pp. 317–22.

BOAL, W. M., AND M. R. RANSOM (1997): "Monopsony in the Labor Market," *Journal of Economic Literature*, 35(1), 86–112.

BOROVICKOVA, K., AND R. SHIMER (2017): "High Wage Workers Work for High Wage Firms," NBER Working Paper 24074, National Bureau of Economic Research.

BROOKS, W. J., J. P. KABOSKI, Y. A. LI, AND W. QIAN (2019): "Exploitation of Labor? Classical Monopsony Power and Labor's Share," NBER Working Paper 25660, National Bureau of Economic Research.

BROWN, C., AND J. MEDOFF (1989): "The employer size-wage effect," *Journal of Political Economy*, 97(5), 1027–1059.

BURDETT, K., AND D. MORTENSEN (1998): "Wage differentials, employer size, and unemployment," *International Economic Review*, 39(2), 257–73.

CARD, D., A. R. CARDOSO, J. HEINING, AND P. KLINE (2018): "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 36(S1), 13–70.

CARD, D., AND A. B. KRUEGER (1994): "Pricing-to-market, trade costs, and international relative prices," *American Economic Review*, 84(4), 772–793.

DE LOECKER, J., AND J. EECKHOUT (2017): "The Rise of Market Power and the Macroeconomic Implications," NBER Working Paper 23687, National Bureau of Economic Research.

DELGADO, M., R. BRYDEN, AND S. ZYONTZ (2014): "Categorization of traded and local industries in the US economy," Discussion paper, Mimeo. Available online at: http://www. clustermapping. us/[Accessed May 2015].

DUBE, A., D. CENGIZ, A. LINDNER, AND B. ZIPPERER (2019): "The effect of minimum wages on low-wage jobs: Evidence from the United States using a bunching estimator," *Quarterly Journal of Economics*, 134(3).

DUBE, A., J. JACOBS, S. NAIDU, AND S. SURI (2019): "Monopsony in Online Labor Markets," *American Economic Review: Insights*, Forthcoming.

EDMOND, C., V. MIDRIGAN, AND D. Y. XU (2018): "How Costly Are Markups?," NBER Working Paper 24800, National Bureau of Economic Research.

FORT, T. C., AND S. D. KLIMEK (2016): "The Effects of Industry Classification Changes on US Employment Composition," *Tuck School at Dartmouth*.

FRIEDRICH, B., L. LAUN, C. MEGHIR, AND L. PISTAFERRI (2019): "Earnings Dynamics and Firm-Level Shocks," NBER Working Paper 25786, National Bureau of Economic Research.

GIANDREA, M. D., AND S. A. SPRAGUE (2017): "Estimating the US labor share," *Monthly Labor Review*.

GIROUD, X., AND J. RAUH (2019): "State Taxation and the Reallocation of Business Activity: Evidence from Establishment-Level Data," *Journal of Political Economy*, Forthcoming.

GRAHAM, J. R., M. T. LEARY, AND M. R. ROBERTS (2015): "A century of capital structure: The leveraging of corporate America," *Journal of Financial Economics*, 118(3), 658–683.

GUTIÉRREZ, G., AND T. PHILIPPON (2016): "Investment-less growth: An empirical investigation," NBER Working Paper 22897, National Bureau of Economic Research.

HANDWERKER, E. W., AND M. DEY (2018): "Megafirms and Monopsonists: Not the same employers, not the same workers," *BLS Working Paper Series*, 2.

HERSHBEIN, B., AND C. MACALUSO (2018): "Labor market concentration and the demand for skills," .

HOTELLING, H. (1990): "Stability in competition," in *The Collected Economics Articles of Harold Hotelling*, pp. 50–63. Springer.

JAROSCH, G., J. S. NIMCSIK, AND I. SORKIN (2019): "Granular Search, Market Structure, and Wages," NBER Working Paper 26239, National Bureau of Economic Research.

KARABARBOUNIS, L., AND B. NEIMAN (2013): "The Global Decline of the Labor Share," *The Quarterly Journal of Economics*, 129(1), 61–103.

KEANE, M., AND R. ROGERSON (2012): "Micro and macro labor supply elasticities: A reassessment of conventional wisdom," *Journal of Economic Literature*, 50(2), 464–76.

KENNAN, J., AND J. R. WALKER (2011): "The effect of expected income on individual migration decisions," *Econometrica*, 79(1), 211–251.

KLINE, P., N. PETKOVA, H. WILLIAMS, AND O. ZIDAR (2019): "Who Profits from Patents? Rent-Sharing at Innovative Firms," *Quarterly Journal of Economics*, 134(3), 1343–404.

LALLEMAND, T., R. PLASMAN, AND F. RYCX (2007): "The Establishment-size Wage Premium: Evidence from European Countries," *Empirica*, 34(5), 427–451.

LAMADON, T., M. MOGSTAD, AND B. SETZLER (2019): "Imperfect Competition and Rent Sharing in the U.S. Labor Market," NBER Working Paper 25954, National Bureau of Economic Research.

MALMBERG, H. (2013): "Random Choice over a Continuous Set of Options," Ph.D. thesis, Department of Mathematics, Stockholm University.

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

——— (2011): "Imperfect Competition in the Labor Market," in *Handbook of labor economics*, vol. 4, pp. 973–1041. Elsevier.

MITMAN, K., AND S. RABINOVICH (2019): "Do Changes in Unemployment Insurance Explain the Emergence of Jobless Recoveries?," Discussion Paper Series ZA DP No. 12365, IZA Institute of Labor Economics.

REICHLING, F., AND C. WHALEN (2012): "Review of Estimates of the Frisch Elasticity of Labor Supply," Working Paper 2012-13, Congressional Budget Office Washington, D.C.

RINZ, K. (2018): "Labor Market Concentration, Earnings Inequality, and Earnings Mobility," CARRA Working Paper Series 2018-10, Center for Administrative Records Research and Applications , US Census Bureau.

ROBINSON, J. (1933): *The Economics of Imperfect Competition*. Palgrave Macmillan.

Rossi-Hansberg, E., P.-D. Sarte, and N. Trachter (2018): "Diverging Trends in National and Local Concentration," NBER Working Paper 25066, National Bureau of Economic Research.

Salop, S. C. (1979): "Monopolistic Competition with Outside Goods," *The Bell Journal of Economics*, pp. 141–156.

Staiger, D. O., J. Spetz, and C. S. Phibbs (2010): "Is There Monopsony in the Labor Market? Evidence from a Natural Experiment," *Journal of Labor Economics*, 28(2), 211–236.

Suárez Serrato, J. C., and O. Zidar (2016): "Who benefits from state corporate tax cuts? A local labor markets approach with heterogeneous firms," *American Economic Review*, 106(9), 2582–2624.

Verboven, F. (1996): "The nested logit model and representative consumer theory," *Economics Letters*, 50(1), 57–63.

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

Appendix For Online Publication

This Appendix is organized as follows. Section A provides additional tables and figures references in the text. Section B provides our micro-foundation for nested-CES preferences used in the main text and references in Section 2. Section C contains details about the data and sample selection criteria. Section E contains derivations of the household labor supply curves, optimal firm markdowns, and other formulas referenced in the main text. Section F contains additional details regarding the computation of the baseline model. Section G provides a model of the effect of corporate taxes on the marginal revenue product of labor. Section H provides additional details regarding the calibration. Section I provides our solution algorithm for the model with a minimum wage.

A Additional tables and figures

Moment	Model	Data
A. Unweighted		
Wage-bill Herfindahl (unweighted)	0.35	0.45
Std. Dev. of Wage-bill Herfindahl (unweighted)	0.33	0.33
Skewness of Wage-bill Herfindahl (unweighted)	1.07	0.48
B. Weighted		
Wage-bill Herfindahl (weighted by market's share of total payroll)	0.14	0.14
Std. Dev. of Wage-bill Herfindahl (weighted by market's share of total payroll)	0.03	0.20
Skewness of Wage-bill Herfindahl (weighted by market's share of total payroll)	3.01	2.20
C. Correlations of Wage-bill Herfindahl		
Number of firms	-0.52	-0.21
Std. Dev. Of Relative Wages	-0.31	-0.51
Employment Herfindahl	1.00	0.98
Market Employment	-0.75	-0.21

Table A1: Labor market concentration and cross-market correlations, model versus data

Notes: Benchmark oligopsonistic equilibrium. See data notes in Section 1.

Variable		Mean	Std. Dev.
Corporate tax rate (percent)	$\tau_{s(k)t}$	7.14	3.19
Change in corporate tax rate	$\Delta\tau_{s(k)t}$	0.05	0.78
Total Pay At Firm (Thousands)	$w_{ijt}n_{ijt}$	2,148	19,010
Employment	n_{ijt}	37.99	215.2
Wage bill Herfindahl	HHI^{wn}_{jt}	0.10	0.16
Employment Herfindahl	HHI^{n}_{jt}	0.09	0.15
Wage bill share	s^{wn}_{ijkt}	0.03	0.12
Employment share	s^{n}_{ijt}	0.03	0.11
Number of firms per market	M_j	1,345	2,813
Log number of firms per market	$\log M_j$	5.56	2.01
Log employment	$\log n_{ijkt}$	2.39	1.32
Log wage	$\log w_{ijkt}$	3.58	0.71
Observations			4,425,000

Table A2: Regression sample summary statistics

Notes: Tradeable C-Corps from 2002 to 2012.

Description	Model	Data (KPZW)
Log change in VAPW (VAPW=$\widetilde{Z}\tilde{z}_{ij}n_{ij}^{\tilde{\alpha}-1}$)	0.13	0.13
Median firm size	25.26	25.26
Mean firm size	50.55	61.49
Median VAPW (dollars)	81900	86870
Mean VAPW (dollars)	83734	120160
Model Simulation Parameters		
Size cutoff (Employees)	7.00	
Fraction of Firms Shocked	0.01	
Shock size ($dlog(\tilde{z}_{ij})$)	0.15	

Table A3: Wage pass-through experiment details

Notes: Summary statistics for replication of Kline, Petkova, Williams, and Zidar (2019) regressions. We randomly sample one percent of firms in our benchmark economy. We draw firms with employment greater than $\underline{n}$. We increase the productivity of treated firms by a factor $\Delta log\tilde{z}_ij$. The values of $\underline{n}$ and Δ are calibrated to match the KPWZ (1) median firm size of 25 employees, (2) increase in post-tax value added per worker of 13 percent. We keep aggregates fixed and solve the new market equilibrium. We treat the untreated and treated observations for each firm as a panel with two observations per firm of wages $\left\{w_{ij0}, w_{ij1}\right\}$ and value added per worker, $\left\{\frac{y_{ij0}}{n_{ij0}}, \frac{y_{ij1}}{n_{ij1}}\right\}$. We then regress the wages in levels on VAPW in levels and a firm-specific fixed effect. The regression coefficient is converted into an elasticity using untreated mean wages and mean value added per worker.

Wage bill Herfindahl	Model	Data
Payroll weighted average	0.14	0.14
Unweighted average	0.33	0.45
Correlation with market employment	-0.80	-0.21

Table A4: Concentration and competition

Notes: Reports the average HHI_j^{wn} weighted by employment across markets. Computed for the baseline calibration with Frisch elasticity of φ.

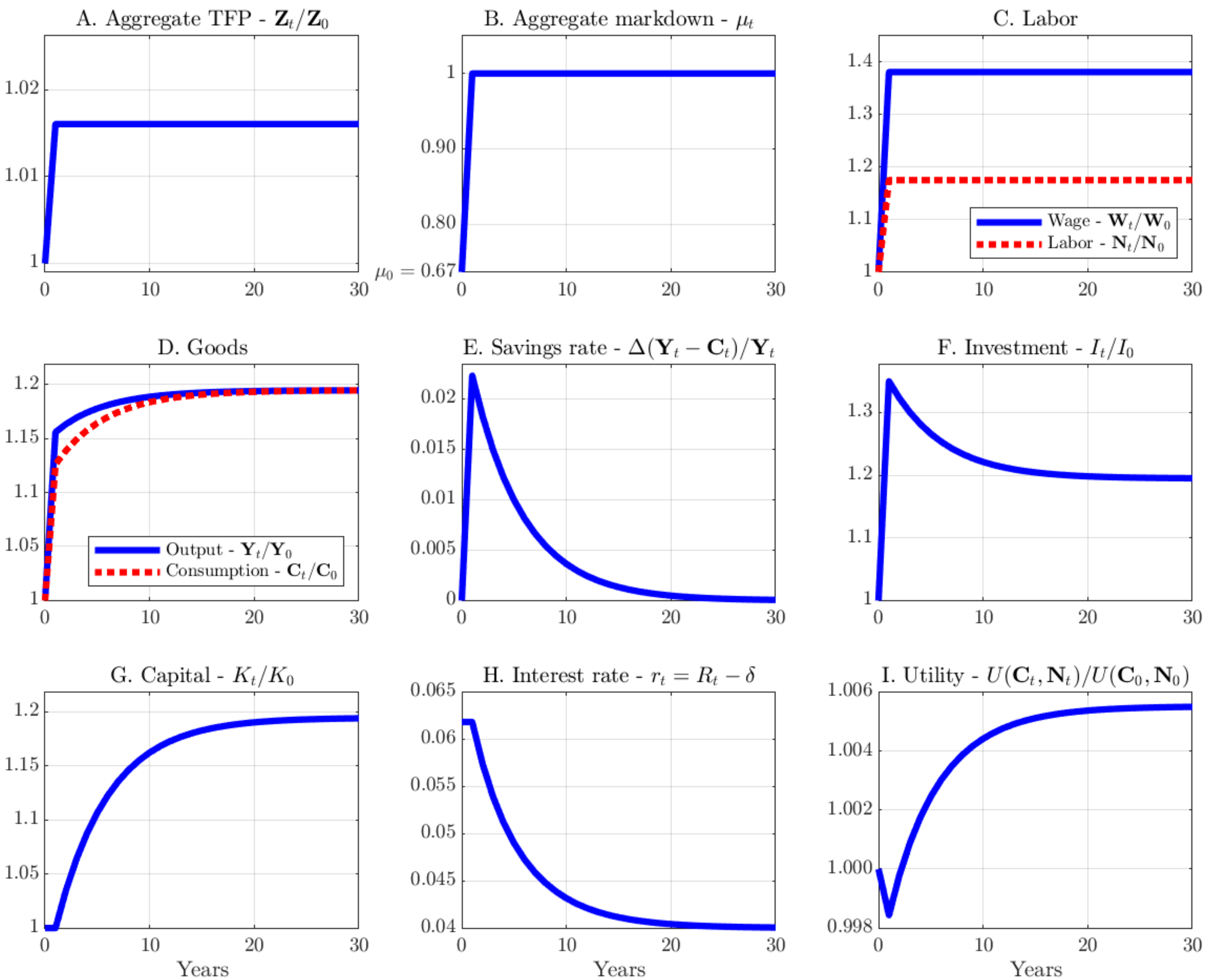

Figure A1: Transition dynamics to change in market structure.

Notes: This figure provides transition dynamics of aggregates to an unexpected change in market structure in period $t = 1$. Transition dynamics are computed under $\varphi = 0.50$. Aggregate TFP $\mathbf{Z}_t$ is as in the representative agent model (19), and similarly the aggregate markdown $\mu_t = \frac{\mathcal{E}_t}{\mathcal{E}_t+1}$ with $\mathcal{E}_t$ from (18). Both depend only on payroll shares, which are determined in labor equilibria and so independent of aggregates. Given shares, we can compute $\mathbf{W}_t$. So $\mathbf{W}_t$ jumps, as does $\mathbf{N}_t$ given the labor supply curve $\mathbf{N}_t = \overline{\varphi}\mathbf{W}_t^{\varphi}$. The path for capital and consumption is then determined by the resource constraint $\mathbf{C}_t = \mathbf{Y}_t + (1-\delta)K_t - K_{t+1}$, household Euler equation $u_C(\mathbf{C}_t, \mathbf{N}_t) = \beta u_C(\mathbf{C}_{t+1}, \mathbf{N}_{t+1})\left[R_{t+1} + 1 - \delta\right]$, and equilibrium price of capital $R_t K_t = (1-\gamma)\alpha\mathbf{Y}_t$. Since capital is undistorted, its paid the competitive factor share equal to its output elasticity.

B Microfounding the nested CES labor supply system

In this section we provide a micro-foundation for the nested CES preferences used in the main text. The arguments used here adapt those in Verboven (1996). We begin with the case of monopsonistic competition to develop ideas and then move to the case of oligopsonistic labor markets studied in the text. We then show that the same supply system occurs in a setting where workers solve a dynamic discrete choice problem and firms compete in a dynamic oligopoly.

B.1 Static discrete choice framework

Agents. There is a unit measure of ex-ante identical individuals indexed by $l \in [0, 1]$. There is a large but finite set of J sectors in the economy, with finitely many firms $i \in \{1, \ldots, M_j\}$ in each sector.

Preferences. Each individual has random preferences for working at each firm ij. Their disutility of labor supply is *convex* in hours worked h_l. Worker l's disutility of working h_{lij} hours at firm ij are:

$$\nu_{lij} = e^{-\mu\varepsilon_{lij}} h_{lij} \qquad , \qquad \log \nu_{lij} = \log h_{ij} - \mu\varepsilon_{ij},$$

where the random utility term ε_{lij} from a multi-variate Gumbel distribution:

$$F(\varepsilon_{i1}, ..., \varepsilon_{NJ}) = \exp\left[-\sum_{ij} e^{-(1+\eta)\varepsilon_{ij}}\right].$$

The term ε_{lij} is a worker-firm specific term which reduces labor disutility and hence could capture (i) an inverse measure of commuting costs, or (ii) a positive amenity.

Decisions. Each individual must earn $y_l \sim F(y)$, where earnings $y_l = w_{ij} h_{lij}$. After drawing their vector $\{\varepsilon_{lij}\}$, each worker solves

$$\min_{ij} \left\{\log h_{ij} - \varepsilon_{lij}\right\} \equiv \max_{ij} \left\{\log w_{ij} - \log y_l + \varepsilon_{lij}\right\}.$$

This problem delivers the following probability that worker l chooses to work at firm ij, which is independent of y_l:

$$Prob_l\left(w_{ij}, w_{-ij}\right) = \frac{w_{ij}^{1+\eta}}{\sum_{ij} w_{ij}^{1+\eta}}. \tag{B1}$$

Aggregation. Total labor supply to firm ij, is then found by integrating these probabilities, multiplied by the hours supplied by each worker l:

$$
\begin{aligned}
n_{ij} &= \int_0^1 Prob_l\left(w_{ij}, w_{-ij}\right) h_{lij} dF\left(y_l\right) \qquad , \qquad h_{lij} = y_l / w_{ij} \\
n_{ij} &= \frac{w_{ij}^{\eta}}{\sum_{i \in j} w_{ij}^{1+\eta}} \underbrace{\int_0^1 y_l dF\left(y_l\right)}_{:=Y}
\end{aligned}
\tag{B2}
$$

Aggregating this expression we obtain the obvious result that $\sum_{i \in j} w_{ij} n_{ij} = Y$. Now define the following indexes:

$$
\mathbf{W} := \left[\sum_{i \in j} w_{ij}^{1+\eta}\right]^{\frac{1}{1+\eta}} \quad , \quad \mathbf{N} := \left[\sum_{i \in j} n_{ij}^{\frac{\eta+1}{\eta}}\right]^{\frac{\eta}{\eta+1}} .
$$

Along with (B2), these indexes imply that $\mathbf{WN} = Y$. Using these definitions along with $\mathbf{WN} = Y$ in (B2) yields the CES supply curve:

$$
n_{ij} = \left(\frac{w_{ij}}{\mathbf{W}}\right)^{\eta} \mathbf{N}.
$$

We therefore have the result that the supply curves that face firms in this model of individual discrete choice are equivalent to those that face the firms when a representative household solves the following income maximization problem:

$$
\max_{\{n_{ij}\}} \sum_{i \in j} w_{ij} n_{ij} \quad s.t. \quad \left[\sum_{i \in j} n_{ij}^{\frac{\eta+1}{\eta}}\right]^{\frac{\eta}{\eta+1}} = \mathbf{N}.
$$

Since at the solution, the objective function is equal to $\mathbf{WN}$, then the envelope condition delivers a natural interpretation of $\mathbf{W}$ as the equilibrium payment to total labor input in the economy for one additional unit of aggregate labor disutility. That is, the following equalities hold:

$$
\frac{\partial}{\partial \mathbf{N}} \sum_{i \in j} w_{ij} n_{ij}^{*}(w_{ij}, w_{-ij}) = \lambda = \mathbf{W} = \frac{\partial}{\partial \mathbf{N}} \mathbf{WN}.
$$

Nested logit and nested CES. Consider changing the distribution of preference shocks as follows:

$$
F(\varepsilon_{i1}, ..., \varepsilon_{NJ}) = \exp\left[-\sum_{j=1}^{J}\left(\sum_{i=1}^{M_j} e^{-(1+\eta)\varepsilon_{ij}}\right)^{\frac{1+\theta}{1+\eta}}\right].
$$

We recover the distribution (B1) above if $\eta = \theta$. Otherwise, if $\eta > \theta$ the problem is convex and the conditional covariance of within sector preference draws differ from the economy wide variance of preference draws. We discuss this more below.

In this setting, choice probabilities can be expressed as the product of the conditional choice probability of supplying labor to firm i conditional on supplying labor to market j, and the probability of supplying labor to market j:

$$Prob_l\left(w_{ij}, w_{-ij}\right) = \underbrace{\frac{w_{ij}^{1+\eta}}{\sum_{i=1}^{M_j} w_{ij}^{1+\eta}}}_{Prob_l(\text{Choose firm } i \,|\, \text{Choose market } j)} \times \underbrace{\frac{\left[\sum_{i=1}^{M_j} w_{ij}^{1+\eta}\right]^{\frac{1+\theta}{1+\eta}}}{\sum_{l=1}^{J}\left[\sum_{k=1}^{M_l} w_{kl}^{1+\eta}\right]^{\frac{1+\theta}{1+\eta}}}}_{Prob_l(\text{Choose market } j)}.$$

Following the same steps as above, we can aggregate these choice probabilities and hours decisions to obtain firm level labor supply:

$$n_{ij} = \frac{w_{ij}^{\eta}}{\sum_{i=1}^{M_j} w_{ij}^{1+\eta}} \frac{\left[\sum_{i=1}^{M_j} w_{ij}^{1+\eta}\right]^{\frac{1+\theta}{1+\eta}}}{\sum_{l=1}^{J}\left[\sum_{k=1}^{M_l} w_{kl}^{1+\eta}\right]^{\frac{1+\theta}{1+\eta}}} Y. \tag{B3}$$

We can now define the following indexes:

$$\mathbf{W}_j = \left[\sum_{i=1}^{M_j} w_{ij}^{1+\eta}\right]^{\frac{1}{1+\eta}} \quad , \quad \mathbf{N}_j = \left[\sum_{i=1}^{M_j} n_{ij}^{\frac{1+\eta}{\eta}}\right]^{\frac{\eta}{1+\eta}},$$
$$\mathbf{W} = \left[\sum_{j=1}^{J} \mathbf{W}_j^{1+\theta}\right]^{\frac{1}{1+\theta}} \quad , \quad \mathbf{N} = \left[\sum_{j=1}^{J} \mathbf{N}_j^{\frac{1+\theta}{\theta}}\right]^{\frac{\theta}{1+\theta}}.$$

Using these definitions and similar results to the above we can show that $\mathbf{W}_j\mathbf{N}_j = \sum_{i=1}^{M_j} w_{ij}n_{ij}$, and $Y = \mathbf{W}\mathbf{N} = \sum_{j=1}^{J} \mathbf{W}_j\mathbf{N}_j$.

Consider the thought experiment of adding more markets J (which is necessary to identically map these formulas to our model). While the min of an infinite number of draws from a Gumbel distribution is not defined (it asymptotes to $-\infty$), the distribution of choices across markets is defined at each point in the limit as we add more markets J (Malmberg (2013)). As a result, the distribution of choices will have a well defined limit, and with the correct scaling as we add more markets (we can scale the disutilities at each step and not affect the market choice), as described in (Malmberg (2013)), the limiting wage indexes will be defined as above. We can then express (B3) as:

$$n_{ij} = \left(\frac{w_{ij}}{\mathbf{W}_j}\right)^{\eta} \left(\frac{\mathbf{W}_j}{\mathbf{W}}\right)^{\theta} \mathbf{N},$$

which completes the CES supply system defined in the text.

Comment. The above has established that it is straightforward to derive the supply system in the model through a discrete choice framework. This is particularly appealing given recent modeling of labor supply using familiar discrete choice frameworks first in models of economic geography and more recently in labor (Borovickova and Shimer (2017), Card, Cardoso, Heining, and Kline (2018), Lamadon, Mogstad, and Setzler (2019)). Since firms take this supply system as given, we can then work with the nested CES supply functions as if they were derived from the preferences and decisions of a representative household. This vastly simplifies welfare computations and allows for the integration of the model into more familiar macroeconomic environments.

A second advantage of this micro-foundation is that it provides a natural interpretation of the somewhat nebulous elasticities of substitution in the CES specification: η and θ. Returning to the Gumbel distribution we observe the following

$$F(\varepsilon_{i1}, ..., \varepsilon_{NJ}) = \exp\left[-\sum_{j=1}^{J}\left(\sum_{i=1}^{M_j} e^{-(1+\eta)\varepsilon_{ij}}\right)^{\frac{1+\theta}{1+\eta}}\right]$$

A higher value of η *increases* the correlation of draws within a market (McFadden, 1978). Within a market if η is high, then an individual's preference draws are likely to be clustered. With little difference in non-pecuniary idiosyncratic preferences for working at different firms, wages dominate in an individual's labor supply decision and wage posting in the market is closer to the competitive outcome. A higher value of θ *decreases* the overall variance of draws across all firms (i.e. it *increases* the correlation across any two randomly chosen sub-vectors of an individual's draws). An individual is therefore more likely to find that their lowest levels of idiosyncratic disutility are in two different markets, increasing across market wage competition.

In the case that $\eta = \theta$, the model collapses to the standard logit model. In this case the following obtains. Take an individual's ε_{lij} for some firm. The conditional probability distribution of some other draw $\varepsilon_{li'j'}$ is the same whether firm i' is in the same market $(j' = j)$ or some other market $(j' \neq j)$. Individuals are as likely to find somewhere local that incurs the same level of labor disability as finding somewhere in another market. In this setting economy-wide monopsonistic competition obtains. When an individual is more likely to find their other low disutility draws in the *same* market, then firms within that market have local market power. This is precisely the case that obtains when $\eta > \theta$.

B.2 Dynamic discrete choice framework

We show that the above discrete choice framework can be adapted to an environment where some individuals draw new vectors ε_l each period and reoptimize their labor supply. Firms therefore compete in a dynamic oligopoly. Restricting attention to the stationary solution to the model where firms keep employment and wages constant—as in the tradition of Burdett and Mortensen (1998)—we show that the allocation of employment and wages once again coincide with the solution to the problem in the main text. To simplify notation we consider the problem for a market with M firms $i \in \{1, \ldots, M\}$ which may be generalized to the model in the text.

Environment. Every period a random fraction λ of workers each draw a new vector ε_l. Let n_i be employment at firm i. Let $\overline{w}_i$ be the *average wage* of workers at firm i, such that the total wage bill in the firm is $\overline{w}_i n_i$. Let the equilibrium labor supply function $h(w_i, w_{-i})$ determine the amount of hires a firm makes if it posts a wage w_i when its competitors wages in the market are given by the vector w_{-i}.

Value function. Let $V(n_i, \overline{w}_i)$ be the firm's present discounted value of profits, where the firm has discount rate $\beta = 1$. Then $V(n_i, \overline{w}_i)$ satisfies:

$$\begin{aligned} V\left(n_i, \overline{w}_i\right) &= \left(Pz_i - \overline{w}_i\right)\left(1-\lambda\right) n_i + \max_{w_i'} \left\{ \left(Pz_i - w_i'\right) h\left(w_i', w_{-i}'\right) + V\left(n_i', \overline{w}_i'\right) \right\} \quad , && \text{(B4)} \\ n'\left(n_i, w_i', w_{-i}'\right) &= \left(1-\lambda\right) n_i + h\left(w_i', w_{-i}'\right) \quad , && \text{(B5)} \\ \overline{w}'\left(n_i, \overline{w}_i, w_i', w_{-i}'\right) &= \frac{\left(1-\lambda\right) \overline{w}_i n_i + h\left(w_i', w_{-i}'\right) w_i'}{\left(1-\lambda\right) n_i + h\left(w_i', w_{-i}'\right)}. && \text{(B6)} \end{aligned}$$

The firm operates a constant returns to scale production function. Of the firm's n_i workers, a fraction $(1-\lambda)$ do not draw new preferences. The total profit associated with these workers is then average revenue (Pz_i) minus average cost $(\overline{w}_i)$. The firm chooses a new wage w_i' to post in the market. In equilibrium, given its competitor's wages w_{-i}', it hires $h(w_i, w_{-i})$ workers. The total profit associated with these workers is again average revenue (Pz_i) minus average cost (w_i'). The second and third equations account for the evolution of the firm's state variables.

Optimality. Given its competitor's prices, the first order condition with respect to w_i' is:

$$\left(Pz_i - w_i'\right) h_1\left(w_i', w_{-i}'\right) - h\left(w_i', w_{-i}'\right) + V_n\left(n_i', \overline{w}_i'\right) n_w'\left(n_i, w_i', w_{-i}'\right) + V_{\overline{w}}\left(n_i', \overline{w}_i'\right) \overline{w}_w\left(n_i, \overline{w}_i, w_i', w_{-i}'\right) = 0$$

The relevant envelope conditions are

$$\begin{aligned} V_n\left(n_i, \overline{w}_i\right) &= \left(Pz_i - \overline{w}_i\right)\left(1-\lambda\right) + V_n\left(n_i', \overline{w}_i'\right) n_n'\left(n_i, w_i', \mathbf{w}_{-i}'\right) + V_{\overline{w}}\left(n_i', \overline{w}_i'\right) \overline{w}_n'\left(n_i, \overline{w}_i, w_i', \mathbf{w}_{-i}'\right) \\ V_{\overline{w}}\left(n_i, \overline{w}_i\right) &= -\left(1-\lambda\right) n_i + V_{\overline{w}}\left(n_i', \overline{w}_i'\right) \overline{w}_{\overline{w}}'\left(n_i, \overline{w}_i, w_i', \mathbf{w}_{-i}'\right) \end{aligned}$$

In a stationary equilibrium $\overline{w}_i = w_i'$, and $n_i' = n_i$. One can compute the partial derivatives involved in these expressions, and evaluate the conditions under stationarity to obtain

$$\left(Pz_i - w_i\right) h_1\left(w_i, w_{-i}\right) = h\left(w_i, w_{-i}\right).$$

Rearranging this expression:

$$w_i = \frac{\varepsilon_i(w_i, w_{-i})}{\varepsilon_i(w_i, w_{-i}) + 1} P z_i \quad , \quad \varepsilon_i(w_i, w_{-i}) := \frac{h_1(w_i, w_{-i}) w_i}{h(w_i, w_{-i})}$$

The solution to the dynamic oligopsony problem for a *given* supply system is identical to the solution of the static problem. In this setting, the supply system is obviously that which is obtained from the individual discrete choice problem in the previous section.

Comments. This setting establishes that the model considered in the main text can also be conceived as a setting where individuals periodically receive some preference shock that causes them to relocate, and firms engage in a dynamic oligopoly given these worker decisions. When $\eta > \theta$ the shock causes a worker to consider all firms in one market very carefully to the exclusion of other markets when they are making their relocation decision. When $\eta = \theta$ the individual considers all firms in all markets equally.

C Data

This section provides additional details regarding the data sources used in the paper, sample restrictions, and construction of a number of variables.

C.1 Census Longitudinal Business Database (LBD)

The LBD is built on the Business Register (BR), Economic Census and surveys. The BR began in 1972 and is a database of all U.S. business establishments. The business register is also called the Standard Statistical Establishment List (SSEL). The SSEL contains records for all industries except private households and illegal or underground activities. Most government owner entities are not in the SSEL. The SSEL includes single and multi unit establishments. The longitudinal links are constructed using the SSEL. The database is annual.

C.2 Sample restrictions

For both the summary statistics and corporate tax analysis, we isolate all plants (lbdnums) with non missing firmids, with strictly positive pay, strictly positive employment, non-missing county codes for the continental US (we exclude Alaska, Hawaii, and Puerto Rico). We then isolate all lbdnums with non-missing 2 digit NAICS codes equal to 11,21,31,32,33, or 55. We use the consistent 2007 NAICS codes provided by Fort and Klimek (2016) throughout the paper. These are the top tradeable 2-digit NAICS codes as defined by Delgado, Bryden, and Zyontz (2014). We winsorize the relative wage at the 1% level to remove outliers. Each plant has a unique firmid which corresponds to the owner of the plant.[64] Throughout the paper, we define a firm to be the sum of all establishments in a commuting zone with a common firmid and NAICS3 classification.

Summary Statistics Sample: Our summary statistics include all observations that satisfy the above criteria in 1976 and 2014.

Corporate Tax Sample: The corporate tax analysis includes all observations that satisfy the above criteria between 2002 and 2012 (note the tax series ends in 2012, but the 'Year t+1' estimates use 2013 observations). We further restrict the sample to firmid-market-year observations which have a 'Corporation' legal form of organization. The legal form of organization changes discontinuously in 2001 and earlier years, and thus we restrict our analysis to post-2002 observations. We must further restrict our attention to corporations that operate in at least two markets, since we use variation across markets, within a state, in order to isolate the impact of the corporate tax shocks on employment and wages.

Sample NAICS Codes and Commuting Zones: Table C1 describes the NAICS 3 codes in our sample. Table C2 provides examples of commuting zones and the counties that are associated with those commuting zones.

[64]Each firm only has one firmid. The firmid is different from the EIN. The firmid aggregates EINS to build a consistent firm identifier in the cross-section and over time.

Table C1: NAICS 3 digit examples

NAICS3	Description	NAICS3	Description
111	Crop Production	322	Paper Manuf.
112	Animal Production and Aquaculture	323	Printing and Related Support Activities
113	Forestry and Logging	324	Petroleum and Coal Products Manuf.
114	Fishing, Hunting and Trapping	325	Chemical Manuf.
115	Support Activities for Agriculture and Forestry	326	Plastics and Rubber Products Manuf.
211	Oil and Gas Extraction	327	Nonmetallic Mineral Product Manuf.
212	Mining (except Oil and Gas)	331	Primary Metal Manuf.
213	Support Activities for Mining	332	Fabricated Metal Product Manuf.
311	Food Manuf.	333	Machinery Manuf.
312	Beverage and Tobacco Product Manuf.	334	Computer and Electronic Product Manuf.
313	Textile Mills	335	Electrical Equipment, Appliance, Component Manuf.
314	Textile Product Mills	336	Transportation Equipment Manuf.
315	Apparel Manufacturing	337	Furniture and Related Product Manuf.
316	Leather and Allied Product Manuf.	339	Miscellaneous Manuf.
321	Wood Product Manuf.	551	Management of Companies and Enterprises

Table C2: Commuting Zone (CZ) examples: Census commuting zones numbers 58 and 47

CZ ID, 2000	County Name	Metro. Area, 2003	County Pop. 2000	CZ Pop. 2000
58	Cook County	Chicago-Naperville-Joliet, IL Metro. Division	5,376,741	8,704,935
58	DeKalb County	Chicago-Naperville-Joliet, IL Metro. Division	88,969	8,704,935
58	DuPage County	Chicago-Naperville-Joliet, IL Metro. Division	904,161	8,704,935
58	Grundy County	Chicago-Naperville-Joliet, IL Metro. Division	37,535	8,704,935
58	Kane County	Chicago-Naperville-Joliet, IL Metro. Division	404,119	8,704,935
58	Kendall County	Chicago-Naperville-Joliet, IL Metro. Division	54,544	8,704,935
58	Lake County	Lake County-Kenosha County, IL-WI Metro. Division	644,356	8,704,935
58	McHenry County	Chicago-Naperville-Joliet, IL Metro. Division	260,077	8,704,935
58	Will County	Chicago-Naperville-Joliet, IL Metro. Division	502,266	8,704,935
58	Kenosha County	Lake County-Kenosha County, IL-WI Metro. Division	149,577	8,704,935
58	Racine County	Racine, WI MSA	188,831	8,704,935
58	Walworth County	Whitewater, WI Micropolitan SA	93,759	8,704,935
47	Anoka County	Minneapolis-St. Paul-Bloomington, MN-WI MSA	298,084	2,904,389
47	Carver County	Minneapolis-St. Paul-Bloomington, MN-WI MSA	70,205	2,904,389
47	Chisago County	Minneapolis-St. Paul-Bloomington, MN-WI MSA	41,101	2,904,389
47	Dakota County	Minneapolis-St. Paul-Bloomington, MN-WI MSA	355,904	2,904,389
47	Hennepin County	Minneapolis-St. Paul-Bloomington, MN-WI MSA	1,116,200	2,904,389
47	Isanti County	Minneapolis-St. Paul-Bloomington, MN-WI MSA	31,287	2,904,389
47	Ramsey County	Minneapolis-St. Paul-Bloomington, MN-WI MSA	511,035	2,904,389
47	Scott County	Minneapolis-St. Paul-Bloomington, MN-WI MSA	89,498	2,904,389
47	Washington County	Minneapolis-St. Paul-Bloomington, MN-WI MSA	201,130	2,904,389
47	Wright County	Minneapolis-St. Paul-Bloomington, MN-WI MSA	89,986	2,904,389
47	Pierce County	Minneapolis-St. Paul-Bloomington, MN-WI MSA	36,804	2,904,389
47	St. Croix County	Minneapolis-St. Paul-Bloomington, MN-WI MSA	63,155	2,904,389

D Labor market concentration in all industries

Table D1 includes summary statistics of labor market concentration across all industries. Similar to tradeable industries, the market-level unweighted and weighted Herfindahls decline. The unweighted wage-bill Herfindahl declines from 0.36 to 0.34. The payroll weighted wage-bill Herfindahl declines from 0.17 to 0.11. The payroll weighted employment Herfindahl declines from 0.15 to 0.09. Similar to tradeable industries, Herfindahls are negatively correlated with the number of firms as well as total employment in the market.

	(A) Firm-market-level averages	
	1976	**2014**
Total firm pay (000s)	209.40	1102.00
Total firm employment	19.43	23.21
Pay per employee	$ 10,777	$ 47,480
Firm-Market level observations	3,746,000	5,854,000

	(B) Market-level averages	
	1976	**2014**
Wage-bill Herfindahl (Unweighted)	0.36	0.34
Employment Herfindahl (Unweighted)	0.33	0.32
Wage-bill Herfindahl (Weighted by market's share of total wage-bill)	0.17	0.11
Employment Herfindahl (Weighted by market's share of total wage-bill)	0.15	0.09
Firms per market	75.70	113.20
Percent of markets with 1 firm	10.4%	9.4%
Market level observations	49,000	52,000

	(C) Market-level correlations	
	1976	**2014**
Correlation of Wage-bill Herfindahl and number of firms	-0.20	-0.17
Correlation of Wage-bill Herfindahl and Employment Herfindahl	0.97	0.97
Correlation of Wage-bill Herfindahl and Market Employment	-0.15	-0.16
Market-level observations	49,000	52,000

Table D1: Summary Statistics, Longitudinal Employer Database 1976 and 2014

Notes: All NAICS. Market defined to be NAICS3 within Commuting Zone. Observations rounded to nearest thousand and numbers rounded to 4 significant digits according to Census disclosure rules. Firm-market-level refers to a 'firmid by Commuting Zone by 3-digit NAICs by Year' observation. Market-level refers to a 'Commuting Zone by 3-digit NAICs by Year' aggregation of observations.

E Mathematical derivations

This section details derivation of mathematical formulae appearing in the main text. It covers: (i) the household problem, (ii) sectoral equilibria of the firm problem, (iii) the labor share, (iv) wage pass-through results.

E.1 Household problem derivations

We solve for demand of the final good by taking the first order condition of the household problem with respect to C_t

$$\beta^t u'\left(\mathbf{C}_t - \frac{1}{\overline{\varphi}^{\frac{1}{\varphi}}}\frac{\mathbf{N}_t^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}\right) = \Lambda_t$$

The optimality condition for capital yields:

$$\Lambda_t = \Lambda_{t+1}\big(R_{t+1} + (1-\delta)\big).$$

To determine labor supply, we proceed with a three-step budgeting problem. Consider the first stage. Suppose the household must earn S_t by choosing labor supply across markets:

$$\mathbf{N}_t = \min_{\{\mathbf{N}_{jt}\}} \left[\int_0^1 \mathbf{N}_{jt}^{\frac{\theta+1}{\theta}} dj\right]^{\frac{\theta}{\theta+1}} \quad \text{s.t} \int_0^1 \mathbf{W}_{jt}\mathbf{N}_{jt} dj \geq S_t$$

The $FOC\left(\mathbf{N}_{jt}\right)$ is[65]

$$\begin{aligned}
\mathbf{N}_t^{-\frac{1}{\theta}}\mathbf{N}_{jt}^{\frac{1}{\theta}} &= \lambda \mathbf{W}_{jt} \\
\mathbf{N}_t^{-\frac{1}{\theta}}\left[\int_0^1 \mathbf{N}_{jt}^{\frac{\theta+1}{\theta}} dj\right] &= \lambda \int_0^1 \mathbf{W}_{jt}\mathbf{N}_{jt} dj \\
\mathbf{N}_t &= \lambda \int_0^1 \mathbf{W}_{jt}\mathbf{N}_{jt} dj
\end{aligned}$$

then define $\mathbf{W}_t$ by the number that satisfies $\mathbf{W}_t\mathbf{N}_t = \int_0^1 \mathbf{W}_{jt}\mathbf{N}_{jt} dj$, which implies that $\lambda = \mathbf{W}_t^{-1}$. Using the wage index in the first-order condition, we obtain:

$$\begin{aligned}
\mathbf{N}_t^{-\frac{1}{\theta}}\mathbf{N}_{jt}^{\frac{1}{\theta}} &= \lambda \mathbf{W}_{jt} \\
\mathbf{N}_{jt} &= \left(\frac{\mathbf{W}_{jt}}{\mathbf{W}_t}\right)^{\theta}\mathbf{N}_t
\end{aligned} \tag{E1}$$

[65] Where we have used $\left[\int_0^1 \mathbf{N}_{jt}^{\frac{\theta+1}{\theta}} dj\right]^{\frac{\theta}{\theta+1}-1} = \left[\int_0^1 \mathbf{N}_{jt}^{\frac{\theta+1}{\theta}} dj\right]^{-\frac{1}{\theta+1}} = \mathbf{N}_t^{-\frac{1}{\theta}}$

We then recover the wage index by multiplying (E1) by $\mathbf{W}_{jt}$ and integrating across markets:

$$\begin{aligned}
\mathbf{W}_{jt}\mathbf{N}_{jt} &= \mathbf{W}_{jt}^{1+\theta}\mathbf{W}_t^{-\theta}\mathbf{N}_t \\
\int_0^1 \mathbf{W}_{jt}\mathbf{N}_{jt}dj &= \int_0^1 \mathbf{W}_{jt}^{1+\theta}dj\mathbf{W}_t^{-\theta}\mathbf{N}_t \\
\mathbf{W}_t\mathbf{N}_t &= \int_0^1 \mathbf{W}_{jt}^{1+\theta}dj\mathbf{W}_t^{-\theta}\mathbf{N}_t \\
\mathbf{W}_t &= \left[\int_0^1 \mathbf{W}_{jt}^{1+\theta}dj\right]^{\frac{1}{1+\theta}}
\end{aligned}$$

Moving to the second stage, suppose that a household must raise resources S_t within a market and chooses labor supply to each firm within that market:

$$\mathbf{N}_{jt} = \min_{\{n_{ijt}\}} \left(\sum_{i=1}^{M} n_{ijt}^{\frac{\eta+1}{\eta}}\right)^{\frac{\eta}{\eta+1}} \quad s.t. \quad \sum_{i=1}^{M} w_{ijt}n_{ijt} \geq S_t$$

Let $\mathbf{W}_{jt}$ be the number such that $\mathbf{W}_{jt}\mathbf{N}_{jt} = \sum_i w_{ijt}n_{ijt}$. Taking first order conditions and proceeding similarly to the first stage we obtain the following:

$$\begin{aligned}
n_{ijt} &= \left(\frac{w_{ijt}}{\mathbf{W}_{jt}}\right)^{\eta}\mathbf{N}_{jt} \qquad \text{(E2)} \\
\mathbf{W}_{jt} &= \left[\sum_i w_{ijt}^{1+\eta}\right]^{\frac{1}{1+\eta}}
\end{aligned}$$

Moving to the third stage, we recast the original problem and take first order conditions for $\mathbf{N}_t$:

$$U = \max_{\{\mathbf{N}_t, C_t, K_t\}} \sum_{t=0}^{\infty} \beta^t u\left(C_t - \frac{1}{\overline{\varphi}^{\frac{1}{\varphi}}}\frac{\mathbf{N}_t^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}\right)$$

subject to the household's budget constraint which is given by,

$$C_t + \left[K_{t+1} - (1-\delta)K_t\right] = \mathbf{N}_t\mathbf{W}_t + R_tK_t + \Pi_t.$$

This yields the following expression for the aggregate labor supply index:

$$\mathbf{N}_t = \overline{\varphi}\mathbf{W}_t^{\varphi} \qquad \text{(E3)}$$

Substituting (E1) and (E3) into equation (E2), we derive the labor supply curve in the main text:

$$
\begin{aligned}
n_{ijt} &= \overline{\varphi} \left(\frac{w_{ijt}}{\mathbf{W}_{jt}} \right)^{\eta} \left(\frac{\mathbf{W}_{jt}}{\mathbf{W}_{t}} \right)^{\theta} \mathbf{W}_t^{\varphi} \\
\mathbf{W}_{jt} &= \left[\sum_i w_{ijt}^{1+\eta} \right]^{\frac{1}{1+\eta}} \\
\mathbf{W}_{t} &= \left[\int_0^1 \mathbf{W}_{jt}^{1+\theta} dj \right]^{\frac{1}{1+\theta}}
\end{aligned}
$$

To obtain the inverse labor supply curve, we use the first order conditions for labor supply within the market:

$$
n_{ijt} = \left(\frac{w_{ijt}}{\mathbf{W}_{jt}} \right)^{\eta} \mathbf{N}_{jt}
$$

Inverting this equation yields,

$$
\mathbf{w}_{ijt} = \left(\frac{\mathbf{n}_{ijt}}{\mathbf{N}_{jt}} \right)^{1/\eta} \mathbf{W}_{jt} \tag{E4}
$$

Labor supply across markets is given by the following expression:

$$
\mathbf{N}_{jt} = \left(\frac{\mathbf{W}_{jt}}{\mathbf{W}_{t}} \right)^{\theta} \mathbf{N}_t
$$

Inverting this equation yields,

$$
\mathbf{W}_{jt} = \left(\frac{\mathbf{N}_{jt}}{\mathbf{N}_{t}} \right)^{1/\theta} \mathbf{W_t} \tag{E5}
$$

Combining (E5), (E4) and (E3) yields the expression in the text.

E.2 Derivation of firm problem under Cournot competition

Let $y_{ijt} = \overline{Z} z_{ijt} \left(k_{ijt}^{1-\gamma} n_{ijt}^{\gamma} \right)^{\alpha}$. The firm problem with capital and decreasing returns to scale is given by,

$$
\max_{k_{ijt}, n_{ijt}} \overline{Z} z_{ijt} \left(k_{ijt}^{1-\gamma} n_{ijt}^{\gamma} \right)^{\alpha} - R_t k_{ijt} - w_{ijt} n_{ijt}
$$

subject to the household labor supply curve. Taking first order conditions for k_{ijt} yields $\frac{R_t k_{ijt}}{y_{ijt}} = (1-\gamma)\,\alpha$. We substitute this expression into the profit function

$$
\max_{k_{ijt}, n_{ijt}} \left[1 - (1-\gamma)\,\alpha \right] y_{ijt} - w_{ijt} n_{ijt}
$$

We solve for capital using the first order condition for capital (again):[66]

$$k_{ijt} = \left(\frac{(1-\gamma)\,\alpha z_{ijt}\overline{Z}}{R_t}\right)^{\frac{1}{1-(1-\gamma)\alpha}} n_{ijt}^{\frac{\gamma\alpha}{1-(1-\gamma)\alpha}}$$

We substitute this into the expression for y_{ijt} to obtain firm-level output as a function of n_{ijt}:

$$y_{ijt} = \left(\frac{(1-\gamma)\,\alpha}{R_t}\right)^{\frac{(1-\gamma)\alpha}{1-(1-\gamma)\alpha}} \left(z_{ijt}\overline{Z}\right)^{\frac{1}{1-(1-\gamma)\alpha}} n_{ijt}^{\frac{\gamma\alpha}{1-(1-\gamma)\alpha}}$$

The firm profit function becomes:

$$\pi_{ijt} = \left[1-(1-\gamma)\,\alpha\right]\left(\frac{(1-\gamma)\,\alpha}{R_t}\right)^{\frac{(1-\gamma)\alpha}{1-(1-\gamma)\alpha}} \left(z_{ijt}\overline{Z}\right)^{\frac{1}{1-(1-\gamma)\alpha}} n_{ijt}^{\frac{\gamma\alpha}{1-(1-\gamma)\alpha}} - w_{ijt}n_{ijt}$$

Defining $\widetilde{\alpha} := \frac{\gamma\alpha}{1-(1-\gamma)\alpha}$, $\widetilde{z}_{ijt} := \left[1-(1-\gamma)\,\alpha\right]\left(\frac{(1-\gamma)\alpha}{R_t}\right)^{\frac{(1-\gamma)\alpha}{1-(1-\gamma)\alpha}} z_{ijt}^{\frac{1}{1-(1-\gamma)\alpha}}$, and $\widetilde{Z} := \overline{Z}^{\frac{1}{1-(1-\gamma)\alpha}}$ yields the firm profit maximization problem, expression (2.3), in the text.

Define $MRPL_{ijt} = \widetilde{\alpha}\widetilde{Z}\widetilde{z}_{ijt}n_{ijt}^{\widetilde{\alpha}-1}$. Define $X_t = \frac{1}{\overline{\varphi}^{\frac{1}{\varphi}}}\mathbf{N}_t^{\frac{1}{\varphi}-1/\theta}$ and substitute this into the inverse labor supply function to derive the following expression:

$$w_{ijt} = n_{ijt}^{1/\eta}\mathbf{N}_{jt}^{1/\theta-1/\eta}X_t \tag{E6}$$

We substitute this expression into the profit function to obtain,

$$\pi_{ijt} = \max_{n_{ijt}} \widetilde{Z}\widetilde{z}_{ijt}n_{ijt}^{\widetilde{\alpha}} - n_{ijt}^{\frac{1}{\eta}+1} n_{jt}^{\frac{1}{\theta}-\frac{1}{\eta}} X_t$$

Before taking first order conditions, we derive a useful result, $\frac{\partial \mathbf{N}_{jt}}{\partial n_{ijt}}\frac{n_{ijt}}{\mathbf{N}_{jt}} = s_{ijt}^{wn}$.

Lemma E.1. $\frac{\partial n_{jt}}{\partial n_{ijt}}\frac{n_{ijt}}{n_{jt}} = s_{ijt}^{wn}$

Proof: Using the definition of $\mathbf{N}_{jt} = \left[\sum_i n_{ijt}^{\frac{\eta+1}{\eta}}\right]^{\frac{\eta}{\eta+1}}$ and taking first order conditions yields:

$$\begin{aligned}\frac{\partial \mathbf{N}_{jt}}{\partial n_{ijt}} &= \left[\sum_i n_{ijt}^{\frac{\eta+1}{\eta}}\right]^{\frac{\eta}{\eta+1}-1} n_{ijt}^{\frac{\eta+1}{\eta}-1} \\ &= \mathbf{N}_{jt}^{-\frac{1}{\eta}} n_{ijt}^{\frac{1}{\eta}}\end{aligned}$$

66

$(1-\alpha)\,\gamma z_i\left(k_i^{1-\alpha}n_i^{\alpha}\right)^{\gamma} = rk_i$ implies $k_i = \left(\frac{(1-\alpha)\gamma z_i}{r}\right)^{\frac{1}{1-(1-\alpha)\gamma}} n_i^{\frac{\alpha\gamma}{1-(1-\alpha)\gamma}}$

This yields the elasticity of market level labor supply:

$$\frac{\partial \mathbf{N}_{jt}}{\partial n_{ijt}} \frac{n_{ijt}}{\mathbf{N}_{jt}} = \left(\frac{n_{ijt}}{\mathbf{N}_{jt}}\right)^{\frac{\eta+1}{\eta}} \tag{E7}$$

Substituting (E7) into the definition of the wage-bill share:

$$s_{ijt}^{wn} = \frac{w_{ijt}n_{ijt}}{\sum_i w_{ijt}n_{ijt}} = \frac{n_{ijt}^{\frac{1}{\eta}+1}\mathbf{N}_{jt}^{\frac{1}{\theta}-\frac{1}{\eta}}X}{\sum_i n_{ijt}^{\frac{1}{\eta}+1}\mathbf{N}_{jt}^{\frac{1}{\theta}-\frac{1}{\eta}}X} = \frac{n_{ijt}^{\frac{1}{\eta}+1}}{\sum_i n_{ijt}^{\frac{1}{\eta}+1}} = \frac{n_{ijt}^{\frac{\eta+1}{\eta}}}{\left[\sum_i n_{ijt}^{\frac{\eta+1}{\eta}}\right]^{\frac{\eta}{\eta+1}\frac{\eta+1}{\eta}}} = \frac{n_{ijt}^{\frac{\eta+1}{\eta}}}{\mathbf{N}_{jt}^{\frac{\eta+1}{\eta}}} \implies s_{ijt}^{wn} = \frac{\partial \mathbf{N}_{jt}}{\partial n_{ijt}} \frac{n_{ijt}}{\mathbf{N}_{jt}}$$

□

Lemma E.2. *The equilibrium markdown μ_{ijt} is a wage bill share weighted harmonic mean of the monopsonistically competitive markup under η or θ.*

$$\begin{aligned} w_{ijt} &= \mu_{ijt} MRPL_{ijt} \\ \mu_{ijt} &= \frac{\varepsilon_{ijt}}{\varepsilon_{ijt}+1} \\ \varepsilon_{ijt} &= \left[\left(1 - s_{ijt}^{wn}\right)\frac{1}{\eta} + s_{ijt}^{wn}\frac{1}{\theta}\right]^{-1} \end{aligned} \tag{E8}$$

Proof: Using Lemma E.1, we take first-order conditions to derive the optimal employment decision:

$$\begin{aligned} 0 &= MRPL_{ijt} - \left(\frac{1}{\eta}+1\right)\left[n_{ijt}^{\frac{1}{\eta}}\mathbf{N}_{jt}^{\frac{1}{\theta}-\frac{1}{\eta}}X_t\right] - \left(\frac{1}{\theta}-\frac{1}{\eta}\right)\left[n_{ijt}^{\frac{1}{\eta}+1}\mathbf{N}_{jt}^{\frac{1}{\theta}-\frac{1}{\eta}}X_t\right]\frac{1}{\mathbf{N}_{jt}}\frac{\partial \mathbf{N}_{jt}}{\partial n_{ijt}} \\ MRPL_{ijt} &= \left[\frac{\eta+1}{\eta} + \left(\frac{\theta+1}{\theta} - \frac{\eta+1}{\eta}\right)s_{ijt}^{wn}\right]w_{ijt} \\ w_{ijt} &= \left[1 + \left(1 - s_{ijt}^{wn}\right)\frac{1}{\eta} + s_{ijt}^{wn}\frac{1}{\theta}\right]^{-1} MRPL_{ijt} \end{aligned}$$

□

E.3 Equilibrium properties - Labor Share

Using Lemma E.2, an individual firm's labor share, ls_{ij}, can be written in terms of the equilibrium markup:

$$\begin{aligned} ls_{ij} &= \frac{w_{ij}n_{ij}}{\widetilde{Z}\widetilde{z}_{ij}n_{ij}^{\widetilde{\alpha}}} \\ ls_{ij} &= (\alpha\gamma)\frac{w_{ij}}{MRPL_{ij}} \\ ls_{ij} &= (\alpha\gamma)\mu_{ij} \end{aligned}$$

Accounting for payments to capital, value added is $y_{ij} = (1/(1-(1-\gamma)\alpha))\widetilde{Z}\widetilde{z}_{ij}n_{ij}^{\widetilde{\alpha}}$. At the market level, the inverse labor share in market j, LS_j^{-1}, is given by the following expression:

$$LS_j^{-1} = \frac{\sum_i y_{ij}}{\sum_i w_{ij}n_{ij}} = \sum_i \left(\frac{w_{ij}n_{ij}}{\sum_i w_{ij}n_{ij}}\right)\frac{y_{ij}}{w_{ij}n_{ij}}$$

Using the definition of the wage-bill share,

$$\begin{aligned}
LS_j^{-1} &= \sum_i s_{ij}^{wn}\frac{y_{ij}}{w_{ij}n_{ij}} \\
LS_j^{-1} &= \sum_i s_{ij}^{wn}\frac{1}{1-(1-\gamma)\alpha}\frac{\widetilde{Z}\widetilde{z}_{ij}n_{ij}^{\widetilde{\alpha}}}{\left\{\mu_{ij}\tilde{\alpha}\widetilde{Z}\widetilde{z}_{ij}n_{ij}^{\widetilde{\alpha}-1}\right\}n_{ij}} \\
LS_j^{-1} &= \sum_i s_{ij}^{wn}(\alpha\gamma)^{-1}\mu_{ij}^{-1} \\
LS_j^{-1} &= (\alpha\gamma)^{-1}\sum_i s_{ij}^{wn}\left[\frac{\eta+1}{\eta} + s_{ij}^{wn}\left(\frac{\theta+1}{\theta} - \frac{\eta+1}{\eta}\right)\right] \\
LS_j^{-1} &= (\alpha\gamma)^{-1}\frac{\eta+1}{\eta} + (\alpha\gamma)^{-1}\left(\frac{\theta+1}{\theta} - \frac{\eta+1}{\eta}\right)HHI_j^{wn}
\end{aligned}$$

Aggregating across markets yields the economy-wide labor share:

$$LS^{-1} = \frac{\int\sum y_{ij}}{\int\sum w_{ij}n_{ij}} = \int\frac{\sum w_{ij}n_{ij}}{\int\sum w_{ij}n_{ij}}\frac{\sum y_{ij}}{\sum w_{ij}n_{ij}} = \int s_j^{wn}LS_j^{-1}dj$$

This yields the expression in the text:

$$LS^{-1} = \frac{1}{\alpha\gamma}\left(\frac{\eta+1}{\eta} + \left(\frac{\theta+1}{\theta} - \frac{\eta+1}{\eta}\right)\int s_j^{wn}HHI_j^{wn}dj\right) \tag{E9}$$

E.4 Representative firm problem

In this appendix, we derive the representative firm problem and prove *Proposition 1*. We proceed as follows: (1) we aggregate firm-level variables, (2) we define the representative firm's first order conditions, (3) we then show under what restrictions factor demands and output of the representative and heterogeneous firm problems coincide, and (4) we then consider an auxiliary economy with a continuum of identical symmetric firms and characterize the equilibrium set of equations and factor shares for the auxiliary economy, (5) we derive the conditions (stated in *Proposition 1*) under which the first order conditions in this auxiliary economy coincide with the representative firm economy, and thus the heterogeneous firm economy.

Step 1. Recall that optimal capital demand as a function of labor solves,

$$k^*\left(n_{ij}\right) = \arg\max\ \overline{Z}z_{ij}\left(k_{ij}^{1-\gamma}n_{ij}^{\gamma}\right)^{\alpha} - Rk_{ij}$$
$$k_{ij} = \left(\frac{(1-\gamma)\,\alpha z_{ij}\overline{Z}}{R}\right)^{\frac{1}{1-(1-\gamma)\alpha}} n_{ij}^{\frac{\gamma\alpha}{1-(1-\gamma)\alpha}}$$

Define ν_{ij} as the number which satisfies $n_{ij} = \nu_{ij}\mathbf{N}$. ν_{ij} is the contribution of n_{ij} to the aggregate labor *index*. Using this definition, labor supply is given by,

$$n_{ij} = \overline{\varphi}\left(\frac{w_{ij}}{\mathbf{W}_j}\right)^{\eta}\left(\frac{\mathbf{W}_j}{\mathbf{W}}\right)^{\theta}\mathbf{N}$$
$$n_{ij} = \nu_{ij}\mathbf{N}$$

Substituting capital demand into the definition of output $y_{ij} = \overline{Z}z_{ij}\left(k_{ij}^{1-\gamma}n_{ij}^{\gamma}\right)^{\alpha}$, and then using the definition of ν_{ij} yields,

$$y_{ij} = \left(z_{ij}\overline{Z}\right)^{\frac{1}{1-(1-\gamma)\alpha}}\left(\frac{(1-\gamma)\,\alpha}{R}\right)^{\frac{(1-\gamma)\alpha}{1-(1-\gamma)\alpha}}\left(\nu_{ij}\mathbf{N}\right)^{\frac{\gamma\alpha}{1-(1-\gamma)\alpha}}$$

Aggregating firm-level output implies that aggregate output is given by,

$$\mathbf{Y} = \overline{Z}^{\frac{1}{1-(1-\gamma)\alpha}}\left\{\int\sum_i\left(z_{ij}\nu_{ij}^{\gamma\alpha}\right)^{\frac{1}{1-(1-\gamma)\alpha}}\right\}\left(\frac{(1-\gamma)\,\alpha}{R}\right)^{\frac{(1-\gamma)\alpha}{1-(1-\gamma)\alpha}}\mathbf{N}^{\frac{\gamma\alpha}{1-(1-\gamma)\alpha}} \tag{E10}$$

Likewise, aggregating firm-level capital demand implies that aggregate capital is given by,

$$K = \overline{Z}^{\frac{1}{1-(1-\gamma)\alpha}}\left\{\int\sum_i\left(z_{ij}\nu_{ij}^{\gamma\alpha}\right)^{\frac{1}{1-(1-\gamma)\alpha}}\right\}\left(\frac{(1-\gamma)\,\alpha}{R}\right)^{\frac{1}{1-(1-\gamma)\alpha}}\mathbf{N}^{\frac{\gamma\alpha}{1-(1-\gamma)\alpha}} \tag{E11}$$

Step 2. We define the *aggregate production function* as follows:

$$\mathbf{Y} = \mathbf{Z}\left(K^{1-\gamma}\mathbf{N}^{\gamma}\right)^{\alpha}$$

Since the capital market is competitive, capital demand by the representative firm would be

$$K = \mathbf{Z}^{\frac{1}{1-(1-\gamma)\alpha}}\left(\frac{(1-\gamma)\,\alpha}{R}\right)^{\frac{1}{1-(1-\gamma)\alpha}}\mathbf{N}^{\frac{\gamma\alpha}{1-(1-\gamma)\alpha}} = \left[\mathbf{Z}\left(\frac{(1-\gamma)\,\alpha}{R}\right)\mathbf{N}^{\gamma\alpha}\right]^{\frac{1}{1-(1-\gamma)\alpha}} \tag{E12}$$

Substituting capital demand back into output, output of the representative firm would be,

$$\mathbf{Y} = \mathbf{Z}^{\frac{1}{1-(1-\gamma)\alpha}} \left(\frac{(1-\gamma)\alpha}{R} \right)^{\frac{(1-\gamma)\alpha}{1-(1-\gamma)\alpha}} \mathbf{N}^{\frac{\gamma\alpha}{1-(1-\gamma)\alpha}} = \left[\mathbf{Z} \left(\frac{(1-\gamma)\alpha}{R} \right)^{(1-\gamma)\alpha} \mathbf{N}^{\gamma\alpha} \right]^{\frac{1}{1-(1-\gamma)\alpha}} \tag{E13}$$

Step 3. This leads us to the first result. Comparing equations (E11) and (E10) with (E12) and (E13), capital demand and output are the same as derived from a representative firm when $\mathbf{Z} = \overline{Z} \left[\int \sum_i \left(z_{ij} \nu_{ij}^{\gamma\alpha} \right)^{\frac{1}{1-(1-\gamma)\alpha}} \right]^{1-(1-\gamma)\alpha}$.

Lemma E1: *The representative firm and heterogeneous firm problems coincide for the following aggregate production function,*

$$\mathbf{Y} = \mathbf{Z} \left(K^{1-\gamma} \mathbf{N}^{\gamma} \right)^{\alpha}$$

where the economy total factor productivity index TFP $\mathbf{Z}$

$$\mathbf{Z} = \overline{Z} \left[\int \sum_i \left(z_{ij} \nu_{ij}^{\gamma\alpha} \right)^{\frac{1}{1-(1-\gamma)\alpha}} \right]^{1-(1-\gamma)\alpha}$$

Step 4. Consider the following economy. Assume there is a continuum of identical firms indexed by ij, as before, with productivity $\mathbf{Z}$. Assume these firms are monopsonistically competitive, in that the across market and within market elasticities of labor substitution are equal and equal to $\mathcal{E}$. The representative household solves the following problem

$$\max_{\{\mathbf{N}_{ij}, K\}} u \left(\mathbf{C} - \overline{\varphi}^{-\frac{1}{\varphi}} \frac{\mathbf{N}^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}} \right) \quad \mathbf{N} = \left[\int_0^1 \sum_i \mathbf{N}_{ij}^{\frac{\mathcal{E}+1}{\mathcal{E}}} dj \right]^{\frac{\mathcal{E}}{\mathcal{E}+1}} \quad \mathbf{C} = \int_0^1 \sum_i c_{ij} dj$$

subject to

$$\mathbf{C} + I = \int \sum_i \mathbf{W}_{ij} \mathbf{N}_{ij} dj + \Pi$$

The solution to this problem is an aggregate labor supply function and labor supply to each firm:

$$\mathbf{N} = \overline{\varphi} \mathbf{W}^{\varphi}$$

$$\mathbf{W} = \left[\int_0^1 \sum_i \mathbf{W}_{ij}^{\mathcal{E}+1} dj \right]^{\frac{1}{\mathcal{E}+1}}$$

$$\mathbf{N}_{ij} = \left(\frac{\mathbf{W}_{ij}}{\mathbf{W}} \right)^{\mathcal{E}} \mathbf{N}$$

The firm has the production function $\mathbf{Y}_{ij} = \mathbf{Z} \left(K_{ij}^{1-\gamma} \mathbf{N}_{ij}^{\gamma} \right)^{\alpha}$ and chooses $\mathbf{N}_{ij}$ and K_{ij} to maximize profits subject to their labor supply curve:

$$\max_{\mathbf{N}_{ij},K_{ij}} \mathbf{Z}\left(K_{ij}^{1-\gamma}\mathbf{N}_{ij}^{\gamma}\right)^{\alpha} - \mathbf{W}_{ij}\mathbf{N}_{ij} - R\mathbf{K}_{ij} \quad s.t. \quad \mathbf{W}_{ij} = \overline{\varphi}^{-\frac{1}{\varphi}}\left(\frac{\mathbf{N}_{ij}}{\mathbf{N}}\right)^{1/\mathcal{E}}\mathbf{W}$$

The firm's optimality condition for capital is

$$K_{ij} = \left[\mathbf{Z}\left(\frac{(1-\gamma)\,\alpha}{R}\right)\mathbf{N}_{ij}^{\gamma\alpha}\right]^{\frac{1}{1-(1-\gamma)\alpha}}$$

This can be substituted back into the firm's maximization problem to yield

$$\max_{\mathbf{N}_{ij}} \mathbf{Z}^{\frac{1}{1-(1-\gamma)\alpha}}\left(\frac{(1-\gamma)\,\alpha}{R}\right)^{\frac{(1-\gamma)\alpha}{1-(1-\gamma)\alpha}}\mathbf{N}_{ij}^{\frac{\gamma\alpha}{1-(1-\gamma)\alpha}} - \mathbf{W}_{ij}\mathbf{N}_{ij}$$

We can group terms to simplify notation. Define $\widetilde{\mathbf{Z}} = \mathbf{Z}^{\frac{1}{1-(1-\gamma)\alpha}}\left(\frac{(1-\gamma)\alpha}{R}\right)^{\frac{(1-\gamma)\alpha}{1-(1-\gamma)\alpha}}$ and $\widetilde{\alpha} = \frac{\gamma\alpha}{1-(1-\gamma)\alpha}$ in order to rewrite the firm problem as follows:

$$\max_{\mathbf{N}_{ij},K_{ij}} \widetilde{\mathbf{Z}}\mathbf{N}_{ij}^{\widetilde{\alpha}} - \mathbf{W}_{ij}\mathbf{N}_{ij}$$

Define $\mathbf{MRPL}_{ij} = \widetilde{\alpha}\widetilde{\mathbf{Z}}\mathbf{N}_{ij}^{\widetilde{\alpha}-1}$. Then the firm's optimality condition for labor implies a wage $\mathbf{W}_{ij}$ that is a markdown, $\mu = \frac{\mathcal{E}}{\mathcal{E}+1}$, on the marginal revenue product of labor:

$$\mathbf{W}_{ij} = \mu\,\mathbf{MRPL}_{ij} \tag{E14}$$

Now we use the homogeneity of firms. Since all firms are identical, $\{\mathbf{W}_{ij}, \mathbf{N}_{ij}, K_{ij}\}$ will be common across all firms.[67] Using this symmetry and integrating the first order conditions (E14) over firms and markets yields $\boldsymbol{W} = \mu\boldsymbol{MRPL}$. The labor share in this auxiliary economy is $\mathbf{LS} = \frac{\mathbf{WN}}{\mathbf{Y}} = \mu\frac{\gamma\alpha}{1-(1-\gamma)\alpha}\left[1-(1-\gamma)\,\alpha\right]$ and the capital share is $\mathbf{KS} = \frac{RK}{\mathbf{Y}} = (1-\gamma)\,\alpha$.

Step 5. Let HA denote values of variables taken from our benchmark heterogeneous agent economy. For example, $\widetilde{IHI}_{HA}^{wn}$ denotes our benchmark heterogneoeus agent economy's inverse wage-bill Herfindahl. Equation (9) in the text defines the aggregate labor share in our benchmark economy, LS_{HA}. First, using the definition $\mu = \frac{\mathcal{E}}{\mathcal{E}+1}$ it is clear that the factor shares in the auxiliary and benchmark economies coincide, $\mathbf{LS} = LS_{HA}$, only when

$$\frac{\mathcal{E}}{\mathcal{E}+1} = \left[HHI_{HA}^{wn}\left(\frac{\theta}{\theta+1}\right)^{-1} + (1-HHI_{HA}^{wn})\left(\frac{\eta}{\eta+1}\right)^{-1}\right]^{-1}, \; HHI_{HA}^{wn} = \int_0^1 s_{j,HA}^{wn}HHI_{j,HA}^{wn}dj$$

Second, we showed in *Lemma E1*, that aggregating representative monopsonists' capital demand, labor

[67]Note $\mathbf{W} = M^{1/(\mathcal{E}+1)}\mathbf{W}_{ij}$, $\mathbf{N} = M^{\mathcal{E}/(\mathcal{E}+1)}\mathbf{N}_{ij}$, $K = MK_{ij}$ where $M = \int M_j dj$.

demand and output yield aggregate capital, labor disutility and output that coincide with the heterogeneous agent economy if

$$\mathbf{Z} = \overline{Z}\left[\int \sum_{i=1}^{M_j} \left(z_{ij}\nu_{ij}{}^{\gamma\alpha}\right)^{\frac{1}{1-(1-\gamma)\alpha}} dj\right]^{1-(1-\gamma)\alpha} \quad , \quad \text{where} \quad , \quad \nu_{ij} = \frac{n_{ij,HA}}{\mathbf{N}_{HA}}$$

We can therefore solve the HA economy, compute $(\mathcal{E}, \mathbf{Z})$ from data generated by that economy, and then solve for aggregate quantities using the RA economy treating $(\mathcal{E}, \mathbf{Z})$ as exogenous parameters.

□

F Non-Constant Returns to Scale Computation $\gamma \neq 1$

We solve the model by (i) guessing a vector of wage-bill shares, $\mathbf{s}_j^{wn} = (s_{1j}^{wn}, \ldots, s_{M_j j}^{wn})$, (ii) solving for firm-level markdowns, firm-level wages, and the sectoral wage index, and (iii) updating the wage-bill share using firm-level wages and the sectoral wage index.

From the main text, we define the marginal revenue product of labor as follows:

$$MRPL_{ij} = \widetilde{Z}\widetilde{z}_{ijt}\widetilde{\alpha} n_{ijt}^{\widetilde{\alpha}-1}$$

Substituting for n_{ijt} using the labor supply equation (6), and defining $\widehat{z}_{ij} = \widetilde{\alpha}\widetilde{z}_{ij}$ and $\omega = \frac{\widetilde{Z}}{\overline{\varphi}^{1-\widetilde{\alpha}}}$, then the marginal revenue product of labor can be written as:

$$MRPL_{ij} = \omega \mathbf{W}^{(1-\widetilde{\alpha})(\theta-\varphi)} \widehat{z}_{ij} \left\{ w_{ij}^{-\eta} \mathbf{W}_j^{\eta-\theta} \right\}^{1-\widetilde{\alpha}}$$

Use Lemma E.2 to write the wage in terms of the marginal revenue product of labor:

$$\begin{aligned} w_{ij} &= \mu_{ij} MRPL_{ij} \\ w_{ij} &= \mu_{ij} \omega \mathbf{W}^{(1-\widetilde{\alpha})(\theta-\varphi)} \widehat{z}_{ij} \left\{ w_{ij}^{-\eta} \mathbf{W}_j^{\eta-\theta} \right\}^{1-\widetilde{\alpha}} \end{aligned}$$

Use the fact that $\mathbf{W}_j = w_{ij} s_{ij}^{-\frac{1}{\eta+1}}$ to write this expression in terms of wage-bill shares, and then solve for w_{ij}. The resulting expression is given below:

$$w_{ij} = \omega^{\frac{1}{1+(1-\widetilde{\alpha})\theta}} \mathbf{W}^{\frac{(1-\widetilde{\alpha})(\theta-\varphi)}{1+(1-\widetilde{\alpha})\theta}} \mu_{ij}^{\frac{1}{1+(1-\widetilde{\alpha})\theta}} \widehat{z}_{ij}^{\frac{1}{1+(1-\widetilde{\alpha})\theta}} s_{ij}^{-\frac{(1-\widetilde{\alpha})(\eta-\theta)}{\eta+1}\frac{1}{1+(1-\widetilde{\alpha})\theta}}$$

We will solve for an equilibrium in 'hatted' variables, and then rescale the 'hatted' variables to recover the equilibrium values of n_{ij} and w_{ij}. Define the following 'hatted' variables:

$$\begin{aligned} \widehat{w}_{ij} &:= \mu_{ij}^{\frac{1}{1+(1-\widetilde{\alpha})\theta}} \widehat{z}_{ij}^{\frac{1}{1+(1-\widetilde{\alpha})\theta}} s_{ij}^{-\frac{(1-\widetilde{\alpha})(\eta-\theta)}{\eta+1}\frac{1}{1+(1-\widetilde{\alpha})\theta}} \\ \widehat{\mathbf{W}}_j &:= \left[\sum_{i\in j} \widehat{w}_{ij}^{\eta+1} \right]^{\frac{1}{\eta+1}} \\ \widehat{\mathbf{W}} &:= \left[\int \widehat{\mathbf{W}}_j^{\theta+1} dj \right]^{\frac{1}{\theta+1}} \\ \widehat{n}_{ij} &:= \left(\frac{\widehat{w}_{ij}}{\widehat{\mathbf{W}}_j} \right)^{\eta} \left(\frac{\widehat{\mathbf{W}}_j}{\widehat{\mathbf{W}}} \right)^{\theta} \left(\frac{\widehat{\mathbf{W}}}{1} \right)^{\varphi} \end{aligned}$$

These definitions imply that

$$
\begin{aligned}
w_{ij} &= \omega^{\frac{1}{1+(1-\tilde{\alpha})\theta}} \mathbf{W}^{\frac{(1-\tilde{\alpha})(\theta-\varphi)}{1+(1-\tilde{\alpha})\theta}} \widehat{w}_{ij} \\
\mathbf{W}_j &= \omega^{\frac{1}{1+(1-\tilde{\alpha})\theta}} \mathbf{W}^{\frac{(1-\tilde{\alpha})(\theta-\varphi)}{1+(1-\tilde{\alpha})\theta}} \widehat{\mathbf{W}}_j \\
\mathbf{W} &= \omega^{\frac{1}{1+(1-\tilde{\alpha})\theta}} \mathbf{W}^{\frac{(1-\tilde{\alpha})(\theta-\varphi)}{1+(1-\tilde{\alpha})\theta}} \widehat{\mathbf{W}}
\end{aligned}
$$

These definitions allow us to compute the equilibrium market shares in terms of 'hatted' variables:

$$
s_j^{wn} = \left(\frac{w_{ij}}{\mathbf{W}_j}\right)^{\eta+1} = \left(\frac{\widehat{w}_{ij}}{\widehat{\mathbf{W}}_j}\right)^{\eta+1} \tag{F1}
$$

For a given set of values for parameters $\{\widetilde{Z}, \overline{\varphi}, \widetilde{\alpha}, \beta, \delta\}$, we can solve for the non-constant returns to scale equilibrium as follows:

1. Guess $\mathbf{s}_j^{wn} = (s_{1j}^{wn}, \ldots, s_{M_j j}^{wn})$
2. Compute $\{\varepsilon_{ij}\}$ and $\{\mu_{ij}\}$ using the expressions in Lemma E.2.
3. Construct the 'hatted' equilibrium values as follows:

$$
\begin{aligned}
\widehat{w}_{ij} &= \mu_{ij}^{\frac{1}{1+(1-\tilde{\alpha})\theta}} \widehat{z}_{ij}^{\frac{1}{1+(1-\tilde{\alpha})\theta}} s_{ij}^{-\frac{(1-\tilde{\alpha})(\eta-\theta)}{\eta+1}\frac{1}{1+(1-\tilde{\alpha})\theta}} \\
\widehat{\mathbf{W}}_j &= \left[\sum_{i\in j} \widehat{w}_{ij}^{\eta+1}\right]^{\frac{1}{\eta+1}} \\
\widehat{\mathbf{W}} &= \left[\int \widehat{\mathbf{W}}_j^{\theta+1} dj\right]^{\frac{1}{\theta+1}} \\
\widehat{n}_{ij} &= \left(\frac{\widehat{w}_{ij}}{\widehat{\mathbf{W}}_j}\right)^{\eta} \left(\frac{\widehat{\mathbf{W}}_j}{\widehat{\mathbf{W}}}\right)^{\theta} \left(\frac{\widehat{\mathbf{W}}}{1}\right)^{\varphi}
\end{aligned}
$$

4. Update the wage-bill share vector using equation (F1).
5. Iterate until convergence of wage-bill shares.

Recovering true equilibrium values from 'hatted' equilibrium: Once the 'hatted' equilibrium is solved, we can construct the true equilibrium values by rescaling as follows:

$$\omega = \frac{\widetilde{Z}}{\overline{\varphi}^{1-\widetilde{\alpha}}} \tag{F2a}$$

$$\mathbf{W} = \omega^{\frac{1}{1+(1-\widetilde{\alpha})\varphi}} \widehat{\mathbf{W}}^{\frac{1+(1-\widetilde{\alpha})\theta}{1+(1-\widetilde{\alpha})\varphi}} \tag{F2b}$$

$$w_{ij} = \omega^{\frac{1}{1+(1-\widetilde{\alpha})\theta}} \mathbf{W}^{\frac{(1-\widetilde{\alpha})(\theta-\varphi)}{1+(1-\widetilde{\alpha})\theta}} \widehat{w}_{ij} \tag{F2c}$$

$$\mathbf{W}_j = \omega^{\frac{1}{1+(1-\widetilde{\alpha})\theta}} \mathbf{W}^{\frac{(1-\widetilde{\alpha})(\theta-\varphi)}{1+(1-\widetilde{\alpha})\theta}} \widehat{\mathbf{W}}_j \tag{F2d}$$

$$n_{ij} = \overline{\varphi} \left(\frac{w_{ij}}{\mathbf{W}_j}\right)^{\eta} \left(\frac{\mathbf{W}_j}{\mathbf{W}}\right)^{\theta} \left(\frac{\mathbf{W}}{1}\right)^{\varphi} \tag{F2e}$$

F.1 Scaling the economy

We set the scale parameters $\overline{\varphi}$ and $\widetilde{Z}$ in order to match average firm size observed in the data ($AveFirmSize^{Data} = 27.96$ from Table **??**), and average earnings per worker in the data ($AveEarnings^{Data} = \$65,773$ from Table A2):

$$\widehat{AveFirmSize}^{Data} = \frac{\int \left\{\sum_{i\in j} n_{ij}\right\} dj}{\int \left\{M_j\right\} dj} \tag{F3a}$$

$$\widehat{AveEarnings}^{Data} = \frac{\int \left\{\sum_{i\in j} w_{ij} n_{ij}\right\} dj}{\int \left\{\sum_{i\in j} n_{ij}\right\} dj} \tag{F3b}$$

To compute the values of $\overline{\varphi}$ and $\widetilde{Z}$ that allow us to match $AveFirmSize^{Data}$ and $AveEarnings^{Data}$, we substitute the model's values for n_{ij}, w_{ij}, and M_j into $AveFirmSize^{Data}$ and $AveEarnings^{Data}$. We repetitively substitute equations (F2a) through (F2e) into (F3a) and (F3b). We then solve for $\overline{\varphi}$ and $\widetilde{Z}$ in terms of 'hatted' variables as follows:

$$\overline{\varphi} = \frac{\frac{AveFirmSize^{Data}}{\widehat{AveFirmSize}^{Model}}}{\left(\frac{AveEarnings^{Data}}{\widehat{AveEarnings}^{Model}}\right)^{\varphi}} \tag{F4}$$

$$\widetilde{Z} = \overline{\varphi}^{1-\widetilde{\alpha}} \left(\frac{AveEarnings^{Data}}{\widehat{AveEarnings}^{Model}}\right)^{1+(1-\widetilde{\alpha})\varphi} \times \widehat{\mathbf{W}}^{-(1-\widetilde{\alpha})(\theta-\varphi)} \tag{F5}$$

where

$$\widehat{AveFirmSize}^{Model} = \frac{\int \left\{ \sum_{i \in j} \widehat{n}_{ij} \right\} dj}{\int \left\{ M_j \right\} dj}$$
$$\widehat{AveEarnings}^{Model} = \frac{\int \left\{ \sum_{i \in j} \widehat{w}_{ij} \widehat{n}_{ij} \right\} dj}{\int \left\{ \sum_{i \in j} \widehat{n}_{ij} \right\} dj}$$

The scaled model equilibrium values (defined by (F2a) through (F2e) evaluated at (F4) and (F5)) will now match $AveFirmSize^{Data}$ and $AveEarnings^{Data}$.

F.2 Recovering productivities

In Section 4 we discuss the potential to invert the model to recover productivities. We show how this may be achieved. Proceeding as below one may obtain, non-parameterically, the distribution of productivities of firms in the economy. In contrast, our approach in the body of the paper is instead to make a parameteric assumption on the distribution of productivities. Why? The approach of non-parameterically determining the distribution requires computation of the model *within* the Census Research Data Center, which is costly. However we detail this procedure here as it is straight-forward to implement on data that researchers have easier access to.

Take the expression for a firm's equilibrium wage from above, where we recognize that μ_{ij} is a closed form function of the wage-bill share s_{ij}:

$$w_{ij} = \omega^{\frac{1}{1+(1-\tilde{\alpha})\theta}} \mathbf{W}^{\frac{(1-\tilde{\alpha})(\theta-\varphi)}{1+(1-\tilde{\alpha})\theta}} \mu(s_{ij})^{\frac{1}{1+(1-\tilde{\alpha})\theta}} \widehat{z}_{ij}^{\frac{1}{1+(1-\tilde{\alpha})\theta}} s_{ij}^{-\frac{(1-\tilde{\alpha})(\eta-\theta)}{\eta+1} \frac{1}{1+(1-\tilde{\alpha})\theta}}$$

We can divide these expressions for two firms 1 and 2 in sector j. We drop extraneous subscripts:

$$\frac{w_1}{w_2} = \left(\frac{\mu(s_1)}{\mu(s_2)}\right)^{\frac{1}{1+(1-\tilde{\alpha})\theta}} \left(\frac{\widehat{z}_1}{\widehat{z}_2}\right)^{\frac{1}{1+(1-\tilde{\alpha})\theta}} \left(\frac{s_1}{s_2}\right)^{-\frac{(1-\tilde{\alpha})(\eta-\theta)}{\eta+1} \frac{1}{1+(1-\tilde{\alpha})\theta}}.$$

Recall that $s_{ij} = (w_{ij}/\mathbf{W}_j)^{\eta+1}$. Which implies that $(s_1/s_2) = (w_1/w_2)^{\eta+1}$. Using this

$$\left(\frac{s_1}{s_2}\right)^{\frac{1}{\eta+1}} = \left(\frac{\mu(s_1)}{\mu(s_2)}\right)^{\frac{1}{1+(1-\tilde{\alpha})\theta}} \left(\frac{\widehat{z}_1}{\widehat{z}_2}\right)^{\frac{1}{1+(1-\tilde{\alpha})\theta}} \left(\frac{s_1}{s_2}\right)^{-\frac{(1-\tilde{\alpha})(\eta-\theta)}{\eta+1} \frac{1}{1+(1-\tilde{\alpha})\theta}}.$$

which implies that

$$\frac{\widehat{z}_2}{\widehat{z}_1} = \frac{\mu(s_1)}{\mu(s_2)} \left(\frac{s_2}{s_1}\right)^{\frac{1+(1-\tilde{\alpha})(\eta)}{\eta+1}}. \tag{F6}$$

Given data on shares, and our estimates of the parameters η and θ which index the function μ, the right hand side can be treated as data. We can therefore invert the model to obtain *relative* productivities. Data on employment and wages in one firm can then be used to determine absolute productivities.

G Corporate Taxes and Labor Demand

Consider a single firm i. Assume constant returns to scale. Let the corporate tax rate be given by τ_c, and let the fraction of capital financed by debt be λ. Accounting profits of a firm (on which taxes are based) are given by

$$\pi^A = Pz_i k_i^{1-\alpha} n_i^{\alpha} - w_i n_i - \underbrace{\lambda r k_i}_{\text{interest expense}} - \delta k_i$$

The pre-tax economic profits of a firm are given by

$$\pi^E = Pz_i k_i^{1-\alpha} n_i^{\alpha} - w_i n_i - r k_i - \delta k_i$$

The after-tax economic profits of a firm are given by

$$\pi = \pi^E - \tau_c \pi^A$$

Define $\widetilde{z_i} = (1-\tau_c) z_i$, $\widetilde{w_i} = (1+\tau_c) w_i$, and $\widetilde{r} = (1+\lambda\tau_c) r + (1+\tau_c)\delta$. After substituting and solving, the profit maximization problem of the firm becomes:

$$\max_{k_i, n_i} \widetilde{z}_i P k_i^{1-\alpha} n_i^{\alpha} - \widetilde{w}_i n_i - \widetilde{r} k_i$$

Substituting for capital, the profit maximization problem becomes

$$\pi = \max_{n_i} \left[\left[(1-\alpha)^{\frac{1-\alpha}{\alpha}} - (1-\alpha)^{\frac{1}{\alpha}} \right] \widetilde{z}_i^{\frac{1}{\alpha}} \widetilde{r}^{-\frac{1-\alpha}{\alpha}} - \widetilde{w}_i \right] n_i$$

We can scale the profits by $\frac{1}{1+\tau_c}$ and then use the definition of $\widetilde{w}_i$ to write profits as follows:

$$\widehat{\pi} = \frac{\pi}{1+\tau_c} = \max_{n_i} \left[\widehat{MRPL}_i - w_i \right] n_i$$

Where the marginal product is given by,

$$\widehat{MRPL}_i = \frac{\left[(1-\alpha)^{\frac{1-\alpha}{\alpha}} - (1-\alpha)^{\frac{1}{\alpha}} \right] \widetilde{z}_i^{\frac{1}{\alpha}} \widetilde{r}^{-\frac{1-\alpha}{\alpha}}}{1+\tau_c}$$

In the estimation, we do not need to take a stance on the value of λ (the share of capital financed by debt), but this expression shows how corporate tax rates map to labor demand.

H Estimation details and bias exercise

H.1 Distribution of firms across markets

We assume there are 5,000 markets. For computational reasons, we must cap the number of firms per market since the Pareto distribution has a fat tail. We set the cap equal to 200 firms per market. Our results are not sensitive to the number of markets or the cap on firms per market. Figure H1 plots the mixture of Pareto distributions from which we draw the number of firms per market, M_j. The distribution of the number of firms per market, $G(M_j)$, is a mixture of Pareto distributions. The thin tailed Pareto has the following parameters: Shape=0.67, Scale=5.7, Location=2.0. The fat tailed Pareto has the following parameters: Shape=0.67, Scale=6.25×5.7, Location=2.0.

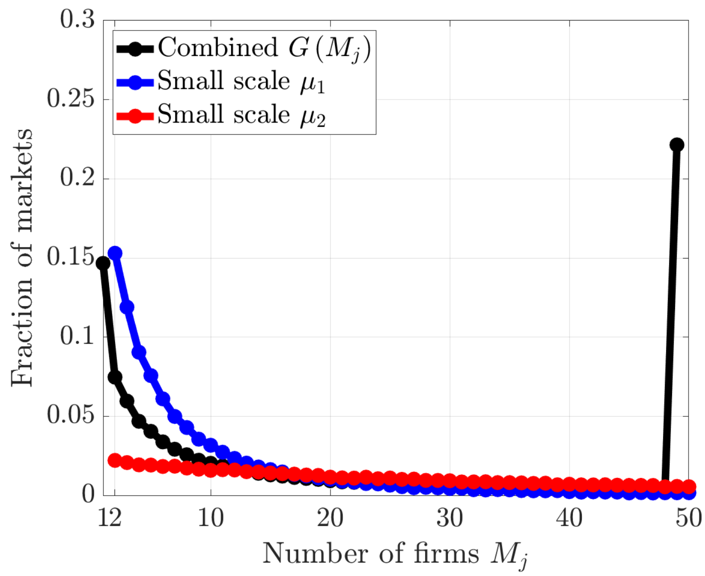

Figure H1: Distribution of the number of firms across sectors

Notes: This is a mixture of Pareto distributions. Thin Tailed: Shape=0.67, Scale=5.7, Location=2.0. Fat Tailed: Shape=0.67, Scale=6.25×5.7, Location=2.0.

Distribution of number of firms M_j	**Mean** (1)	**Std.Dev.** (2)	**Skewness** (3)
Data (LBD, 2014)	51.6	264.9	29.9
Model	51.6	264.9	28.7

Table H1: Distribution of firms across markets, $M_j \sim G(M_j)$

H.2 Tax Experiment Details

In each simulation of the model, we conduct a tax experiment where we simulate a common corporate tax change of $\Delta_\tau = \tau'_C - \tau_C = .01$, holding aggregate quantities fixed. We rerun our reduced-form regressions on the simulated data in order to recover average reduced form labor supply elasticities as a function of wage-bill shares. These market-share-dependent reduced form labor supply elasticities are the moments used to recover η and θ in Section 4. We describe the details of the exercise below:

1. Simulate the benchmark equilibrium, treat as date $t = 1$ 'data.'

2. C-corps in the model economy (recall there is a share ω_C of C-corps in all markets) have their taxes raised by 1 percentage point.

3. Simulate the 'post-shock' equilibrium, treat as date $t = 2$ 'data.'

4. Estimate the same reduced form regressions as Section 4 using the $t = 1, 2$ simulated data. Estimate the following regressions for each firm i in region j:

$$\log(n_{ijt}) = \alpha_i + \beta_n \tau_{Ct} + \gamma_{0s} s_{ijt} + \beta_{ns} \tau_{Ct} * s_{ijt} + \epsilon_{ijt}$$
$$\log(w_{ijt}) = \alpha_i + \beta_w \tau_{Ct} + \omega_{0s} s_{ijt} + \beta_{ws} \tau_{Ct} * s_{ijt} + u_{ijt}$$

5. Compute the employment and wage elasticities with respect to productivity, $\frac{d\log(n_{ijt})}{d\tau_{Ct}}$ and $\frac{d\log(w_{ijt})}{d\tau_{Ct}}$. Use these expressions to recover the average reduced form labor supply elasticities using the formula:
$$\widehat{\epsilon}(s_{ij}) = \frac{\beta_n + \beta_{ns} s_{ij}}{\beta_w + \beta_{ws} s_{ij}}$$

6. Use the recovered $\{\widehat{\epsilon}(s_{ijt}), s_{ijt}\}$ pairs as moments to recover η and θ.

H.3 Biases

To explore the difference between structural and reduced form labor supply elasticities, we conduct a Monte Carlo exercise where we simulate a perfectly idiosyncratic shock and then rerun our reduced-form regressions on the simulated data to recover reduced form labor supply elasticities. We average these across firms within payroll share bins and compare these to the structural labor supply elasticity implied by (8). We repeat this exercise for 5,000 simulations and report the averages in Figure 6. We describe the details of the exercise below:

1. Simulate the benchmark equilibrium, treat as date $t = 1$ 'data.'

2. Randomly select 1 firms in each market and increase their productivity by 1% (20% or 50%), holding aggregates fixed (assuming partial equilibrium).

3. Simulate 'post-shock' partial equilibrium (industry competitors adjust but aggregates are held fixed), treat as date $t = 2$ 'data.'

4. Estimate the same reduced form regressions as Section 4 using the $t = 1, 2$ simulated data. Estimate the following regressions for firms with payroll share s_{ij}^{wn} in bins with nodes $[.1, \ldots, .9]$.

$$\log(n_{ijt}) = \alpha_i + \beta_n \log(z_{ijt}) + \epsilon_{ijt} \;\; , \;\; \log(w_{ijt}) = \alpha_i + \beta_w \log(z_{ijt}) + u_{ijt}$$

5. In each bin, compute the employment and wage elasticities with respect to productivity, $\frac{d\log(n_{ijt})}{d\log(z_{ijt})}$ and $\frac{d\log(w_{ijt})}{d\log(z_{ijt})}$. Use these expressions to recover the predicted average reduced form labor supply elasticities using the formula:
$$\widehat{\epsilon} = \frac{\beta_n}{\beta_w}$$

6. Figure 6 plots these values at the upper cutoff of these bins. For shares equal to 0 and 1, the solution is exact $\varepsilon(1) = \widehat{\epsilon}(1) = \theta$, $\varepsilon(0) = \widehat{\epsilon}(0) = \eta$.

I Minimum wage

For ease of exposition, we lay out the minimum wage problem ignoring capital. Consider the household problem with the added constraint $n_{ijt} \leq \underline{n}_{ijt}$. For ease of interpretation we attach multiplier $\lambda_t \nu_{ijt}$ to the new labor supply constraint, normalized by the household budget multiplier λ_t:

$$U_0 = \max_{\{n_{ijt}, c_{ijt}\}} \sum_{t=0}^{\infty} \beta^t u \left(\mathbf{C}_t - \frac{1}{\overline{\varphi}^{\frac{1}{\varphi}}} \frac{\mathbf{N}_t^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}} \right) \quad , \quad \beta \in (0,1) \quad , \quad \varphi > 0$$

$$\begin{aligned} C_t &= \int \sum_i w_{ijt} n_{ijt} + \Pi_t & (\lambda_t) \\ n_{ijt} &\leq \underline{n}_{ijt} \;\; \forall \{ij\} & (\lambda_t \nu_{ijt}) \\ C_t &= \int \sum_i c_{ijt} dj & \end{aligned}$$

$$\mathbf{N}_t = \left[\int \left\{ \left[\sum_i n_{ijt}^{\frac{\eta+1}{\eta}} \right]^{\frac{\eta}{\eta+1}} \right\}^{\frac{\theta+1}{\theta}} dj \right]^{\frac{\theta}{\theta+1}}$$

In order to solve the firm's problem, we will have to take account of the normalized households multipliers, ν_{ijt}, on equation $(*)$. The firm's problem is given by:

$$\pi_{ijt} = \max_{n_{ijt}} \widetilde{Z} \widetilde{z}_{ijt} n_{ijt}^{\widetilde{\alpha}} - w_{ijt} n_{ijt}$$

$$s.t.$$

$$w_{ijt} = \nu_{ijt} + \overline{\varphi}^{-\frac{1}{\varphi}} \mathbf{N}_t^{\frac{1}{\varphi}} \left(\frac{\mathbf{N}_{jt}}{\mathbf{N}_t} \right)^{\frac{1}{\theta}} \left(\frac{n_{ijt}}{\mathbf{N}_{jt}} \right)^{\frac{1}{\eta}}$$

$$w_{ijt} \geq \overline{w}$$

Define the *shadow* wage-bill share:

$$\widetilde{s}_{ijt} = \frac{(w_{ijt} - \nu_{ijt}) n_{ijt}}{\sum_{i \in j} (w_{ijt} - \nu_{ijt}) n_{ijt}}$$

Define the *shadow* sectoral and aggregate wage indexes:

$$\widetilde{\mathbf{W}}_{jt} := \left[\sum_{i \in j} \left(w_{ijt} - \nu_{ijt} \right)^{1+\eta} \right]^{\frac{1}{1+\eta}} \qquad , \qquad \widetilde{\mathbf{W}}_t := \left[\int \widetilde{\mathbf{W}}_{jt}^{1+\theta} dj \right]^{\frac{1}{1+\theta}} .$$

I.1 Minimum wage solution algorithm

We implement the following solution algorithm. Initialize the algorithm by (i) guessing a value for $\widetilde{\mathbf{W}}_t^{(0)}$, (ii) assuming all firms are in *Region I*, which implies guessing $\nu_{ijt}^{(0)} = 0$. These will all be updated in the algorithm.

1. Solve the sectoral equilibrium:

   (a) Guess shadow shares $\widetilde{s}_{ijt}^{(0)}$.

   (b) In *Region I*, where minimum wage does not bind, solve for the firm's wage as before, except with the shadow aggregate wage index $\widetilde{\mathbf{W}}_t$ instead of $\mathbf{W}_t$:

$$w_{ijt} = \left[\omega\mu\left(\widetilde{s}_{ijt}\right)\widetilde{\mathbf{W}}_t^{(1-\widetilde{\alpha})(\theta-\varphi)}\widetilde{z}_{ijt}\widetilde{s}_{ijt}^{(l)-\frac{(1-\widetilde{\alpha})(\eta-\theta)}{\eta+1}}\right]^{\frac{1}{1+(1-\widetilde{\alpha})\theta}}$$

   (c) In all other regions *Region II, III, IV*, set $w_{ijt} = \underline{w}$.

   (d) Compute shadow wages using the guess $\nu_{ijt}^{(k)}$: $\widetilde{w}_{ijt} = w_{ijt} - \nu_{ijt}^{(k)}$

   (e) Update shares using $\widetilde{w}_{ijt}$:

$$\widetilde{s}_{ijt}^{(l+1)} = \frac{\widetilde{w}_{ijt}^{1+\eta}}{\sum_{i\in j}\widetilde{w}_{ijt}^{1+\eta}} \quad \left(:= \frac{\widetilde{w}_{ijt}n_{ijt}}{\sum_{i\in j}\widetilde{w}_{ijt}n_{ijt}} = \frac{\widetilde{w}_{ijt}\overline{\varphi}\left(\frac{\widetilde{w}_{ijt}}{\widetilde{\mathbf{W}}_{jt}}\right)^{\eta}\left(\frac{\widetilde{\mathbf{W}}_{jt}}{\widetilde{\mathbf{W}}_t}\right)^{\theta}\widetilde{\mathbf{W}}_t^{\varphi}}{\sum_{i\in j}\widetilde{w}_{ijt}\overline{\varphi}\left(\frac{\widetilde{w}_{ijt}}{\widetilde{\mathbf{W}}_{jt}}\right)^{\eta}\left(\frac{\widetilde{\mathbf{W}}_{jt}}{\widetilde{\mathbf{W}}_t}\right)^{\theta}\widetilde{\mathbf{W}}_t^{\varphi}}\right)$$

   (f) Iterate over (b)-(e) until $\widetilde{s}_{ijt}^{(l+1)} = \widetilde{s}_{ijt}^{(l)}$.

2. Recover employment n_{ijt} according to the current guess of firm region. First use $\widetilde{w}_{ijt}$ to compute $\widetilde{\mathbf{W}}_{jt}$, $\widetilde{\mathbf{W}}_t$. Then by region:

   (I) Firm is unconstrained:

$$n_{ijt} = \overline{\varphi}\left(\frac{w_{ijt}}{\widetilde{\mathbf{W}}_{jt}}\right)^{\eta}\left(\frac{\widetilde{\mathbf{W}}_{jt}}{\widetilde{\mathbf{W}}_t}\right)^{\theta}\widetilde{\mathbf{W}}_t^{\varphi}$$

   (II) Firm is constrained and employment is determined by the household labor supply curve at $\underline{w}$:

$$n_{ijt} = \overline{\varphi}\left(\frac{\underline{w}}{\widetilde{\mathbf{W}}_{jt}}\right)^{\eta}\left(\frac{\widetilde{\mathbf{W}}_{jt}}{\widetilde{\mathbf{W}}_t}\right)^{\theta}\widetilde{\mathbf{W}}_t^{\varphi}$$

   (III),(IV) Firm is constrained and employment is determined by firm $MRPL_{ij}$ curve at $\underline{w}$:

$$n_{ijt} = \left(\frac{\widetilde{\alpha}\widetilde{Z}\widetilde{z}_{ijt}}{\underline{w}}\right)^{\frac{1}{1-\widetilde{\alpha}}}$$

3. Update $\nu_{ijt}^{(k)}$:
   (a) Use n_{ijt} to compute $\mathbf{N}_{jt}$, $\mathbf{N}_t$.
   (b) Update ν_{ijt} from the *household's* first order conditions:

$$\nu_{ijt}^{(k+1)} = w_{ijt} - \overline{\varphi}^{-\frac{1}{\varphi}} \left(\frac{n_{ijt}}{\mathbf{N}_{jt}}\right)^{\frac{1}{\eta}} \left(\frac{\mathbf{N}_{jt}}{\mathbf{N}_t}\right)^{\frac{1}{\theta}} \mathbf{N}_t^{\frac{1}{\varphi}}$$

4. Update $\widetilde{\mathbf{W}}_t^{(k)}$:
   (a) Compute $\widetilde{w}_{ijt} = w_{ijt} - \nu_{ijt}^{(k+1)}$
   (b) Use $\widetilde{w}_{ijt}$ to update the aggregate wage index to $\widetilde{\mathbf{W}}_t^{(k+1)}$.
5. Update firm regions:
   (a) Compute profits for all firms: $\pi_{ijt} = \widetilde{Z}\widetilde{z}_{ijt} n_{ijt}^{\widetilde{\alpha}} - \underline{w} n_{ijt}$.
   (b) If in sector j there exists a firm with $w_{ijt} < \overline{w}$, then move the firm with the lowest wage into *Region II*.
   (c) If in sector j there exists a firm that was initially in *Region II* and has negative profits $\pi_{ijt} < 0$, move that firm into *Region III*.[68]
6. Iterate over (1) to (5) until $\nu_{ijt}^{(k+1)} = \nu_{ijt}^{(k)}$ and $\widetilde{\mathbf{W}}_t^{(k+1)} = \widetilde{\mathbf{W}}_t^{(k)}$.

J Discussion of empirical estimation

As discussed in Section 3, the model predicts that the labor supply elasticity faced by firms varies by their market share (equation 8). If this relationship were known in the data, it would precisely pin down the elasticities of substitution of labor within and across sectors. Existing work estimating labor supply elasticities to firms has focused either on specific markets (e.g. (Webber, 2016) or in well identified responses to small experimental variations in wages (Dube, Jacobs, Naidu, and Suri, 2019; Dube, Cengiz, Lindner, and Zipperer, 2019). A contribution of this paper is to estimate a share-elasticity relationship through a novel quasi-natural experiment using a large cross-section of firms.

The intuition for our procedure is as follows. We first estimate the rate at which labor demand shocks *pass-through* to wages and employment and the reduced form relationship between these labor supply elasticities and local labor market shares. We then invert this empirical relationship using our model to recover estimates of the structural parameters that control the relative substitutability of labor within and between markets. To identify how pass-through rates vary by market share, we compare how the

[68] We do not need to distinguish *Region III* from *Region IV* in the algorithm, since it the determination of equilibrium wages and employment are the same in each region.

firm responds to these labor demand shocks differentially across markets within the same state, but in which their shares of the labor market differ.

This procedure requires a shock to labor demand in order to trace out the labor supply curve. We use state corporate tax changes which constitute a shock to firm labor demand via their distortion of accounting profits relative to economic profits, shifting the marginal revenue product of labor.[69] Both Suárez Serrato and Zidar (2016) and Giroud and Rauh (2019) have studied the impact of state-level corporate tax shocks on local economic activity. We address three issues that may arise: (i) apportionment of state taxes across multi-state production units may mean that state corporate taxes do not affect firms within a state, (ii) taxes are anticipated, (iii) such shocks affect all firms in a region and so can only be used to identify θ.

First, Suárez Serrato and Zidar (2016) show that the impact of corporate taxes on local economic activity is extremely similar for both (i) the statutory corporate taxes that we use and (ii) effective corporate taxes adjusted for apportionment weights.[70] Since establishment sales and company property values are not available to us, we focus on statutory taxes rates compiled by Giroud and Rauh (2019) and based on Suárez Serrato and Zidar (2016) we do not adjust for the apportionment regime of the state.

Second, both Suárez Serrato and Zidar (2016) and Giroud and Rauh (2019) establish that the inclusion of other aspects of changes to fiscal policy around the corporate tax changes does not affect their measured elasticities of local economic activity to corporate taxes.[71]

Third, the fact that (i) only C-corps pay statutory corporate tax rates, (ii) the structure of our model and (iii) Monte Carlo exercises, provide support that we may infer η and θ from a shock that affects some but not all firms. We briefly discuss this in more detail.

J Competitive vs. Oligopolistic Economies

Characterization. To illustrate, Figure J1 extends our example Figure **??**, adding the competitive outcomes for the three labor markets studied. In the sector with three firms, the payroll share of the most productive firm increases, while that of the two least productive firms fall. As a consequence, concentration increases. Meanwhile, the employment at the most productive firm also increases, while their competitors' fall, improving the allocation of employment in the economy and increasing output.

[69] We have not included corporate taxes in our benchmark model. We show that the mapping of our model to the data does not require us to take a stance on the transmission mechanism linking corporate taxes to productivity. Nevertheless, Appendix G shows how corporate tax rates map to shocks to the marginal revenue productivity of labor in our framework.

[70] See their discussion of Table A21, p.19 (emphasis added): "Column (6) of Table 5 and Appendix Table A21 show that using statutory state corporate tax rates in Equation 21 (instead of business tax rates τ_b) results in similar and significant estimates, indicating that *our measure of business tax rates is not crucial for the results*."

[71] Giroud and Rauh (2019) establish the plausible exogeneity of state-corporate tax changes. From a public finance perspective they study the effects of state corporate tax changes on employment and wages. Their focus is *within firm*, *across state* responses, and the reallocation of firm employment across states following tax changes. For an exhaustive description of these tax changes we point the interested reader to their paper.

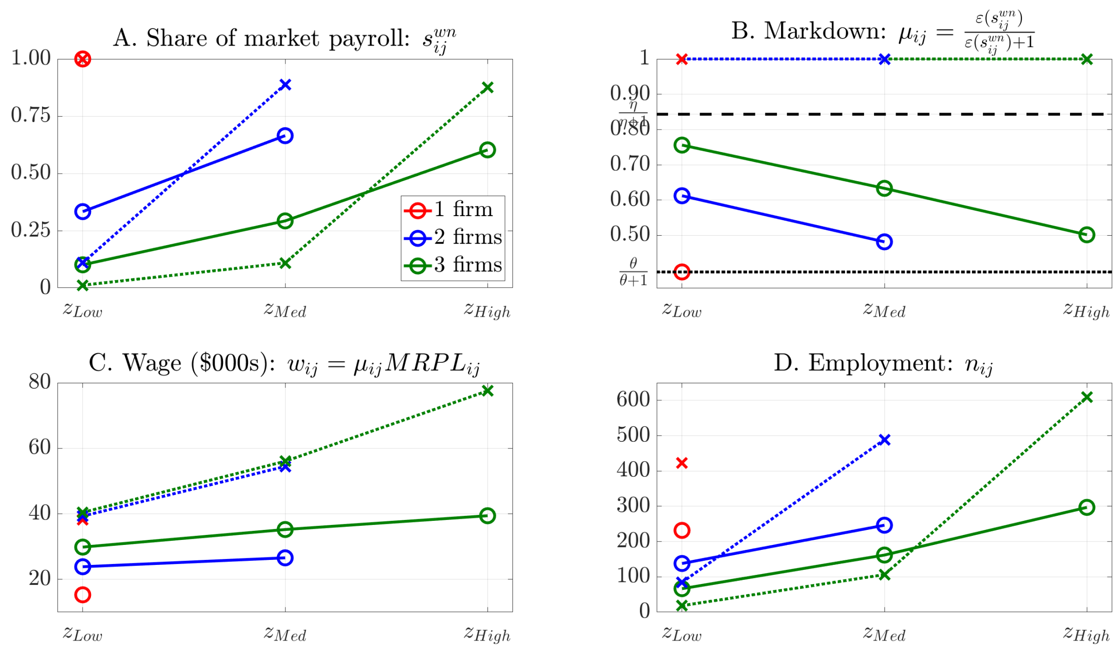

Figure J1: Oligopsonistic [solid] and competitive [dashed] equilibrium in three labor markets

Notes: Figure constructed from model under estimated parameters (Table 3). Low, medium and high productivities of the firms correspond to the 10^{th}, 50^{th} and 90^{th} percentiles of the productivity distribution.