Earnings Inequality in Production Networks*

Federico Huneeus† Kory Kroft‡ Kevin Lim §

Abstract

Why do firms differ in the wages paid to otherwise identical workers and in the share of revenue that they allocate to labor? This paper explores the role of production networks. Using linked employer-employee and firm-to-firm trade transactions data from Chile, we show that firms with better access to both buyers and suppliers of intermediate inputs tend to have higher earnings premia and lower labor shares. Motivated by these two facts, the paper develops and estimates a model with labor market power, worker and firm heterogeneity, and heterogeneity in firm-to-firm linkages in the production network. Greater access to larger downstream buyers and more efficient upstream suppliers raises the marginal revenue product of labor and lowers the relative cost of intermediates to labor. This leads to higher wages in the presence of labor market power and lower labor shares when labor and materials are substitutes. Through counterfactual simulations of the estimated model, we find a substantial role for production networks in explaining the variances of earnings premia and labor shares across firms.

Keywords: Earnings inequality, production networks, monopsony power.

JEL Codes: J31, F16.

*We thank Lorenzo Caliendo, Arnaud Costinot, Dave Donaldson, Emmanuel Farhi, Cecilia Fieler, Gregor Jarosch, Pete Klenow, Sam Kortum, Andres Rodriguez-Clare, Bradley Setzler, Isaac Sorkin, Felix Tintelnot, Jose Vasquez, Ivan Werning and Daniel Xu for invaluable comments on our paper. We are also grateful for feedback from seminar participants at the Georgetown, Harvard University, MIT, University of Arizona, UBC, University of Chicago, Universidad de Chile, Universidad Mayor, UT Austin, University of Toronto; and from conference participants at the Yale University Cowles Trade Day conference, the Central Bank of Chile Heterogeneity in Macroeconomics workshop, the West Coast Trade Conference, the Econometric Society World Congress, the European Economics Association Congress, the Stanford Institute for Theoretical Economics, and the ASSA 2021 Annual Meeting. We thank the Munk School of Global Affairs and Public Policy for financial support.

†Central Bank of Chile (fhuneeus@bcentral.cl).

‡University of Toronto and NBER (kory.kroft@utoronto.ca).

§University of Toronto (kvn.lim@utoronto.ca).

1 Introduction

Firms differ in the wages that they pay to otherwise identical workers and in the share of revenue that they allocate to labor, with larger firms tending to have higher premia and lower labor shares.[1] In Chile, a developing economy with high income inequality, earnings premia and labor shares of value-added are 2.2 times higher and 8 times lower respectively for firms at the 75th percentile of the corresponding distribution compared with firms at the 25th percentile. What explains these patterns? Understanding the sources of firm heterogeneity in earnings premia and labor shares has been a longstanding focus in economics. Firm earnings premia have been shown to be an important driver of earnings inequality across workers (see Card et al. (2018) for a survey). Further, firm-level heterogeneity in the labor share has been shown to be relevant for understanding trends in the aggregate labor share (Autor et al. (2020)).

This paper investigates the role of production networks in shaping firm heterogeneity in earnings premia and labor shares, using a framework with firm and worker heterogeneity, labor market power, and a flexible elasticity of substitution between labor and materials. Production networks may matter for earnings premia for two reasons: (i) firms who tend to match with larger downstream customers face higher effective demand and (ii) firms who source materials from more efficient upstream suppliers will incur a lower cost of materials. These two mechanisms lead to a higher marginal revenue product of labor. Combined with labor market power, they lead to a higher firm wage premia.[2] Similarly, when labor and materials are gross substitutes, a firm's labor share of cost is decreasing in the relative cost of labor to materials. Thus, a firm may have a lower labor share due to a higher cost of labor arising from (i) and (ii), or alternatively due to a lower cost of materials arising directly from greater access to more efficient suppliers.

Existing studies have typically interpreted firm heterogeneity in earnings premia and labor shares as arising from differences in the innate characteristics of firms themselves. For example, Van Reenen (1996), Kline et al. (2019), and Lamadon et al. (2022) consider heterogeneity in firm earnings premia as arising in part from differences in firm productivities.[3] Similarly, Kehrig and Vincent (2021) attribute differences in labor shares to differences in firm-specific productivity and demand shocks, while Karabarbounis and Neiman (2014) focus on heterogeneity in the cost of capital inputs. Our framework also allows a role for innate firm characteristics to shape earnings premia and labor shares. However, our focus is on the production network because much less is known about its importance for driving these patterns. Furthermore, production network heterogeneity has been shown to be crucial for explaining differences in firm outcomes such as

[1] For example, see Abowd et al. (1999), Card et al. (2018), and Bonhomme et al. (2019) on heterogeneity in earnings premia, Oi and Idson (1999) on the relationship between firm size and earnings premia, and Autor et al. (2020) and Kehrig and Vincent (2021) on labor share heterogeneity and how this correlates with firm size.

[2] In our framework, we show that (i) always leads to higher earnings premia, while (ii) leads to higher earnings premia under reasonable restrictions on model parameters that we discuss below.

[3] Dunne et al. (2004), Faggio et al. (2010), and Barth et al. (2016) also interpret trends in wage dispersion as being related to productivity dispersion across industries and firms.

size (Bernard et al. (2022)), that are well-known to be *correlated* with earnings premia and labor shares, but its implications for labor market outcomes *directly* are not well understood.[4]

Our paper makes three key contributions. First, using linked employer-employee and firm-to-firm transactions data from Chile, we document new stylized facts showing that firms with greater access to larger customers and more efficient suppliers in the production network tend to have higher earnings premia and lower labor shares of both value-added and cost. Second, we develop a structural model featuring labor market power and production networks that is capable of rationalizing these new facts, as well as firm heterogeneity in earnings premia and labor shares more generally.[5] To validate the key mechanisms in our model relating production network linkages to firm earnings premia and labor shares, we provide reduced-form evidence showing that exogenous network demand and material cost shocks translate into changes in worker earnings. Third, we structurally estimate the model using the Chilean administrative data, showing in particular how to estimate the labor-materials substitution elasticity in a model-consistent way when both inputs are heterogeneous within the firm. Using the estimated model, we perform counterfactual simulations to investigate the sources of firm heterogeneity in earnings premia and labor shares, focusing in particular on quantifying the role of production network heterogeneity in driving these patterns.

Our three key findings from the estimated structural model are as follows. First, we estimate that labor and materials are gross substitutes and statistically reject the hypothesis of Cobb-Douglas production functions. This is important because, as we show, Cobb-Douglas technology is necessary for the production function to have a value-added representation, as is typically assumed in models of firm labor market power. Second, we find that production network heterogeneity is a key driver of firm heterogeneity in earnings premia and labor shares, accounting for: (i) one-third of the variation in firm-specific earnings premia and 13% of the overall variation in worker earnings; (ii) half of the positive covariance between firm size and earnings premia; (iii) one-quarter of the variation in labor shares of value-added; and (iv) two-thirds of the negative covariance between firm size and labor value-added shares. Third, we find that the importance of the production network for explaining firm heterogeneity in earnings premia differs by almost a factor of two when the production function is restricted to be of a value-added form. Taken together, these findings highlight the importance of accounting

[4] At the same time, we show that size is not a sufficient statistic for a firm's earnings premium, so that unpacking the determinants of heterogeneity in size is complementary but not equivalent to unpacking the determinants of heterogeneity in firm earnings premia.

[5] Understanding the sources of firm heterogeneity in earnings premia and labor shares in a single framework is important. The frameworks that have been proposed to explain firm heterogeneity in earnings premia typically assume value-added production functions with constant labor shares, which imply counterfactual predictions about the passthrough of firm productivity shocks to worker earnings (e.g., see the discussion in Berger et al. (2019)). Similarly, the frameworks used to examine firm heterogeneity in labor shares typically assume perfectly competitive labor markets, implying that firms can scale employment without affecting the unit cost of labor and thus the incentive to substitute toward other inputs.

for firm-to-firm production network linkages and of moving away from value-added production functions in explaining empirical patterns of firm heterogeneity in earnings premia and labor shares.

We now explain in greater detail the contributions of our paper. First, to motivate our focus on production network linkages, we begin in section 2 by showing that differences in earnings and labor shares across firms are empirically related to production network heterogeneity in Chile. We establish this using linked employer-employee and firm-to-firm transactions data, which allow us to observe annual earnings for every employee at each firm in our data, as well as annual transaction values between buyer-seller firm pairs. We use these data to perform two two-way statistical decompositions. First, a decomposition of worker earnings into worker and firm effects (as in Bonhomme et al. (2019)), which is used to obtain a measure of firm-specific earnings premia. This improves on approaches that study firm earnings premia using average wages, which may be confounded by compositional differences in worker quality across firms. Second, a decomposition of firm-to-firm sales into buyer and seller effects (as in Bernard et al. (2022)), which is used to construct measures of a firm's access to customers *downstream* and suppliers *upstream* in the production network. This improves on approaches that measure a firm's position in the production network using coarser data (e.g., industry-level input-output tables). We then establish two novel stylized facts: firms with greater downstream and upstream access tend to have (i) higher earnings premia and (ii) lower labor shares.

In section 3, we develop a structural model to study firm heterogeneity in earnings premia and labor shares. We model the production network as a set of heterogeneous linkages between firms that trade intermediate inputs (as in Lim (2019), Huneeus (2019), Bernard et al. (2022), and Dhyne et al. (2022)) and allow for labor market power arising from horizontal employer differentiation (as in Manning (2003), Card et al. (2018), Lamadon et al. (2022), Azar et al. (2022), Chan et al. (2022), and Kroft et al. (2022)). In addition, the framework allows for worker heterogeneity in ability, firm heterogeneity in employer amenities (as in Rosen (1986) and Sorkin (2018)) that vary at the worker-firm level, and complementarities in production between workers and firms. These features allow the model to speak directly to the richness of the Chilean administrative data, including the heterogeneous sorting of different worker types to different firms. We show that the model can reconcile firm heterogeneity in both earnings premia and labor shares, and that it rationalizes the relationships between labor market outcomes and production network linkages that we document in section 2. This is useful since it allows us to probe the deeper economic mechanisms underlying these correlations.

In section 4, we use the model to characterize the key mechanisms linking firm earnings premia and labor shares to production network heterogeneity. First, we show that the reduced-form measures of upstream and downstream access described in section 2 are in fact sufficient statistics for the relevance of the production network in firms' wage-setting decisions in the model. Second, we establish *why* firms with greater downstream and upstream access in the production network tend to have higher earnings premia. Greater downstream access implies higher demand and hence higher scale, requiring higher wages to attract more workers since the labor supply elasticity (γ) is finite. Similarly, greater upstream access implies a lower cost of materials and tends to increase wages through higher scale. However, this also induces substitution away from labor toward materials, which may increase or decrease wages depending on the elasticity of substitution between the two inputs (ϵ). We show that a necessary and sufficient condition for greater upstream access to lead to higher wages (one satisfied by our empirical estimates) is that ϵ is small relative to the elasticity of substitution across suppliers in the production function (σ). Third, we show how production network heterogeneity can alter the covariance between a firm's size and its earnings premium. Intuitively, firms with lower costs of materials are able to achieve the same size with fewer workers and hence offer lower wages. A corollary of this result is that size is not a sufficient statistic for a firm's earnings premium. Finally, we show that firms facing lower costs of materials relative to labor optimally choose smaller labor shares of both cost and value-added. Hence, larger firms, which tend to have greater upstream access and lower costs of materials relative to labor, will also tend to have smaller labor shares.

To externally validate our model, section 5 provides reduced-form evidence of the passthrough of demand and material cost shocks into changes in labor market outcomes at the firm level. This analysis relies on transactions-level customs data for Chile to construct external export demand and import cost shocks using a Bartik shift-share design, accounting for both direct exposure to these shocks (i.e., firms that directly export or import) as well as indirect exposure through the production network (i.e., firms that sell to or buy from a direct exporter or importer). As predicted by our model, increases in demand and reductions in material costs have positive effects on firm wage bills and employment. These findings also contribute to the empirical literature studying the relationship between firm shocks and worker earnings (for example, Guiso et al. (2005) and Chan et al. (2021)) by extending the analysis to account for passthrough via production network linkages.

In section 6, we formally establish identification of model parameters (including ϵ, γ, and σ) and structurally estimate our model using the Chilean administrative data. Three steps of our identification strategy are particularly important. First, while it is well-known that the labor-materials substitution elasticity ϵ can be identified from the relationship between firms' relative expenditures on these inputs and their relative prices, the literature offers little guidance as to how input prices should be aggregated when both wages and material prices are heterogeneous within the firm.[6] We show that the two-way statistical decompositions of worker earnings (Bonhomme et al. (2019)) and firm-to-firm transactions (Bernard et al. (2022)) can be used to

[6] For example, existing approaches typically use average wages as a measure of the cost of labor (as in Doraszelski and Jaumandreu (2018)) or treat material prices as an industry rather than firm characteristic (as in Oberfield and Raval (2019)).

construct price indices for labor and materials that control for heterogeneity in these inputs. We construct these price indices and use an instrumental variables strategy (following Doraszelski and Jaumandreu (2018)) to estimate $\epsilon = 1.5$, indicating gross substitutability of labor and materials. We also statistically reject the hypothesis of a value-added production function.[7] Second, we show how to identify the labor supply elasticity γ using the passthrough of changes in firm wage bills into changes in worker earnings. In particular, we demonstrate that existing approaches using the passthrough of changes in firm value-added shocks to identify γ (as in Guiso et al. (2005) and Lamadon et al. (2022), for example) are not valid in the presence of firm heterogeneity in material cost shares. Using our approach, we estimate $\gamma = 5.5$. Third, we identify the elasticity of substitution across products σ using the ratio of aggregate sales to profits, which yields an estimate of $\sigma = 3.1$.

Finally, in section 7, we use the model to investigate what drives the variances of worker earnings, firm earnings premia, and labor shares, as well as the covariances between firm earnings premia, labor shares, and size. We first show that our estimated model provides a good fit to these outcomes in the data. We then solve for counterfactual equilibria in which heterogeneity in different sets of model primitives – including a firm's network linkages – is removed and use these simulations to quantify the importance of production network heterogeneity for the labor market variances and covariances. In each case, we find that production network heterogeneity plays a key role. Furthermore, restricting the production function to be of a value-added form ($\epsilon = 1$) greatly overstates the importance of production network heterogeneity for explaining earnings inequality. Thus, seemingly small differences in the value of ϵ can have quantitatively large effects on the importance of production networks for earnings inequality.

To our knowledge, there are only three other papers that study linked employer-employee and firm-to-firm transactions data. Adao et al. (2020) use data from Ecuador to measure the effects of international trade on individual-level factor prices, while Demir et al. (2018) study the effects of trade-induced product quality upgrading on wages in Turkey. Both of these analyses assume a market price for skill and focus on the effects of trade shocks. In contrast, we allow for imperfect competition in labor markets and use our data to speak to the role of the production network itself in shaping earnings inequality. Finally, Alfaro-Ureña et al. (2020) adopt an event study research design to examine the effects on worker earnings in Costa Rica when a local firm starts interacting with multinationals. In contrast, we use our data to address both worker-level earnings and aggregate outcomes such as earnings inequality, which requires a general equilibrium model.

[7] As discussed in section 4.2.1 below, value-added production functions are only consistent with a Cobb-Douglas structure on the aggregation of labor and materials ($\epsilon = 1$).

2 Motivating Facts

We first document the severity of earnings inequality in Chile and show that, as in many other countries, differences in firm earnings premia matter for explaining differences in worker earnings. We also document the extent of labor share heterogeneity amongst Chilean firms. To motivate our focus on production network linkages, we then present two new stylized facts relating a firm's access to customers and suppliers in the production network to its earnings premium (Fact 1) and its labor shares of cost and value-added (Fact 2). In documenting these facts, we rely on employer-employee and firm-to-firm trade transactions data from Chile for 2005-2010, a detailed discussion of which we defer to section 6.[8]

2.1 Heterogeneity in firm earnings premia and labor shares in Chile

Figure 1 shows a scatter plot of the Gini coefficient of worker earnings against the 90-10 earnings percentile ratio for current OECD members and a few other countries, where values are averaged across 2005-2010, while row (i) of Table 1 provides other moments of the log earnings distribution in Chile. Evidently, earnings inequality in Chile is severe. Workers at the 90^{th} percentile of the earnings distribution have earnings that are almost nine times higher than workers at the 10^{th} percentile of the distribution, while in comparison, the average 90-10 ratio among all other OECD members is 4.3, less than half of the ratio in Chile.[9] Similarly, the earnings Gini coefficient in Chile is the highest among OECD members and lies above the 90^{th} percentile across all countries during this period based on data from the World Bank. This underscores the importance of unpacking the determinants of earnings inequality in a country such as Chile.

To obtain a preliminary look at the role of firms in driving earnings inequality in Chile, we first decompose the log earnings of worker m at firm i and time t as:

$$\log w_{imt} = \theta_i x_m + \log f_{it} + \hat{x}_{mt} \tag{2.1}$$

where x_m is a worker fixed effect, f_{it} is a time-varying firm effect, θ_i allows for worker-firm interactions, and $\hat{x}_{mt}$ is an orthogonal residual.[10] In particular, the firm effect f_{it} provides a measure of an employer's premium on wages after adjusting for differences in worker composition, which would otherwise be reflected in measures such as the average wage at the firm.

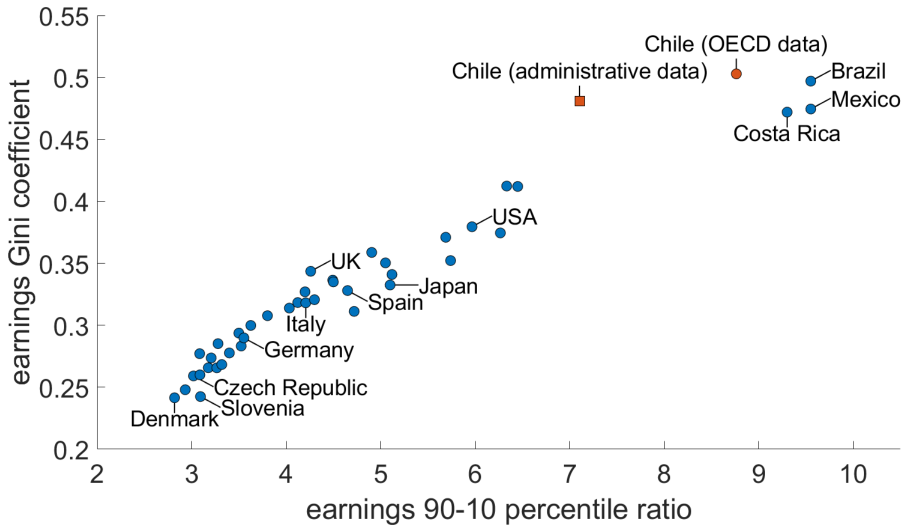
We estimate the decomposition (2.1) using matched employer-employee data following a procedure described in detail below (see sections 6.3.1 and 6.3.2). Row (ii) of Table 1 shows

[8] Since we do not observe clear trends in earnings inequality or labor share heterogeneity over our sample period, we focus below on cross-sectional patterns in our data.

[9] The 90-10 earnings ratio based on the administrative employer-employee data that we use in the paper is slightly lower at 7.1, but still higher than the 90-10 ratios for all other OECD members.

[10] As we discuss below, this decomposition is consistent with the structural model that we develop in the paper and can be viewed as an extension of the well-known earnings model in Abowd et al. (1999) to allow for worker-firm interactions (as in Bonhomme et al. (2019)) and time-variation in the firm effect.

Figure 1: Earnings 90-10 percentile ratio vs. Gini coefficient, 2005-2010 average



Notes: Statistics for Chile are obtained from the OECD Income Inequality database (red circle) and from an administrative employer-employee dataset described in section 6.1 (red square). Data for all other countries are obtained from the OECD Income Inequality database.

Table 1: Moments of the distributions of key labor market and production network variables

		mean	s.d.	p10	p25	p50	p75	p90
(i)	worker earnings	9.37	0.75	8.51	8.77	9.26	9.81	10.40
(ii)	firm earnings effect	0.53	0.44	0.00	0.06	0.56	0.87	1.11
(iii)	firm average wage	8.85	0.62	8.19	8.38	8.68	9.21	9.74
(iv)	firm downstream access	4.99	2.05	2.50	3.47	4.78	6.26	7.73
(v)	firm upstream access	4.90	1.81	2.67	3.73	4.88	5.96	7.07
(vi)	firm labor share of cost	0.43	0.37	0.05	0.11	0.26	0.88	1.00
(vii)	firm labor share of VA	0.53	0.57	0.09	0.17	0.33	0.61	1.28

Notes: Moments in row (i) are calculated at the worker level and all other moments are calculated at the firm-level (unweighted). Variables in all rows except (vi) and (vii) are in logs.

moments of the distribution of firm effects f_{it}, while for comparison, row (iii) shows moments from the distribution of average wages across firms. Evidently, there is substantial variation in firm effects on earnings. For instance, firms at the 75^{th} percentile of the firm effect distribution have earnings premia that are 2.2 ($\approx e^{0.87-0.06}$) times greater than firms at the 25^{th} percentile of the distribution. As discussed in Appendix A, we also find that the variation in firm earnings premia explains around 11% of the overall variation in worker earnings, while the covariance between firm earnings premia and worker effects explains around 20%.[11] In addition, between-firm variation in log earnings explains 46% of total log earnings variance in the average year.[12] These results highlight the importance of differences in firm earnings premia for explaining differences in worker earnings more broadly.

Furthermore, firms are heterogeneous not only in the wages that they offer to workers, but also in the extent to which they allocate revenue to labor versus other productive inputs (including firm owners, who capture profits). To fix ideas, suppose that firms produce output using labor and materials, and let R_{it}, E_{it}^L, and E_{it}^M denote firm i's sales, expenditures on labor, and expenditures on materials, respectively. We consider the share of labor in total production costs and the share of labor in value-added, defined respectively as:

$$s^{L/C_{it}} \equiv \frac{E_{it}^L}{E_{it}^L + E_{it}^M}, \qquad s_{it}^{L/VA} \equiv \frac{E_{it}^L}{R_{it} - E_{it}^M} \tag{2.2}$$

Note that these shares can be interpreted as measures of within-firm inequality: between payments made to labor versus the firm's suppliers in the production network ($s^{L/C}$) and between payments made to labor versus firm owners ($s^{L/VA}$).[13] Rows (vi) and (vii) show moments of the distributions of $s^{L/C}$ and $s^{L/VA}$, from which we observe substantial heterogeneity in both of these labor shares across firms. For example, firms at the 75^{th}-percentile of the labor cost share distribution spend 88% of their production costs on workers, compared with only 11% for firms at the 25^{th}-percentile of the distribution.

[11] In comparison, using US data, Lamadon et al. (2022) find that the variance of log $\tilde{f}_{it}$ explains 4.3% of log earnings variance – less than half of our value for Chile – and that the sorting covariance explains 13.0% – about two-thirds of our value for Chile.

[12] In comparison, Song et al. (2019) report between-firm shares of 38-41% for the US over the same time period and a maximum share of 42% over their entire sample period (1978-2013).

[13] These shares are closely linked through the identity $s^{L/VA} = \frac{s^{L/C}}{\tilde{\mu}-1+s^{L/C}}$, where the firm's sales-cost ratio $\tilde{\mu} \equiv \frac{R}{EL+EM}$ is reflective of its output markup.

2.2 Stylized facts: the production network, earnings premia, and labor shares

Fact 1: Firms that have greater access to customers and suppliers in the production network tend to pay higher wages.

The statistical decompositions presented above establish the importance of differences in employers for earnings inequality, but do not shed light on *why* employers matter for wages. Here, we present preliminary evidence highlighting the role of heterogeneity in employers' access to customers and suppliers in the production network. To measure this, we first decompose log sales by a seller j to a buyer i at time t as follows:

$$\log r_{ijt} = \log d_{it} + \log s_{jt} + \log e_{ijt} \tag{2.3}$$

where d_{it} is a buyer effect, s_{jt} is a seller effect, and e_{ijt} is an orthogonal residual.[14] Intuitively, firms with larger buyer effects tend to spend more on inputs from their suppliers conditional on their suppliers' characteristics, while firms with larger seller effects tend to sell more to their customers conditional on their customers' characteristics. We estimate this decomposition using firm-to-firm trade data following a procedure described in detail below (see section 6.3.3). We then construct measures of a firm's *downstream access* (D_{it}^{net}) and *upstream access* (S_{it}^{net}):

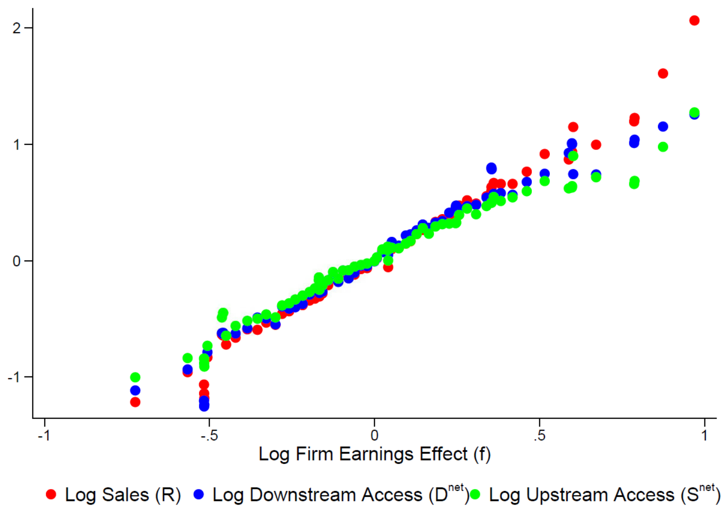
$$D_{it}^{net} \equiv \sum_{j \in \Omega_{it}^{C}} d_{jt} e_{jit}, \qquad S_{it}^{net} \equiv \sum_{j \in \Omega_{it}^{S}} s_{jt} e_{ijt} \tag{2.4}$$

where Ω_{it}^{C} and Ω_{it}^{S} denote the set of firm i's customers and suppliers, respectively. Intuitively, D_{it}^{net} summarizes the extent to which firm i is connected to customers that have high demand for intermediate inputs, while S_{it}^{net} summarizes the extent to which the firm is connected to suppliers that tend to have high sales to other firms in the production network. As we show in the structural model that we develop below, D_{it}^{net} and S_{it}^{net} are sufficient statistics for the relevance of the production network for firms' wage-setting decisions.

The relationship between firm earnings premia and network access is documented in Figure 2, which shows bin scatter plots of a firm's sales, downstream access, and upstream access against the firm earnings effect f_{it} estimated from the worker earnings decomposition (2.1), where all variables are residualized by industry-municipality-year fixed effects. For reference, rows (iv) and (v) of Table 1 also provide moments of the distributions of D_{it}^{net} and S_{it}^{net}. As expected, we find that larger firms tend to pay higher wages. In addition, we see that firms with greater access to customers and suppliers in the production network have higher wage premia, which suggests an important role for production network heterogeneity in explaining earnings inequality.

[14] This decomposition is also studied by Bernard et al. (2022) using firm-to-firm trade data from Belgium.

Figure 2: Plots of sales and network access against firm earnings effects



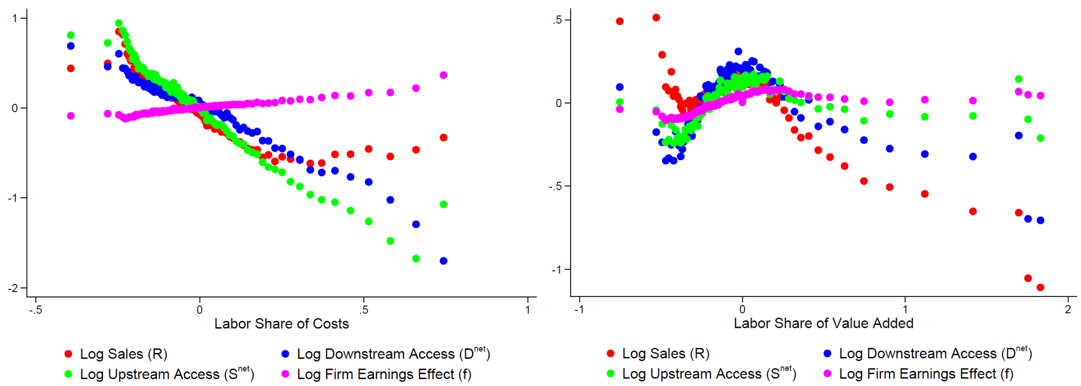
Notes: All plots are generated using the bin scatter program provided by Michael Stepner: https://michaelstepner.com/software. The firm earnings effect is measured as f_{it} from the worker earnings decomposition in equation (2.1). The network access measures D^{net} and S^{net} are as defined in (2.4). All variables are parsed of industry-municipality-year means.

Fact 2: Firms that have greater access to customers and suppliers in the production network tend to have lower labor shares of production costs and lower labor shares of value-added.

Preliminary evidence suggests that network heterogeneity also plays a key role in shaping differences in both labor shares of cost and value-added. This is documented in Figure 3, which plots firm sales, downstream access D_{it}^{net}, upstream access S_{it}^{net}, and firm earnings effects f_{it} against the two labor share measures defined above, where all variables are parsed of industry-municipality-year means. In the left panel of the figure, we see that firms with greater network access tend to have lower labor shares of cost. In the right panel, we see that the same is approximately true for labor shares of value-added as well. Although we observe some non-monotonicity in the middle of the labor share distribution, the overall correlation between labor shares of value-added and both our downstream and upstream access measures is negative. These patterns suggest an important role for heterogeneity in production network linkages in shaping these shares. In addition, we find a negative relationship between firm size and both labor shares, which is consistent with recent evidence documented by Autor et al. (2020) regarding the relationship between firm sales and labor shares of value-added.[15]

[15]Interestingly, we also find a weakly positive relationship between a firm's labor shares and its earnings effect. In comparison, Kehrig and Vincent (2021) find that average wages are essentially unrelated to a firm's labor share of value-added, although as we show in our structural model below, it is the firm earnings effect rather than the average wage that is relevant for a firm's labor shares.

Figure 3: Plots of sales and network access against various labor share measures



Notes: All plots are generated using the bin scatter program provided by Michael Stepner: https://michaelstepner.com/software. The firm earnings effect is measured as f_{it} from the worker earnings decomposition in equation (2.1). The network access measures D^{net} and S^{net} are as defined in (2.4). All variables are parsed of industry-municipality-year means.

3 A Model of Labor Markets and Production Networks

To investigate more carefully the relationship between production network linkages and earnings inequality, we now develop a structural model that can shed light on the stylized facts documented above. The economy is populated by a set of workers Ω^L and a set of firms Ω^F. Workers are heterogeneous in a characteristic that we refer to as *ability*, denoted by a, with an exogenous measure of each ability type denoted by $L(a)$ and the set of abilities denoted by A. Firms are also heterogeneous in a variety of characteristics that we specify below. Time is discrete and indexed by t.

3.1 Labor market

Firms and workers interact in the labor market as follows. Each firm i chooses wages $w_{it}(a)$ and offers exogenous amenities $g_i(a)$ for each worker of ability type a. In addition, workers derive idiosyncratic utility values ξ_{it} from employment at firm i, which are independent across workers and firms for a given t and follow a Gumbel distribution with cumulative distribution function $F_\xi(\xi_{it}) = e^{-e^{-\gamma\xi_{it}}}$, where the variance of the distribution is declining in the shape parameter γ.[16] Each worker then observes the wage offers and amenities corresponding to her ability and chooses an employer to maximize utility. Workers are also residual claimants to firm

[16] It is simple to allow for correlation in ξ_{it} across firms by instead assuming that the vector $\{\xi_{it}\}_{i\in\Omega^F}$ has joint cumulative distribution function given by $exp\left[-\left(\sum_{i\in\Omega^F} e^{-\rho\gamma\xi_{it}}\right)^{\frac{1}{\rho}}\right]$, where the correlation of the distribution is increasing in the parameter $\rho \in [1,\infty)$. This version of the model is observationally equivalent to the version with independent draws of ξ_{it} across firms, with $\rho\gamma$ replacing γ.

profits, which are rebated through transfers in proportion to labor income, so that the rebate received by a worker earning wage w is equal to $\tau_t w$.[17] Total income is then used to finance final consumption, which is a CES aggregate of products produced by all firms in the economy with elasticity of substitution σ across products.

Formally, the potential utility of a worker of ability a with a vector $\xi_t \equiv \{\xi_{it}\}_{i\in\Omega^F}$ of idiosyncratic utility values is given by:

$$u_t\left(a|\,\xi_t\right) = \max_{i\in\Omega^F} \left\{\log\left(1+\tau_t\right) w_{it}\left(a\right) + \log g_i\left(a\right) + \xi_{it}\right\} \tag{3.1}$$

where we treat the price of the CES final consumption aggregate as the numeraire so that all income is in real terms. As is well known, under the Gumbel distribution of idiosyncratic utilities, the measure of workers of ability a that choose employment at firm i is given by:

$$L_{it}\left(a\right) = \kappa_{it}\left(a\right) w_{it}\left(a\right)^{\gamma} \tag{3.2}$$

where $\kappa_{it}\left(a\right)$ is a firm-specific labor supply shifter:

$$\kappa_{it}\left(a\right) \equiv L\left(a\right)\left[\frac{g_i\left(a\right)}{I_t\left(a\right)}\right]^{\gamma} \tag{3.3}$$

and $I_t\left(a\right)$ is a labor market index summarizing the wages and amenities offered by all firms for workers of ability a:

$$I_t\left(a\right) \equiv \left[\sum_{i\in\Omega^F}\left[g_i\left(a\right) w_{it}\left(a\right)\right]^{\gamma}\right]^{\frac{1}{\gamma}} \tag{3.4}$$

We assume that the cardinality of the set of firms Ω^F is large enough such that each firm views itself as atomistic in the labor market and hence takes the labor market indices $I_t\left(\cdot\right)$ as given when choosing wages. Each firm therefore behaves as though it faces an upward-sloping labor supply curve with a constant elasticity γ that is common to all firms and worker ability types. Intuitively, labor supply is more sensitive to differences in wages when there is less dispersion in preference shocks across firms.[18]

[17] Rebating profits in proportion to labor income ensures that these transfers do not affect the sorting of workers across firms.

[18] Note that instead of arising from employer differentiation, labor market power could also stem from concentration (Chan et al. (2022), Berger et al. (2019), Jarosch et al. (2019)) or search frictions (Burdett and Mortensen (1998), Postel-Vinay and Robin (2002), Taber and Vejlin (2018)). Like ours, most of these models imply that wages are a markdown below the marginal revenue product of labor (MRPL) at a firm, where the firm effect on earnings is the component of the MRPL that is common to all workers at a firm. Hence, the mechanisms that we highlight below regarding the interaction between the production network and worker earnings are relevant for a broader class of models of the labor market.

3.2 Production technology

Firm i produces output X_{it} using labor and materials via the following production technology:

$$X_{it} = T_{it} \sum_{a \in A} F\left[\omega_{it}\phi_i(a) L_{it}, M_{it}(a)\right] \tag{3.5}$$

where T_{it} is total factor productivity (TFP), ω_{it} is labor productivity, $\phi_i(a)$ reflects worker-firm complementarities in production, and $M_{it}(a)$ is the quantity of materials assigned to workers of ability a.[19] The function F is a CES aggregator, $F(L, M) = \left(\lambda^{\frac{1}{\epsilon}} L^{\frac{\epsilon-1}{\epsilon}} + (1-\lambda)^{\frac{1}{\epsilon}} M^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon-1}{\epsilon}}$, where λ controls the importance of labor relative to materials in production and ϵ is the elasticity of substitution between labor and materials.

While firms hire workers in the labor market by posting wages, materials are sourced through firm-to-firm trade in the production network. In particular, firm i produces a materials bundle by combining inputs from all of its suppliers $\Omega_{it}^S \subset \Omega^F$ using a CES technology, so that the total quantity of materials used in production $M_{it} \equiv \sum_{a \in A} M_{it}(a)$ satisfies:

$$M_{it} = \left[\sum_{j \in \Omega_{it}^S} \psi_{ijt}^{\frac{1}{\sigma}} \left(x_{ijt}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \tag{3.6}$$

where x_{ijt} denotes the quantity of inputs purchased by i from j and ψ_{ijt} is a relationship-specific productivity shifter. We assume that the latter can be decomposed as:

$$\psi_{ijt} = \psi_{it}\psi_{jt}\tilde{\psi}_{ijt} \tag{3.7}$$

where we refer to ψ_{it} as the *relationship capability* of firm i and $\tilde{\psi}_{ijt}$ as the relationship productivity residual. This decomposition allows firms to differ systematically in the produtivity of their buyer-seller relationships. As is standard in the literature, we also assume the same elasticity of substitution σ across products in production as in final consumption, which simplifies the firm's profit maximization problem as it ensures that both final and intermediate demand have the same price elasticity.

We highlight several important features of the production technology. First, although it is straightforward to incorporate imperfect substitutability between workers of different abilities, the linear aggregation across worker types in equation (3.5) is necessary for the model to generate an earnings equation that is consistent with well-known reduced-form models of earnings such as those in Abowd et al. (1999) and Bonhomme et al. (2019). Second, although we allow

[19] One can also think of certain types of capital inputs as sourced from suppliers in the production network under the label of "materials" if these inputs are chosen statically. Alternatively, it is straightforward to extend the production function to allow for a separate static capital input. See Appendix B for a formal discussion of this extension.

for time-varying labor productivities ω_{it}, we restrict worker-firm complementarities $\phi_i(\cdot)$ to be time-invariant, which will be important for identification of these terms. Third, in the limit as $\lambda \to 1$, output is produced using labor alone and the model simplifies to a version of the model studied in Lamadon et al. (2022). Finally, the production network is not restricted to be bipartite: firms can simultaneously be buyers and sellers, with $\Omega_{it}^C \equiv \left\{ j \in \Omega^F | i \in \Omega_{jt}^S \right\}$ denoting the set of customers for firm i. However, for tractability, we treat the set of active buyer-seller relationships in the economy as an exogenous primitive of the model and do not model network formation.[20] Nonetheless, this imposes no restrictions on how the distribution of buyer-seller links is correlated with other firm primitives or how the network changes over time. For example, more productive firms may have more links and add links at a faster rate than less productive firms. As we discuss below, our identification of model parameters does not require such restrictions.

3.3 Price setting

Firms are monopolistically competitive in output markets, setting prices for their customers while taking the prices set by other firms as given. As with firm behavior in labor markets, we assume that firms behave atomistically in output markets and hence perceive a constant price elasticity of demand equal to $-\sigma$. Note that a firm's relationships with each of its customers are inherently interlinked: a reduction in the price charged to one customer increases demand and hence raises both output and marginal cost, which in turn affects the choice of prices charged to other customers. However, even though we allow firms to charge different prices to different customers, the following result establishes that it is never optimal for them to do so.[21]

Claim **1.** *The profit-maximizing price charged by a firm i to each of its customers (including final consumers) does not vary across customers:*

$$p_{jit} = p_{it}, \, \forall j \in \Omega_{it}^C \cup \{F\} \tag{3.8}$$

Intuitively, each firm maximizes profits by choosing prices such that marginal revenue from each customer is equal to marginal cost. Since demand features a constant and common price elasticity of $-\sigma$, marginal revenue is proportional to price. Furthermore, even though marginal cost is increasing, it depends only on total output of the firm and hence is common across customers. As a result, each firm optimally chooses to charge a common price to each of its customers in equilibrium.

[20] Existing models of endogenous production network formation such as those in Huneeus (2019) and Lim (2019) require that marginal production costs are independent of output, so that the decision of a firm to sell to one customer can be analyzed independently of who else the firm sells to. This assumption is violated in our model due to the upward-sloping labor supply curves, which generates marginal costs that are increasing with output.

[21] Proofs of all claims and propositions are relegated to Section D of the appendix.

3.4 Firm network characteristics

With the result in equation (3.8), we can express total sales for firm i as:

$$R_{it} \equiv p_{it} X_{it} = D_{it} p_{it}^{1-\sigma} \tag{3.9}$$

where D_{it} is a demand shifter. This is given by:

$$D_{it} = E_t + \sum_{j \in \Omega_{it}^C} \Delta_{jt} \psi_{jit} \tag{3.10}$$

where E_t denotes aggregate consumer expenditure and Δ_{jt} is a firm-specific intermediate demand shifter that we refer to as the *buyer effect*:

$$\Delta_{jt} = E_{jt}^M \left(Z_{jt}\right)^{\sigma-1} \tag{3.11}$$

with $E_{jt}^M \equiv Z_{jt} M_{jt}$ denoting total material cost and Z_{jt} denoting the unit cost of materials corresponding to the CES materials bundle in equation (3.6). The unit cost of materials for firm i can in turn be expressed as:

$$Z_{it} = \left[\sum_{j \in \Omega_{it}^S} \psi_{ijt} \Phi_{jt} \right]^{\frac{1}{1-\sigma}} \tag{3.12}$$

where Φ_{jt} is an inverse measure of supplier j's output price that we refer to as the *seller effect*:

$$\Phi_{jt} \equiv p_{jt}^{1-\sigma} \tag{3.13}$$

Note that the relevance of the production network for firm i's production decisions is summarized by the sufficient statistics $\{D_{it}, Z_{it}\}$, which we henceforth refer to as the *network characteristics* of the firm. The demand shifter D_{it} summarizes firm i's downstream connections with its customers (including final consumers), while the unit cost of materials Z_{it} summarizes its upstream connections with suppliers.

3.5 Profit maximization and wage setting

The profit-maximization problem for firm i can now be described as:

$$\pi_{it} = \max_{\{w_{it}(a), M_{it}(a)\}_{a \in A}} \left\{ D_{it}^{\frac{1}{\sigma}} X_{it}^{\frac{\sigma-1}{\sigma}} - \sum_{a \in A} w_{it}(a) L_{it}(a) - Z_{it} \sum_{a \in A} M_{it}(a) \right\} \tag{3.14}$$

where the maximization is subject to the labor supply curves (3.2) and production technology (3.5). Since the price of materials is invariant with respect to worker ability, the marginal revenue product of materials must first of all be equalized across worker ability types in equilibrium, as implied by the first-order condition for (3.14) with respect to materials:

$$Z_{it} = \frac{1}{\mu} D_{it}^{\frac{1}{\sigma}} X_{it}^{-\frac{1}{\sigma}} T_{it} F_M \left(1, \nu_{it}\right) \tag{3.15}$$

where F_M denotes the derivative of F with respect to its second argument and $\nu_{it} \equiv \frac{M_{it}(a)}{\phi_i(a)\omega_{it}L_{it}(a)}$ is materials per efficiency unit of labor, which does not vary by worker ability.

The first-order condition for (3.14) with respect to $w_{it}(a)$ then allows us to express equilibrium wages as:

$$w_{it}(a) = \eta \phi_i(a) W_{it} \tag{3.16}$$

Equation (3.16) states the familiar result that wages are a constant markdown $\eta \equiv \frac{\gamma}{1+\gamma} \in (0,1)$ over the marginal revenue product of labor (MRPL) of the respective worker types, $\phi_i(a) W_{it}$. The component of wages that is common to all workers employed at firm i, W_{it}, is given by:

$$W_{it} = \frac{1}{\mu} D_{it}^{\frac{1}{\sigma}} X_{it}^{-\frac{1}{\sigma}} \omega_{it} T_{it} F_L \left(1, \nu_{it}\right) \tag{3.17}$$

where F_L denotes the derivative of F with respect to its first argument and we define the output markup $\mu \equiv \frac{\sigma}{\sigma-1}$ for brevity. We henceforth refer to W_{it} as the *firm earnings effect.*[22] Note that in the limit as labor supply becomes infinitely elastic ($\gamma \to \infty$), the markdown η approaches unity as in the benchmark with perfectly competitive labor markets.

Equilibrium output for firm i can then be characterized as follows:

$$X_{it} = T_{it} F\left(1, \nu_{it}\right) \bar{L}_{it} \tag{3.18}$$

$$\bar{L}_{it} = \left(\eta W_{it}\right)^{\gamma} \omega_{it} \bar{\phi}_{it} \tag{3.19}$$

$$\bar{\phi}_{it} \equiv \sum_{a \in A} \kappa_{it}(a) \phi_i(a)^{1+\gamma} \tag{3.20}$$

while expenditures on labor and materials are given respectively by:

$$E_{it}^{L} = \eta W_{it} \bar{L}_{it} / \omega_{it} \tag{3.21}$$

$$E_{it}^{M} = Z_{it} \nu_{it} \bar{L}_{it} \tag{3.22}$$

where $\bar{L}_{it} \equiv \sum_{a \in A} \phi_i(a) \omega_{it} L_{it}(a)$ is the total efficiency units of labor hired by the firm and we define the $\bar{\phi}_{it}$ as the *sorting composite* for firm i, since this varies across firms only due to primitives that affect differential sorting of worker types across firms ($g_i(\cdot)$ and $\phi_i(\cdot)$).

[22] As we show below, W_{it} is equivalent to ηf_{it}, where f_{it} is the firm effect in the reduced-form decomposition (2.1) presented in section 2.

Given the firm's technological primitives $\{T_{it}, \omega_{it}\}$, network characteristics $\{D_{it}, Z_{it}\}$, and sorting composite $\bar{\phi}_{it}$, equations (3.15), (3.17), (3.18), and (3.19) define a system of equations in the firm-level variables $\left\{W_{it}, \nu_{it}, X_{it}, \bar{L}_{it}\right\}$. In particular, the production network shapes earnings through the dependence of W_{it} on D_{it} and Z_{it}.

3.6 General equilibrium

To close the model, it remains to characterize total consumer expenditure E_t, which is equivalent to aggregate value-added. This is simply given by the sum of labor income and firm profits:

$$E_t = \sum_{i \in \Omega^F} \left(E_{it}^L + \pi_{it}\right) \tag{3.23}$$

We can then define the primitives and an equilibrium of the model as follows.

Definition **1** (model primitives). *The primitives of the model at time t, Θ_t, are TFPs T_{it}, labor productivities ω_{it}, relationship capabilities ψ_{it}, relationship productivity residuals $\tilde{\psi}_{ijt}$, buyer-supplier linkages Ω_{it}^S, production complementarities $\phi_i(\cdot)$, amenities $g_i(\cdot)$, the worker ability distribution $L(\cdot)$, elasticities $\{\gamma, \epsilon, \sigma\}$, and the weight on labor in production λ.*

Definition **2** (equilibrium). *Given a set of primitives, an equilibrium of the model at time t is a set of values for aggregate value-added E_t, demand shifters D_{it}, material costs Z_{it}, buyer effects Δ_{it}, seller effects Φ_{it}, firm earnings effects W_{it}, output prices p_{it}, output X_{it}, labor efficiency units $\bar{L}_{it}$, sorting composites $\bar{\phi}_i$, labor expenditures E_{it}^L, material expenditures E_{it}^M, wages $w_{it}(\cdot)$, labor supply shifters $\kappa_{it}(\cdot)$, and labor market indices $I_t(\cdot)$, all of which satisfy equations (3.3), (3.4), (3.9), (3.10), (3.11), (3.12), (3.13), (3.14), (3.15) (3.17), (3.18), (3.19), (3.20), (3.21), (3.22), and (3.23).*

4 Equilibrium Analysis

We now use the model to shed light on the stylized facts presented in section 2 by providing a theoretical characterization of how the production network shapes worker earnings, the firm size wage premium, and labor shares of value-added and cost. This analysis will also provide context for the empirical results that follow.

4.1 Structural interpretation of reduced-form network access measures

We begin by providing a structural interpretation of the network access measures D_{it}^{net} and S_{it}^{net} presented in section 2. First, log sales between buyer i and seller j in our model can be expressed as:

$$\log r_{ijt} = \log\left(\Delta_{it}\psi_{it}\right) + \log\left(\Phi_{jt}\psi_{jt}\right) + \log \tilde{\psi}_{ijt} \tag{4.1}$$

Comparing this with the reduced-form decomposition of firm-to-firm sales (2.3), we hence have $d_{it} \equiv \Delta_{it}\psi_{it}$, $s_{jt} \equiv \Phi_{jt}\psi_{jt}$, and $e_{ijt} \equiv \tilde{\psi}_{ijt}$. Therefore, in our model, the network access measures defined in equation (2.4) are given by:

$$D_{it}^{net} \equiv \sum_{j \in \Omega_{it}^{C}} \Delta_{jt}\psi_{jt}\tilde{\psi}_{jit}, \qquad\qquad S_{it}^{net} \equiv \sum_{j \in \Omega_{it}^{S}} \Phi_{jt}\psi_{jt}\tilde{\psi}_{ijt} \tag{4.2}$$

In other words, a firm's downstream access D_{it}^{net} is increasing in the buyer effects Δ_{jt} and relationship capabilities ψ_{jt} of its customers, while its upstream access S_{it}^{net} is increasing in the seller effects Φ_{jt} and relationship capabilities ψ_{jt} of its suppliers. Furthermore, from equations (3.10) and (3.12), we can express a firm's demand shifter as:

$$D_{it} = E_t + \psi_{it} D_{it}^{net} \tag{4.3}$$

and its unit cost of materials as:

$$Z_{it} = \left(\psi_{it} S_{it}^{net}\right)^{\frac{1}{1-\sigma}} \tag{4.4}$$

Hence, the network access measures $\{D_{it}^{net}, S_{it}^{net}\}$ are sufficient statistics for how the production network shapes $\{D_{it}, Z_{it}\}$ and therefore the firm effect on earnings W_{it}.

4.2 The firm effect on earnings

Recall that firms with better downstream access (higher D_{it}^{net} and thus higher D_{it}) and better upstream access (higher S_{it}^{net} and thus lower Z_{it}) also tend to have higher firm effects on earnings (Fact 1). Here, we discuss the key mechanisms in our model that rationalize these observed relationships.

4.2.1 Case with Cobb-Douglas technology

We begin by examining a special case of the model where production technologies are of the Cobb-Douglas form ($\epsilon \to 1$). This case admits a closed-form solution for the firm effect on earnings and hence is useful for providing a more transparent discussion of the key mechanisms. At the same time, we highlight properties of the model that obtain *only* under Cobb-Douglas technology, which underscores the importance of proper identification of ϵ.

Under Cobb-Douglas technology, the firm's profit maximization problem (3.14) can be rewritten by first solving out for the optimal choice of material inputs:

$$\max_{\{w_{it}(a)\}_{a \in A}} \left\{ A_{it}\tilde{X}_{it}^{1-\alpha} - \sum_{a \in A} w_{it}(a) L_{it}(a) \right\} \tag{4.5}$$

$$\text{s.t. } \tilde{X}_{it} = \sum_{a \in A} \phi_{it}(a) L_{it}(a) \tag{4.6}$$

where $A_{it}\tilde{X}_{it}^{1-\alpha}$ is equal to nominal value-added for firm i, $\alpha \equiv \frac{1}{\sigma\lambda+(1-\lambda)} > 0$ reflects curvature in value-added arising from imperfectly elastic demand ($\sigma < \infty$), and A_{it} is a composite term that can be interpreted as *value-added productivity*:

$$A_{it} \equiv \text{const.} \times T_{it}^{\frac{\sigma-1}{\sigma\lambda+1-\lambda}} \omega_{it}^{\frac{\lambda(\sigma-1)}{\sigma\lambda+1-\lambda}} D_{it}^{\frac{1}{\sigma\lambda+1-\lambda}} Z_{it}^{-\frac{(1-\lambda)(\sigma-1)}{\sigma\lambda+1-\lambda}} \tag{4.7}$$

Equations (4.5) and (4.6) hence represent the firm's profit maximization problem in terms of a value-added production function, which is a common approach in the literature (see Lamadon et al. (2022), for example). The firm effect W_{it} can then be solved for explicitly as:

$$W_{it} = \text{const.} \times A_{it}^{\frac{\sigma\lambda+1-\lambda}{\gamma+\sigma\lambda+1-\lambda}} \bar{\phi}_{it}^{-\frac{1}{\gamma+\sigma\lambda+1-\lambda}} \tag{4.8}$$

This special case of the model allows for several important takeaways.

First, from equation (4.7), demand and supply shocks in the network that operate through $\{D_{it}, Z_{it}\}$ act as shifters of value-added productivity A_{it}. In this sense, the introduction of production networks provides a microfoundation for value-added productivity. From equation (4.8), it is then immediately clear that increases in demand D_{it} and reductions in material costs Z_{it} both lead to increases in the earnings effect W_{it}. Second, it is immediately obvious from equations (4.7) and (4.8) that without further information, identification of A_{it} alone does not allow one to separately identify the components of A_{it} (and hence of the firm effect W_{it}) that stem from TFP, labor productivity, and network characteristics. Hence, the value-added approach naturally leaves open the question of how heterogeneity in production network linkages shapes earnings inequality, which we examine in section 7.

Finally, note that the above discussion only applies when $\epsilon = 1$. When this condition does not hold, the firm's profit maximization problem generally does not admit the representation in equations (4.5) and (4.6), and hence the concept of value-added productivity is no longer meaningful.[23] In our empirical estimates below, we find that ϵ is statistically greater than 1. Hence, while the Cobb-Douglas case is useful as a heuristic for developing the intuition behind the model, it remains a simplification that is unsupported by our data.[24]

4.2.2 General case

Beyond the case of Cobb-Douglas technology, one cannot solve for W_{it} as a closed-form function of a firm's network characteristics $\{D_{it}, Z_{it}\}$, technological primitives $\{T_{it}, \omega_{it}\}$, and sorting composite $\bar{\phi}_{it}$. Nonetheless, comparative static results are useful for highlighting the mechanisms

[23] The value-added representation is valid for any ϵ in the limit as $\sigma \to \infty$. This case is not empirically relevant since it corresponds to perfect competition in output markets.

[24] The assumption of Cobb-Douglas technology also implies complete passthrough of changes in firm value-added per worker to changes in worker earnings, which is at odds with existing empirical evidence. In contrast, our model implies incomplete passthrough outside of the Cobb-Douglas case. See Appendix C for details.

that shape differences in firm effects. To focus on the role played by the production network, we examine here the passthrough of demand and material cost shocks in the network into changes in earnings, which we will evaluate empirically below in section 5 (see Appendix D.2 for a full discussion of comparative statics with respect to T_{it}, ω_{it}, and $\bar{\phi}_{it}$).

First, in what follows, let $\hat{X}_{it}$ denote the marginal log change in any variable X and let $\mathbf{X}_t \equiv \{X_{it}\}_{i\in\Omega^F}$ denote a vector of firm-specific variables. Furthermore, let $s_{it}^M \equiv \frac{E_{it}^M}{\frac{1}{\eta}E_{it}^L + E_{it}^M}$ denote the share of materials in firm i's production costs adjusted for wage markdowns and let $\left\{S_t^{sales}, S_t^{mat}\right\}$ denote the sales share and material cost share matrices, with (i, j)-elements equal to the share of firm i's sales accounted for by firm j and the share of firm i's material expenditures accounted for by firm j, respectively.

Now, consider a vector of *exogenous* shocks $\left\{\hat{\mathbf{D}}_t, \hat{\mathbf{Z}}_t\right\}$ to demand and material costs for all firms in the production network.[25] The following Proposition then summarizes how these shocks affect wages through the firm effects on earnings. As we discuss in Appendix D.2, these results abstract from feedback effects arising from the fact that marginal costs are scale dependent due to upward-sloping labor supply curves, which are second-order and likely to be small empirically.

PROPOSITION **1.** *The first-order effects of demand and material cost shocks $\left\{\hat{\mathbf{D}}_t, \hat{\mathbf{Z}}_t\right\}$ in the production network on firm i's earnings effect are given by:*

$$\hat{W}_{it} = \Gamma_{it}\hat{D}_{it}^{total}, \qquad \hat{W}_{it} = -\left(\sigma - \epsilon\right)\Gamma_{it}s_{it}^M\hat{Z}_{it}^{total} \tag{4.9}$$

where $\Gamma_{it} \equiv \frac{1}{\gamma+\sigma\left(1-s_{it}^M\right)+\epsilon s_{it}^M} > 0$ is the scale elasticity for firm i and firms' total demand and material cost shocks are defined as:

$$\hat{\mathbf{D}}_t^{\mathbf{total}} \equiv \left[I - S_t^{sales}\delta_t^D\right]^{-1}\hat{\mathbf{D}}_t, \qquad \hat{\mathbf{Z}}_t^{\mathbf{total}} \equiv \left[I - S_t^{mat}\delta_t^U\right]^{-1}\hat{\mathbf{Z}}_t \tag{4.10}$$

where δ_t^D and δ_t^U are diagonal matrices of downstream and upstream weights with (i, i)-elements given by $\delta_{it}^D \equiv \frac{\gamma+\epsilon}{\gamma+\sigma\left(1-s_{it}^M\right)+\epsilon s_{it}^M}$ and $\delta_{it}^U \equiv \frac{\gamma+\epsilon}{\gamma+\sigma\left(1-s_{it}^M\right)+\epsilon s_{it}^M} \cdot s_{it}^M$.

The intuition for these results can be understood as follows. First, $\hat{D}_{it}^{total}$ and $\hat{Z}_{it}^{total}$ are sufficient statistics summarizing a firm's total exposure to demand and material cost shocks in the production network, where the Leontief inverses of the weighted sales and material cost share matrices $S_t^{sales}\delta_t^D$ and $S_t^{mat}\delta_t^U$ account for own-firm shocks, shocks to direct customers and suppliers, as well as shocks to indirect customers and suppliers (those connected via other customers and suppliers). Intuitively, the importance of each customer for a firm's total exposure to demand shocks depends on sales shares, while the importance of each supplier for a firm's total exposure to material cost shocks depends on material cost shares.

[25] In our empirical analysis below, we will treat these as arising from export demand and import cost shocks, which are outside the model.

Second, demand shocks have a positive effect on firm earnings effects ($\Gamma_{it} > 0$). This occurs because higher demand raises the output price of a firm, which translates into a higher MRPL and hence higher wages given the upward-sloping labor supply curves faced by each firm. The scale elasticity Γ_{it} summarizes the three conditions that are necessary for the existence of such scale effects: (i) the labor market is imperfectly competitive ($\gamma < \infty$), so that marginal costs of labor are increasing; (ii) the output market is imperfectly competitive ($\sigma < \infty$), so that higher demand raises output prices at the level of the firm; and (iii) labor and materials are imperfect substitutes ($\epsilon < \infty$), so that firms cannot fully escape from increasing marginal costs of labor by purchasing materials at constant marginal cost.

Third, the sign of the relationship between the firm effect and material cost shocks depends on the sign of $\sigma - \epsilon$. This is because a change in material cost has both a scale and a substitution effect on W_{it}. On one hand, a higher cost of materials is akin to a negative productivity shock, which induces the firm to contract in scale and hire fewer workers at lower wages. The strength of this scale effect is increasing in the parameter σ, since a firm's scale is more sensitive to changes in production costs when products are more differentiated. On the other hand, an increase in the cost of materials induces firms to substitute away from materials toward labor, which may either increase wages (if $\epsilon > 1$, so that labor and materials are substitutes) or decrease wages (if $\epsilon < 1$, so that labor and materials are complements). Hence, the net effect of material cost shocks on W_{it} depends on the relative magnitude of σ versus ϵ. In our estimation of the model's parameters below, we find that $\sigma > \epsilon$ and hence higher material costs induce lower wages.

Finally, the downstream and upstream weights $\left\{\delta_t^D, \delta_t^U\right\}$ in (4.10), which are strictly less than 1 if $\sigma > \epsilon$, reflect the fact that demand and material cost shocks are only incompletely passed through into demand for inputs from suppliers and into output prices charged to customers respectively. When a customer experiences an increase in demand, this not only raises demand for materials, but also raises the wages paid by the firm, hence increasing the firm's marginal cost of output and partially offsetting the increase in its material demand. Similarly, when a supplier faces a higher cost of materials, this not only increases the firm's marginal cost of output, but also reduces the wages that the firm pays (given $\sigma > \epsilon$), which partially offsets the increase in marginal cost.

In sum, increases in demand for a firm or its customers lead to higher wages, while increases in material costs for a firm or its suppliers reduce wages. An immediate corollary is that, all else equal, firms with higher values of demand D_{it} and lower material costs Z_{it} will pay higher wages, which provides a theoretical rationalization for the relationships described in Fact 1. In section 5, we provide a validation of the mechanisms highlighted above by estimating the passthrough of export demand and import cost shocks into changes in earnings.

4.3 Labor shares of value-added and cost

Recall that firms with greater downstream and upstream access in the production network tend to have lower labor shares of value-added and lower labor shares of cost (Fact 2). To see how these patterns might arise in our model, first note that labor cost shares, $s_{it}^L \equiv 1 - s_{it}^M$, are completely determined by the productivity-adjusted cost of labor relative to materials:

$$s_{it}^L = 1 - \left[1 + \frac{\lambda}{1-\lambda}\left(\frac{W_{it}/\omega_{it}}{Z_{it}}\right)^{1-\epsilon}\right]^{-1} \tag{4.11}$$

As discussed below, we estimate that $\epsilon > 1$, so that the labor cost share is *declining* in the relative price $\frac{W_{it}}{Z_{it}}$. Furthermore, as established in Proposition 1, increases in downstream and upstream access in the network raise W_{it} (given that we estimate $\sigma > \epsilon$), while greater upstream access directly lowers Z_{it}. Hence, firms with greater downstream and upstream network access tend to have lower labor shares of cost because they tend to face higher costs of labor relative to materials. Furthermore, since larger firms tend to have greater downstream and upstream access in the network, production network heterogeneity tends to *amplify* the negative relationship between labor cost shares and firm size.

As for labor shares of value-added, $s_{it}^{L/VA} \equiv \frac{E_{it}^L}{R_{it}-E_{it}^M}$, first note that a simple identity allows us to write:

$$s_{it}^{L/VA} \equiv \frac{\eta s_{it}^L}{R_{it}/TC_{it} - (1 - s_{it}^L)} \tag{4.12}$$

where $TC_{it} \equiv \frac{1}{\eta}E_{it}^L + E_{it}^M$ is a measure of the firm's total production cost adjusted for markdowns on wages. As a matter of accounting, a firm's labor share of value-added is lower when, *ceteris paribus*, labor makes up a smaller share of its production cost or its sales-cost ratio is large (since the latter implies that more revenue is allocated to firm profits). It is straightforward to show that the sales-cost ratio R_{it}/TC_{it} in our model is simply equal to the output markup μ. Hence, labor shares of value-added vary across firms only in relation to differences in input cost shares. In particular, firms with lower labor cost shares also have lower labor shares of value-added. Consequently, firms with greater downstream and upstream access in the production network tend to have lower labor shares of value-added, while production network heterogeneity also tends to *amplify* the negative relationship between labor value-added shares and firm size.

4.4 The firm size wage premium

As discussed in the motivation, larger firms tend to have higher firm effects on earnings (Figure 2). To examine the role of production network linkages in driving this relationship, first note that we can express a firm's earnings effect in terms of its sales, sorting composite, and labor cost share:

$$\log W_{it} = \text{const.} + \frac{1}{1+\gamma}\log R_{it} - \frac{1}{1+\gamma}\log\bar{\phi}_{it} - \frac{1}{1+\gamma}\log s_{it}^{L} \tag{4.13}$$

The covariance between firm earnings effects and firm size can therefore be expressed as:

$$\begin{aligned}\text{cov}\left(\log W_{it}, \log R_{it}\right) = \frac{1}{1+\gamma}\text{var}\left(\log R_{it}\right) - \frac{1}{1+\gamma}\text{cov}\left(\log\bar{\phi}_{it}, \log R_{it}\right) \\ - \frac{1}{1+\gamma}\text{cov}\left(\log s_{it}^{L}, \log R_{it}\right)\end{aligned} \tag{4.14}$$

Hence, conditioning on the variance in firm sales, the firm size wage premium is stronger if: (i) larger firms have smaller sorting composites (since these firms must then pay higher wages to compensate for poorer amenities, for example); or (ii) larger firms tend to have smaller labor cost shares. It is through the latter channel that production network linkages play a key role. As discussed in the preceding section, larger firms tend to have greater downstream and upstream access in the network, which tends to generate a negative relationship between the labor cost share and firm size. Consequently, production network heterogeneity *amplifies* the firm size wage premium.

In addition, note from equation (4.13) that firm size is generally not a sufficient statistic for the firm effect on earnings. Even if one accounts for differences in sorting through $\bar{\phi}_{it}$ (as, for example, in Lamadon et al. (2022)), differences in material cost shares due to heterogeneity in the production network still contribute to differences in worker earnings conditional on sales (as long as $\epsilon \neq 1$). This further implies that unpacking the determinants of heterogeneity in firm size (as in Bernard et al. (2022), for example) is complementary but not equivalent to unpacking the determinants of heterogeneity in worker earnings.

5 Reduced-form Passthrough Evidence

Before taking our model to data, we first provide reduced-form evidence to validate some of its key mechanisms. We focus here on the predictions encapsulated in Proposition 1, which characterize how demand and material cost shocks in the production network affect worker earnings.

To test these predictions, we require firm-level demand and material cost shocks $\left\{\hat{D}_{it}, \hat{Z}_{it}\right\}$. To provide validation that is external to the model, we rely on Bartik shift-share shocks to export demand and import costs for Chilean firms that participate directly in exporting and importing. In addition, Appendix G describes a simple extension of our model that rationalizes the following approach to constructing these export and import shocks.

First, we define an international trade market m as a product-by-foreign-country pair. We then construct the following shift-share shocks to export demand and import costs:

$$\hat{\Delta}^X_{it} \equiv \sum_{m\in\Omega^{M,X}_{i1}} s^X_{mi1}\hat{s}^I_{mt}, \qquad \hat{p}^I_{it} \equiv -\sum_{m\in\Omega^{M,I}_{i1}} s^I_{im1}\hat{s}^X_{mt} \tag{5.1}$$

where $\Omega^{M,X}_{i1}$ ($\Omega^{M,I}_{i1}$) denotes the markets in which firm i actively exports (imports) in the first year of our sample, s^X_{mi1} (s^I_{im1}) denotes the share of firm i's exports (imports) accounted for by market m in the first year of our sample, and $\hat{s}^I_{mt}$ ($\hat{s}^X_{mt}$) denotes the annual log change in market m's share of world imports (exports) excluding trade with Chile within the corresponding product category. Intuitively, if a Chilean firm initially exports to markets that subsequently become more important sources of global demand for imports, we interpret this as an increase in export demand for the Chilean firm. Similarly, if a Chilean firm initially imports from markets that subsequently become more important suppliers of global exports, we interpret this as a decline in the cost of imports for the Chilean firm.

To translate export demand and import cost shocks into overall demand and material cost shocks, we weight these respectively by the share of each firm's sales accounted for by exports (s^{sales}_{Xit}) and the share of each firm's material cost accounted for by imports (s^{mat}_{iIt}):

$$\hat{D}_{it} \equiv s^{sales}_{Xit}\hat{\Delta}^X_{it}, \qquad \hat{Z}_{it} \equiv s^{mat}_{iIt}\hat{p}^I_{it} \tag{5.2}$$

We then construct each firm's total exposure to these shocks $\left\{\hat{D}^{total}_{it}, \hat{Z}^{total}_{it}\right\}$ following equation (4.10). Finally, we estimate the following regression via OLS using long differences between 2005 and 2010:

$$\hat{Y}_{it} = \alpha_D\hat{D}^{total}_{it} + \alpha_Z\hat{Z}^{total}_{it} + f_{\text{ind}(i)} + \zeta_{it} \tag{5.3}$$

where $\hat{Y}_{it}$ is the log change in an outcome of interest Y_{it}, $f_{\text{ind}(i)}$ is an industry fixed effect corresponding to the industry $\text{ind}(i)$ of firm i, and ζ_{it} is a residual that accounts for changes in Y_{it} arising from shocks other than our constructed demand and material cost shocks (for example, fluctuations in TFP and labor productivities, which by construction are orthogonal to our regressors).

Note that even though Proposition 1 concerns changes in firm effects W_{it}, firm wage bills are also affected by demand and material cost shocks only through changes in firm effects (since $E^L_{it} = (\eta W_{it})^{1+\gamma}\,\bar{\phi}_{it}$). Since we can measure wage bills directly in the data whereas firm effects must be estimated, we test the model's predictions using changes in firm wage bills as the main outcome of interest. Proposition 1, together with our structural estimates of σ and ϵ, then implies the following sign restrictions: $\alpha_D > 0$ (higher demand raises wage bills) and $\alpha_Z < 0$ (higher material costs lower wage bills, given $\sigma > \epsilon$).[26]

[26]From the expressions in (4.9), the regression coefficients α_D and α_Z can also be interpreted as average values

Column (1) of Table 2 shows our estimation results treating Y_{it} as a firm's wage bill. For comparison, we also include results for specifications where Y_{it} is a firm's average wage (column (2)) and sales (column (3)). We highlight the following takeaways. First, we estimate positive and statistically significant effects of demand shocks on firm wage bills, average wages, and sales (row A).[27] Second, we estimate negative and statistically significant effects of cost shocks on firm wage bills, average wages, and sales (row B). In sum, we find evidence of the passthrough of demand and cost shocks into changes in earnings that is broadly consistent with the predictions of our estimated structural model.

Table 2: Reduced-form passthrough estimates

	(1) Wage Bill	(2) Employment	(3) Sales
A. demand shocks, α_D	0.357	0.283	0.521
	(0.128)	(0.109)	(0.127)
B. cost shocks, α_Z	0.413	0.466	0.426
	(0.127)	(0.108)	(0.126)
sector fixed effects	yes	yes	yes
N	27,696	27,696	27,696

Notes: This table presents our estimates of the passthrough coefficients in equation (5.3) for different outcome variables $\hat{Y}_{it}$. Each of the outcome variables in columns (1)-(3) are measured in logs. Standard errors are shown in parentheses.

6 Connecting the Model to Data

We now turn towards identification and estimation of the model's parameters. We first describe the main sources of data used in our empirical analysis (section 6.1). We then describe a set of additional assumptions that are helpful for identification (section 6.2), before presenting our identification strategy and estimation results (section 6.3).

6.1 Data sources

We use four administrative datasets from the Internal Revenue Service (IRS, or SII for its acronym in Spanish) in Chile, covering the entire formal private sector. In each of the datasets described below, firms are assigned a unique tax ID, which facilitates the merging of these datasets. Hence, in what follows, we define a firm as a tax ID.[28] In addition, we convert all

of the structural terms $(1+\gamma)\,\Gamma_{it}$ and $-(\sigma-\epsilon)\,(1+\gamma)\,\Gamma_{it}s_{it}^{M}$ respectively, where the additional factor of $1+\gamma$ reflects the elasticity of wage bills with respect to firm effects.

[27] We also note that evidence of positive passthrough from both own demand shocks and customer demand shocks has been documented by Dhyne et al. (2021) using data from Belgium.

[28] As all tax forms are reported at the headquarter-level, plant-level information is not available. Furthermore, while it is possible that a firm has several tax IDs, information that allows us to observe firm ownership is not

nominal variables to real 2015 dollars.

Matched employer-employee data. These data are obtained from IRS tax affidavits 1887 and 1879, and report annual earnings from each job that a worker has from 2005-2018, including wages, salaries, bonuses, tips, and other sources of labor income deemed taxable by the IRS. As earnings are reported net of social security payments, we adjust the earnings measure to include these payments.

Firm-to-firm trade data. These data are obtained from IRS tax forms 3323 and 3327, and are based on value-added tax (VAT) records from 2005-2010. Each firm in this dataset reports the full list of its intermediate buyers and suppliers, as well as the total gross value of transactions with each buyer and supplier. As reporting occurs semi-annually, we aggregate this data to the annual level to make it consistent with the other datasets. Since this dataset reports transactions gross of taxes, we measure transactions net of taxes by using the flat value-added tax rate of 19% that was in effect in Chile during the sample period.

Worker age data. We obtain the year of birth of each individual who is alive in 2018 from a civil registry database. Merging this dataset with the employer-employee dataset using workers' unique tax IDs gives us a measure of the age of every worker. As we discuss below, we use these data to remove age effects from worker earnings.

Firm balance sheet data. These data are obtained from IRS tax form 29 and contain a set of firm balance sheet characteristics from 2005-2010. We use this dataset to measure total sales and material cost for each firm.

To prepare the data for use in our empirical analysis, we first clean the firm-to-firm trade and employer-employee data following procedures described in detail in Appendix E.1. We refer to these cleaned datasets respectively as the *baseline firm-to-firm dataset* and the *baseline employer-employee dataset*. The former contains 32 million firm-to-firm-year observations and 17 million observations of unique firm pairs, with 593 (923) thousand supplier-year (buyer-year) observations and 195 (289) thousand unique suppliers (buyers). The latter contains 42 (2) million worker-year (firm-year) observations and 6,497 (488) thousand unique workers (firms).

Starting from the baseline employer-employee dataset, we then define two subsamples that we will use in different parts of the paper. The first, which we refer to as the *stayers* sample, restricts the baseline sample to workers observed with the same employer for at least 8 consecutive years and to employers that have at least 10 stayers in each year. These restrictions allow for a flexible specification of how worker's earnings evolve over time at a given firm and ensures a sufficient sample size to perform the analyses at the firm level. We also omit the first and last years of workers' employment spells to avoid concerns over exit and entry into employment during the year, which confound our measure of annual earnings.

The second, which we refer to as the *movers* sample, restricts the baseline sample to workers

available.

observed at multiple firms over time. In other words, the firm that pays a worker her greatest earnings in a given year is not the same firm in all years. Following previous work and motivated by concerns about limited mobility bias, we also restrict the movers sample to firms with at least two movers (Lamadon et al., 2022). Finally, as in the previous literature (Abowd et al., 1999; Lamadon et al., 2022), we restrict this sample to firms that belong to the largest connected set of firms, which in our dataset represents 99.9% of workers.

Finally, for the purpose of estimating the elasticity of substitution between labor and materials, we merge the baseline employer-employee and the baseline firm-to-firm dataset using the unique tax IDs discussed above. We implement this merge at the firm-year level and thus exclude in the merged dataset the set of firms that do not have information in either the employer-employee or the firm-to-firm dataset. The sample includes 126 thousand firm-year observations and 48 thousand unique firms. We refer to this merged dataset as the *baseline firm-level dataset*.

Appendix Table A.1 compares the size of the three employer-employee datasets, the firm-to-firm dataset and the firm dataset we use throughout the paper. Detailed summary statistics of these samples are provided in Appendix Table A.2. The samples are broadly similar. The most noticeable differences are that the stayers sample has older, higher-earning workers and higher labor shares, as well as larger firms in terms of employment and degree (number of suppliers and buyers). Nonetheless, the firms in the stayers sample are broadly similar to the firms in the baseline employer-employee dataset in terms of value-added per worker, materials share of sales, and intermediate sales as a share of total sales.

6.2 Assumptions for identification

We now impose additional assumptions that will be helpful for identification of the model's primitives. These assumptions relate mainly to functional forms and the underlying stochastic processes for time-varying primitives. As we move toward connecting the model with worker-level data, we now also explicitly index individual workers by m.

Assumption **6.1.** *The ability of worker m at time t, a_{mt}, is comprised of a permanent component $\bar{a}_m$ and a time-varying component $\hat{a}_{mt}$, where $\log \hat{a}_{mt}$ follows a stationary mean-zero stochastic process that is independent of $\bar{a}_m$.*

This distinction between permanent and transient worker ability will be important for a decomposition of worker earnings into firm and worker effects that we implement below.

Assumption **6.2.** *Worker-firm production complementarity takes the following form:*

$$\log \phi_i (a_{mt}) = \theta_i \log \bar{a}_m + \log \hat{a}_{mt} \tag{6.1}$$

and the firm amenity function depends only on permanent worker ability, $g_i (a_{mt}) = g_i (\bar{a}_m)$.

This functional form for $\phi_i(\cdot)$ allows the model to generate an earnings equation that is consistent with the reduced-form model in Bonhomme et al. (2019), which features time-invariant worker-firm interactions. More generally, this assumption imposes that the two sources of worker-firm interactions in the model – $\phi_i(\cdot)$ and $g_i(\cdot)$ – are time-invariant. In what follows, we refer to the primitive θ_i simply as the *production complementarity* of firm i.

Assumption **6.3.** *Relationship productivity residuals $\tilde{\psi}_{ijt}$ are iid across firm pairs and time.*

This assumption will be important for the decomposition of firm-to-firm transactions into buyer and seller effects that we implement below.

Assumption **6.4.** *Time-varying firm primitives $\{T_{it}, \omega_{it}, \psi_{it}\}$ follow stationary first-order Markov processes with innovations that are iid across both firms and time.*

This follows well-known papers in the literature on production function estimation such as Olley and Pakes (1996) and Doraszelski and Jaumandreu (2018). As described below, we adopt the approach in the latter paper to estimate parameters of the production function and hence consider this Markov structure.

Assumption **6.5.** *The stochastic processes for transient worker ability $\hat{a}_{mt}$, time-varying firm primitives $\{T_{it}, \omega_{it}, \psi_{it}\}$, and relationship productivity residuals $\tilde{\psi}_{ijt}$ are mutually independent.*

Independence of the stochastic processes for worker and firm characteristics ensures that residual worker earnings due to transient ability shocks are uncorrelated with the characteristics of the worker's firm and is the same as the orthogonality assumption imposed in Lamadon et al. (2022). Furthermore, independence of firm primitives and relationship productivity residuals does not imply that firms match at random, only that they do not match based on $\tilde{\psi}_{ijt}$.

6.3 Identification strategy and estimation results

6.3.1 Labor supply elasticity

Given the functional form for $\phi_i(\cdot)$ in Assumption 6.2, the earnings equation (3.16) first allows us to express the log wage of worker m at firm i and time t as:

$$\log w_{imt} = \theta_i \log \bar{a}_m + \log \eta W_{it} + \log \hat{a}_{mt} \tag{6.2}$$

Note that this is consistent with the decomposition of log earnings in equation (2.1) presented in the motivation, with $x_m \equiv \log \bar{a}_m$ and $f_{it} \equiv \eta W_{it}$. Using equations (3.19) and (3.21) to substitute for W_{it}, we then obtain:

$$\log w_{imt} = \theta_i \log \bar{a}_m - \frac{1}{1+\gamma} \log \bar{\phi}_{it} + \frac{1}{1+\gamma} \log E^L_{it} + \log \hat{a}_{mt} \tag{6.3}$$

Now, note that the sorting composite $\bar{\phi}_{it}$ is time-varying only through the labor market indices $I_t(\cdot)$. Since Assumptions 6.1 and 6.4 impose stationarity on the distributions of time-varying worker and firm primitives respectively, we treat the aggregate indices $I_t(\cdot)$ and hence $\bar{\phi}_{it}$ as time-invariant. Restricting attention to workers that do not change employers between t and $t+1$ (stayers) and taking first-differences of equation (6.3) then gives:

$$\Delta \log w_{imt} = \frac{1}{1+\gamma} \Delta \log E_{it}^{L} + \Delta \log \hat{a}_{mt} \tag{6.4}$$

Intuitively, the change in a firm's wage bill is a sufficient statistic for all firm-level shocks that matter for changes in the earnings of stayers at the firm, including shocks to a firm's customers and suppliers in the production network. Since the labor supply elasticity γ controls the extent of imperfect competition in the labor market and mediates the extent of rent-sharing between a firm and its employees, the passthrough of changes in wage bills to changes in wages is informative about the magnitude of γ. In particular, stronger passthrough implies greater labor market power and a smaller value of γ.

This approach to the identification of γ resembles the passthrough analysis in Guiso et al. (2005) and Lamadon et al. (2022), but with one key difference: both of these papers construct shocks at the firm-level using changes in value-added, whereas our model motivates using changes in wage bills instead.[29] This distinction is moot in two special cases of our model: if output markets are perfectly competitive ($\sigma \to \infty$), so that profits are zero and value-added is equivalent to the wage bill; or if intermediates are not used in production ($\lambda \to 1$), in which case the wage bill is a constant fraction of value-added for every firm. In the general case, however, wage bills are not proportional to value-added and identification of γ requires leveraging changes in the former instead of the latter.[30]

To estimate γ in practice, we first remove age and year effects from log wages (since these are outside of our model) by regressing the latter on a vector of year dummy variables and a cubic polynomial in worker age, then treating the residual as our measure of $\log w_{imt}$. We then estimate the passthrough elasticity in equation (6.4) using the stayers sample. As we show formally in Appendix F.1, an IV approach that instruments $\Delta \log E_{it}^{L}$ with its own lags of at least 3 and greater is robust to allowing for measurement error in observed log wage bills (of an MA(1) form), whereas OLS estimation of equation (6.4) is not. Hence, in our preferred

[29] There are also subtle differences in the assumptions placed on the stochastic processes for firm-level shocks. Guiso et al. (2005) assume that log value-added follows an AR(1) process with innovations comprised of a unit root process plus an MA(1) process. Lamadon et al. (2022) make the same assumptions as Guiso et al. (2005) but constrain the AR(1) coefficient to be zero. In contrast, we allow for non-linear first-order Markov processes in firm primitives that determine firm wage bills (Assumption 6.4) and MA(1) measurement errors in wage bills, but consider only stationary processes for firm and worker shocks.

[30] In appendix F.1, we document our estimates of the labor supply elasticity γ using value-added shocks instead of wage bill shocks and show that we obtain different results. Hence, the distinction is both theoretically and empirically relevant.

specification, we instrument the change in the log wage bill using a cubic polynomial in 3, 4 and 5 of its own lags (stopping at 5 lags due to sample size considerations). The results obtained from this specification are shown in Column 1 of Table 3. We estimate a passthrough elasticity of around 0.15, which implies a labor supply elasticity of $\gamma = 5.5$.

For comparison, we also report results obtained from other specifications. In Column 2, we instrument $\Delta \log E_{it}^L$ using only a cubic polynomial in its third lag. In this case, the passthrough elasticity increases to 0.18 ($\gamma = 4.6$), although the first-stage F-statistic for this specification is substantially smaller than the corresponding F-statistic in our preferred specification. Nonetheless, both the estimates reported in Columns 1 and 2 are in line with estimates of passthrough elasticities reported in the literature.[31] In Column 3, we show the OLS estimate that ignores potential measurement error in wage bills. We find that the passthrough elasticity is substantially larger at 0.27. This implies $\gamma = 2.7$, which is half of our preferred estimate.[32]

Table 3: Estimation of labor supply elasticity (γ)

	$\Delta \log w_{imt}$		
	(1)	(2)	(3)
$\Delta \log \tilde{E}_{it}^L$	0.155	0.177	0.268
	(0.006)	(0.007)	(0.001)
γ	5.5	4.6	2.7
Strategy	GMM	GMM	OLS
Instruments Accumulated Lags	5	3	
First Stage F-Stat	2325	1426	
Number of Observations	2,507,868	2,507,868	2,507,868

Notes: This table presents results from the passthrough regression based on equation (6.4). All GMM strategies use different instruments of cubic polynomials of lags of wage bill and is implemented in two stages with a robust weighting matrix used to compute standard errors. Column 1 (our preferred specification) uses changes of wage bill lagged for 3, 4 and 5 periods as instruments. Column 2 uses changes of wage bill lagged for 3 periods as instruments. Column 3 estimates the model with OLS, which ignores measurement error on the wage bill. Standard errors are shown in parentheses.

[31] For example, in a survey, Card et al. (2018) report values for this elasticity between 0.10 and 0.15. Lamadon et al. (2022) in particular estimate a passthrough elasticity of 0.15. Note that these estimates rely on different sources of variation: whereas we use changes in wage bills (as justified by our model), Card et al. (2018) review estimates using value added per worker while Lamadon et al. (2022) use changes in value added.

[32] For additional robustness, we also consider a difference-in-difference estimator for γ proposed by Lamadon et al. (2022), which considers firms with above-median values of $\Delta \log E_{it}^L$ as treated and others as untreated. We find estimates of γ that are similar to our preferred estimate using this approach, the details of which are relegated to Appendix F.1 for brevity.

6.3.2 Worker abilities and firm production complementarities

We identify worker abilities $\{\bar{a}_m, \hat{a}_{mt}\}$ and firm production complementarities θ_i using the decomposition of log worker earnings studied by Bonhomme et al. (2019). We first move all time variation in wage bills to the left-hand side of the earnings equation (6.3) by rewriting this as:

$$\log \tilde{w}_{imt} = \underbrace{\theta_i \log \bar{a}_m}_{\text{worker-firm interaction}} + \underbrace{\log \bar{W}_i}_{\text{firm FE}} + \underbrace{\log \hat{a}_{mt}}_{\text{residual}} \tag{6.5}$$

where $\bar{W}_i \equiv \left(\bar{E}_i^L/\bar{\phi}_i\right)^{\frac{1}{1+\gamma}}$ is a time-invariant firm effect, $\log \bar{E}_i^L$ is the time-average of firm i's log wage bill, and $\log \tilde{w}_{imt} \equiv \log w_{imt} - \frac{1}{1+\gamma}\left(\log E_{it}^L - \log \bar{E}_i^L\right)$ is log worker earnings residualized by the innovation in its employer's log wage bill. Given the orthogonality of $\hat{a}_{mt}$ to both $\bar{a}_m$ (Assumption 6.1) and employer primitives (Assumptions 6.2 and 6.5), we then obtain the key identifying restriction in Bonhomme et al. (2019):

$$\mathbb{E}\left[\frac{1}{\theta_j}\left(\log \tilde{w}_{jm,t+1} - \log \bar{W}_j\right) - \frac{1}{\theta_i}\left(\log \tilde{w}_{im,t} - \log \bar{W}_i\right) | m \in M_{t,t+1}^{i\to j}\right] = 0 \tag{6.6}$$

where the expectation is taken over the set of workers $M_{t,t+1}^{i\to j}$ that move from firm i at time t to firm j at time $t+1$. In principle, this restriction gives $\left|\Omega^F\right|^2$ moment conditions for identification of $2\left|\Omega^F\right|$ parameters (θ_i and $\bar{W}_i$ for every firm), where intuitively, changes in earnings accompanying changes in employers are informative about the firm-specific determinants of earnings.

In practice, we follow Bonhomme et al. (2019) and first assign each firm in our data to one of ten clusters via a K-means clustering algorithm that targets moments of the within-firm distribution of residualized earnings $\tilde{w}_{imt}$, with $k(i)$ denoting the *earnings cluster* of firm i.[33] Although not strictly necessary for identification, this reduces the dimension of the parameter set that needs to be estimated and ameliorates the well-known limited mobility bias issue. We then estimate $\left\{\bar{W}_{k(i)}, \theta_{k(i)}\right\}$ via limited information maximum likelihood using the movers sample and the moment condition (6.6).[34] Permanent worker ability is then recovered as $\log \bar{a}_m = \mathbb{E}\left[\frac{\log \tilde{w}_{imt} - \log \bar{W}_{k(i)}}{\theta_{k(i)}}\right]$, while transient worker ability is recovered as the residual in earnings given our estimates of all other determinants of earnings. This allows us to estimate the worker ability distribution $L(\cdot)$. Furthermore, the time-varying firm earnings effect W_{it} can be recovered as $\log W_{it} = -\log \eta + \log \bar{W}_{k(i)} + \frac{1}{1+\gamma}\left(\log E_{it}^L - \log \bar{E}_i^L\right)$, which is firm-specific even though the time-invariant firm effect $\bar{W}_{k(i)}$ is restricted to vary only by cluster. Note that this approach allow us to estimate the decomposition of worker earnings (2.1) presented in section 2.

Our estimates of $\log \bar{W}_k$ and θ_k obtained using this procedure are presented in Table 4, where

[33] Appendix F.2 provides more details including diagnostics of the clustering procedure and robustness of our results with respect to the number of clusters.

[34] We thank Bradley Setzler for providing the code for this step of the estimation procedure.

clusters are sorted according to the former variable. We observe a positive correlation between $\log \bar{W}_k$ and θ_k, indicating that firms with higher wage premia are also those where workers of higher ability are more productive.[35] In addition, the estimates that we obtain for θ_k are indicative of strong production complementarities. For example, they imply that workers in the top 2% of the permanent ability distribution are around 40% more productive when employed at firms in the highest $\bar{W}_k$ cluster than at firms in the lowest $\bar{W}_k$ cluster.

Table 4: Estimates of firm fixed effects and production complementarities

Cluster	1	2	3	4	5	6	7	8	9	10
$\log \bar{W}_k$	0	0.25	0.61	0.89	1.06	1.24	1.50	1.69	1.80	1.92
θ_k	1	1.13	1.42	1.66	1.77	1.91	2.19	2.37	2.44	2.26

Notes: This table presents estimates of firm fixed effects $\log \bar{W}_k$ and production complementarities θ_k in the earnings equation (6.5) by earnings cluster k using the movers sample. Clusters are sorted in ascending order of $\log \bar{W}_k$. Note that $\log \bar{W}_k$ and θ_k are normalized to zero and one respectively for firms in the first earnings cluster.

6.3.3 Relationship capabilities and productivity residuals

Rewriting the firm-to-firm sales equation (4.1), we have:

$$\log r_{ijt} = \log \tilde{\Delta}_{it} + \log \tilde{\Phi}_{jt} + \log \tilde{\psi}_{ijt} \tag{6.7}$$

where $\tilde{\Delta}_{it} \equiv \Delta_{it}\psi_{it}$ and $\tilde{\Phi}_{jt} \equiv \Phi_{jt}\psi_{jt}$. Since the assignment of buyers to sellers is independent of $\tilde{\psi}_{ijt}$ under Assumption 6.3, $\tilde{\Delta}_{it}$ is identified from purchases by firm i from all its suppliers controlling for total sales by these suppliers, $\tilde{\Phi}_{jt}$ is identified from sales by firm j to all its customers controlling for total expenditures by these customers, and $\tilde{\psi}_{ijt}$ is identified from the residual.[36]

In practice, we estimate the terms on the right-hand side of equation (6.7) by regressing log firm-to-firm transactions on buyer-year and seller-year fixed effects. Details of the implementation are discussed in Appendix F.4. Of the total variance in log transaction values across all relationships, we find that 11.8% is explained by $\log \tilde{\Delta}_{it}$, 33.6% by $\log \tilde{\Phi}_{it}$, -0.5% by the covariance of the first two terms, and the remaining 55.1% by $\log \tilde{\psi}_{ijt}$. Hence, both firm-specific and relationship-specific characteristics are important drivers of variation in firm-to-firm sales.

To separately identify buyer effects Δ_{it}, seller effects Φ_{it}, and relationship capabilities ψ_{it} from $\tilde{\Delta}_{it}$ and $\tilde{\Phi}_{it}$, first note that the share of a firm's total sales that come from the network

[35] This positive correlation is also documented in Lamadon et al. (2022) using US data.

[36] Since matching in intermediate input markets can occur many-to-many (each seller can have several buyers at once and each buyer can have several sellers), this identification strategy only requires cross-sectional moments. This is in contrast with identification of the worker and firm earnings effects in equation (6.5), which requires movements of workers across firms over time.

(i.e. excluding final sales) can be expressed as $s_{it}^{net} = \frac{\psi_{it}\sum_{j\in\Omega_{it}^{C}}\tilde{\Delta}_{jt}\tilde{\psi}_{jit}}{E_t+\psi_{it}\sum_{j\in\Omega_{it}^{C}}\tilde{\Delta}_{jt}\tilde{\psi}_{jit}}$. Solving for ψ_{it}, we obtain:

$$\psi_{it} = E_t\left(\frac{s_{it}^{net}}{1-s_{it}^{net}}\right)\frac{1}{\sum_{j\in\Omega_{it}^{C}}\tilde{\Delta}_{jt}\tilde{\psi}_{jit}} \tag{6.8}$$

which therefore allows identification of ψ_{it} up to a normalizing constant given observable network sales shares s_{it}^{net}.[37] Intuitively, a higher value of ψ_{it} increases sales only within the network but not to final consumers. Buyer and seller effects are then easily recovered from $\tilde{\Delta}_{it}$ and $\tilde{\Phi}_{it}$.

6.3.4 Product substitution elasticity

In Appendix F.5, we show that the product substitution elasticity σ is identified from the following moment condition:

$$\sigma = \frac{\mathbb{E}\left[R_{it}\right]}{\mathbb{E}\left[R_{it}-TC_{it}\right]} \tag{6.9}$$

where expectations are taken over all firms in the economy. Intuitively, σ controls the extent of output market power and hence determines the aggregate ratio of sales to profits that appears on the right-hand side of (6.9), where total production costs TC_{it} in the denominator are adjusted for markdowns on wages. Using the sample moment analog of the right-hand side of equation (6.9), we obtain an estimate of $\sigma = 3.1$ for the average year in our sample.

6.3.5 Labor-materials substitution elasticity and labor productivities

Given the first-order Markov structure of firm productivity primitives in Assumption 6.4, we can first express log labor productivity as $\log\omega_{it} = F^{\omega}\left(\log\omega_{i,t-1}\right)+\xi_{it}^{\omega}$, where F^{ω} is a Markov transition function and ξ_{it}^{ω} is an innovation. Combining equations (3.15), (3.17), (3.21), and (3.22), we then obtain:

$$\log\frac{E_{it}^{M}}{E_{it}^{L}} = \log\left[\frac{1}{\eta}\left(\frac{1-\lambda}{\lambda}\right)\right]+(\epsilon-1)\log\frac{W_{it}}{Z_{it}}+(1-\epsilon)\,F^{\omega}\left(\log\omega_{i,t-1}\right)+(1-\epsilon)\,\xi_{it}^{\omega} \tag{6.10}$$

which is the standard relationship between relative factor expenditures $\left(\frac{E_{it}^{M}}{E_{it}^{L}}\right)$ and relative factor prices $\left(\frac{W_{it}}{Z_{it}}\right)$ implied by cost minimization under CES technologies. Note, however, that W_{it} and Z_{it} are both firm-specific statistics reflecting the heterogeneous wages and material prices that a firm pays to its workers and suppliers respectively, which generally differ from simple averages of these input costs. Instead, W_{it} is identified from the decomposition of worker earnings discussed in section 6.3.2, while $Z_{it} = \left(\sum_{j\in\Omega_{i}^{S}}\Phi_{jt}\psi_{ijt}\right)^{\frac{1}{1-\sigma}}$ can be constructed from the seller effects

[37] The normalizing constant for ψ_{it} is irrelevant for the same reason that one can normalize either Hicks neutral productivity or one factor-biased productivity without loss of generality.

and relationship productivities obtained from the firm-to-firm sales decomposition discussed in section 6.3.3, given a value for the product substitution elasticity σ. This highlights the importance of merged employer-employee and firm-to-firm data for identifying these factor price aggregates within the firm.

Under Assumption 6.4, identification of the labor-materials substitution elasticity ϵ from equation (6.10) then follows the strategy in Doraszelski and Jaumandreu (2018). We implement this on the baseline firm-level dataset, using polynomials in one-period lagged factor prices and expenditures to instrument for $\log \frac{W_{it}}{Z_{it}}$, as well as a cubic polynomial control function in $\log \frac{E^M_{it-1}}{E^L_{it-1}}$ and $\log \frac{W_{it-1}}{Z_{it-1}}$ to control for $F^{\omega}(\log \omega_{i,t-1})$, with a detailed derivation of the approach relegated to Appendix F.6. Since there are many potential instruments available, we implement estimation using all possible combinations of the instruments and vary the order of the polynomials used. Among specifications that deliver a p-value of the Hansen J test above 0.1, we then choose the specification that yields the highest F-statistic.

Table 5 presents our results. Our preferred specification based on the criteria above is shown in Column 1. This specification uses quadratic polynomials in $\{E^M_{it-1}, E^L_{it-1}\}$ as instruments and delivers an estimate of $\epsilon = 1.5$, implying that labor and materials are substitutes ($\epsilon > 1$), a result that holds with statistical significance. For comparison, we also present results obtained from other specifications. In Column 2, we use estimates of W_{it} obtained from the wage model and estimation strategy in Abowd et al. (1999), which does not address the issue of limited mobility bias and rules out worker-firm interactions. Applying the instrument selection criteria above, we use a linear polynomial in $\{E^M_{it-1}, E^L_{it-1}, W_{it-1}, Z_{it-1}\}$ as instruments and find $\epsilon = 1.6$, which is not statistically different from our preferred estimate in Column 1. In Column 3, we follow the standard approach in the literature of using average firm wages instead of the model-consistent firm earnings effect W_{it}. Our instrument set in this case is comprised of quadratic polynomials in $\{W_{it-1}, Z_{it-1}\}$. We find $\epsilon = 1.05$, which is not statistically different from one.[38] In all cases, we estimate that $\sigma > \epsilon$ with statistical significance (recall our baseline estimate of $\sigma = 3.1$), which from Proposition 1 implies that reductions in material costs Z_{it} have *positive* effects on wages.

Given an estimate of ϵ, labor productivities are then easily recovered as residuals in the relationship between relative input expenditures and prices. Furthermore, the weight on labor in the production function λ is not separately identified from the average level of labor productivity

[38] Oberfield and Raval (2019) and Doraszelski and Jaumandreu (2018) estimate values of ϵ below one using US and Spanish data respectively. However, their measures of factor prices differ fundamentally from ours. Both papers use average wages in place of W_{it}, while Oberfield and Raval (2019) use an industry fixed effect in place of Z_{it} and Doraszelski and Jaumandreu (2018) use a weighted-average of intermediate input prices in place of Z_{it}. Thus, our estimates, which are based on constructed price indices, are not strictly comparable. Nevertheless, in Column 3, we move closer to the empirical specification in Doraszelski and Jaumandreu (2018) by using the average wage instead of our model-based labor price index. Our estimate of ϵ falls and becomes more similar to their estimates.

ω_{it} across firms, which is intuitive since both λ and ω_{it} control the productivity of labor relative to materials. Hence, in what follows we set λ to an arbitrary constant in the interval $(0, 1)$ without any loss of generality.

Table 5: Estimation of labor-materials substitution elasticity, ϵ

	$\log E^M/E^L$		
	(1)	(2)	(3)
$\log W/Z$	0.553	0.623	
	(0.058)	(0.094)	
$\log \bar{w}/Z$			0.052
			(0.043)
ϵ	1.55	1.62	1.05
Model for Wage Component	BLM	AKM	Average
Instruments	$\{E^M_{it-1}, E^L_{it-1}\}$	$\{E^M_{it-1}, E^L_{it-1}, W_{it-1}, Z_{it-1}\}$	$\{W_{it-1}, Z_{it-1}\}$
Instrument Polynomial	Quadratic	Linear	Quadratic
First Stage F-Stat	130	84	186
Hansen's J Test	0.121	0.379	0.003
Number of Observations	44,967	44,967	44,967

Notes: This table presents estimates of equation (6.10) using the baseline firm-level dataset. Column 1, our preferred specification, is based on the specification selection criteria described in section 6.3.5. Column 2 uses the AKM wage model to estimate the firm effect W_{it} while Column 3 uses the average firm wage instead of W_{it}. All specifications are estimated using two-stage GMM with a robust weighting matrix. Standard errors are shown in parentheses.

6.3.6 Amenities

We identify firm amenities from variation in employment shares that is unexplained by differences in observable wages. Just as we restrict production complementarities θ_i to vary only by a firm's earnings cluster $k(i)$, we impose a similar restriction on amenities to reduce the dimension of parameters that need to be estimated:

$$g_i(\bar{a}) = \tilde{g}_i \bar{g}_{k(i)}(\bar{a}) \tag{6.11}$$

where $\bar{g}_{k(i)}(\bar{a})$ allows for worker-firm variation in amenities but restricts this to be the same for firms within a cluster, while variation in amenities across firms within a cluster is accounted for by $\tilde{g}_i$. As shown in Appendix F.3, the cluster-ability component of amenities can be identified from:

$$\bar{g}_k(\bar{a}) = \frac{1}{(\bar{a}_m)^{\theta_k}} \left[\Lambda_{kt}(\bar{a})\right]^{\frac{1}{\gamma}} \tag{6.12}$$

where $\Lambda_{kt}(\bar{a})$ is the share of workers of permanent ability $\bar{a}$ that are employed by firms in earnings cluster k. Similarly, the firm-specific component of amenities can be identified from:

$$\tilde{g}_i = \frac{1}{W_{it}} \left(\bar{\Lambda}_{it}\right)^{\frac{1}{\gamma}} \tag{6.13}$$

where $\bar{\Lambda}_{it}$ is the share of employment of all worker types by firms in cluster $k(i)$ accounted for by firm i. Since a firm with a high value of amenities is able to attract a large share of workers at a lower wage, amenities are intuitively identified from employment shares after controlling for relevant determinants of earnings – $\bar{a}^{\theta_k}$ at the cluster-ability level and W_{it} at the firm-level.

For a given worker type, we find lower amenity values at larger firms, with this negative relationship being stronger for workers of higher permanent ability. Furthermore, our estimates of amenities and production complementarities jointly imply the sorting of high-ability workers to firms with high wage premia (large values of $\bar{W}_i$). Details of these findings are relegated to Appendix F.3 for brevity.

6.3.7 Firm TFP

Identification of firm TFPs requires at least as many moment conditions as there are firms. For these, we rely on the firm earnings effects, which one can write in general as:

$$W_{it} = F_i\left(T_t|\Theta_t^{-T}\right) \tag{6.14}$$

where Θ_t is the set of model primitives listed in Definition 1, $T_t \equiv \{T_{it}\}_{i\in\Omega^F}$, $\Theta_t^{-T} \equiv \Theta_t \setminus T_t$, and $\{F_i\}_{i\in\Omega^F}$ is a set of *known* functions that depend on the structural relationships of the model. Given identification of all other primitives Θ_t^{-T}, equation (6.14) therefore provides a set of moments for exact identification of firm TFPs.

We choose this approach because it ensures that the model replicates the firm effects on earnings that we estimate from the data, which in turn guarantees that the model matches observed earnings for a given worker conditional on also replicating the worker's observed choice of employer. This allows us to examine changes in labor market outcomes under various counterfactual scenarios with the confidence that the baseline model provides a good fit to observed data. Note that in the limit of our model without intermediates ($\lambda \to 1$), $\log W_{it}$ is linear in $\log T_{it}$ and hence identification is trivial. With intermediates, however, the functions F_i are generally implicit and involve complex non-linearities, which hence requires a numerical solution for the TFP vector (see Appendix F.7 for details).[39]

7 The Production Network and Inequality Outcomes

We now use the estimated model to investigate the importance of the production network for explaining the four patterns of earnings inequality highlighted at the start of the paper. In particular, we quantify how production network heterogeneity shapes the following outcomes:

[39] Due to these features of the F_i functions, establishing a unique solution for T_t given a vector of firm effects W_t is not trivial. Nonetheless, we have explored the potential for multiplicity by varying the initial guess for the TFP vector and never find multiplicity to occur in practice.

(i) the variance of firm earnings effects and worker earnings; (ii) the covariance between firm earnings effects and firm size; (iii) the variance of labor shares of value-added; and (iv) the covariance between labor shares of value-added and firm size. For brevity, we henceforth refer to these variances and covariances as *inequality outcomes*.

7.1 Numerical solution approach

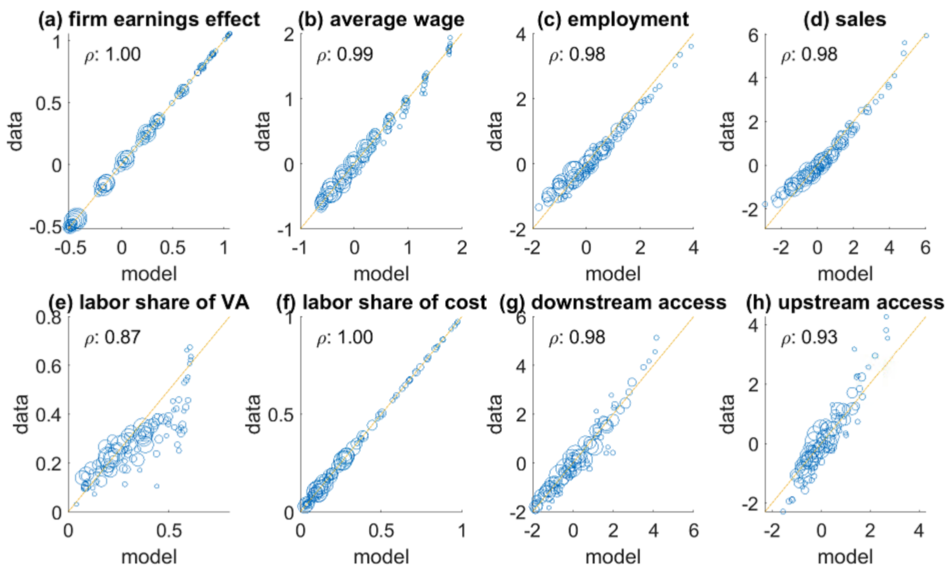
We begin by solving for a baseline equilibrium in which all model primitives are set to their estimated values (with time-varying primitives averaged across 2005-2010). Note that in our model, key outcomes such as the firm effect on earnings do not admit closed-form solutions in terms of model primitives due to non-linearities arising from the input-output structure of the production network. Hence, we require a numerical solution procedure to solve for model equilibria. Since this is not computationally feasible at the level of individual workers and firms (we have over 6 million workers and over 48 thousand firms), we proceed as follows.

First, we discretize the permanent and transient worker ability distributions into 50 quantiles each, which gives us 2,500 worker types. We then set primitives for each worker type (i.e., abilities and amenities) equal to the corresponding average across workers of each type. Second, within each of the ten firm earnings clusters (see section 6.3.2), we again cluster firms into ten subclusters via a K-means clustering algorithm targeting primitives $\{\omega_{it}, \psi_{it}, \tilde{g}_i\}$ that have been estimated at the firm-level. This gives us 100 firm cluster-subcluster pairs that we henceforth simply refer to as firm *groups*. We then set primitives for each firm group equal to the corresponding average across firms within each group, except for TFP, which we solve for numerically at the firm group level (see section 6.3.7). Finally, for the production network, we measure the fraction of potential buyer-seller firm pairs that are active between each group of buyers b and each group of sellers s in the average year, denoting this by m_{bs}. We then assume that each buyer in b matches with a random fraction m_{bs} of suppliers in s. We also set relationship productivity residuals for each buyer-seller group pair to the corresponding average across active relationships between each pair.

Using this approach, we solve for the model's equilibrium at the worker type and firm group level using a numerical solution algorithm described in Appendix H. Figure 4 shows the fit of the model's baseline equilibrium to key moments in the data, where each circle in a plot represents a firm group and the size of each circle is increasing in the number of firms in the group. The figure also shows the correlation (ρ) between each variable in the data and model at the firm group level, weighted by the number of firms in each group. Evidently, the model provides a good fit to all the key moments even after discretizing worker and firm primitives.[40] By implication, the model is able to closely replicate all of the inequality outcomes observed in the data.

[40] We observe a greater discrepancy between the model and data for labor shares of value-added, largely due to the fact that we do not consider firm heterogeneity in output markups. Even in this case, however, the model fit is good ($\rho = 0.87$).

Figure 4: Fit of the baseline equilibrium to key empirical moments



Notes: Each marker in the figure represents a firm group, with the size of each marker increasing in the number of firms in the group. ρ indicates the correlation between model and data moments at the firm group level, weighted by the number of firms in each group.

7.2 Sources of variation and a Shapley approach for counterfactuals

All inequality outcomes in the model are driven by heterogeneity in the following worker and firm primitives: (i) the production network, Ω^S_{it}; (ii) relationship productivity residuals, $\tilde{\psi}_{ijt}$ (iii) firm productivities, $\{T_{it}, \omega_{it}, \psi_{it}\}$; (iv) production complementarities, θ_i; (v) amenities, $g_i(\cdot)$; (vi) and worker abilities, $\{\bar{a}_m, \hat{a}_{mt}\}$. We refer to each of these six sets of primitives as a *source of variation* in the model.

To quantify the contribution of each source of variation to a given inequality outcome, we then proceed as follows. First, we simulate counterfactual equilibria of the model in which each source of variation v is eliminated by setting its value for all workers or firms equal to the mean of v across the respective sample. We then compute inequality outcomes and compare these to the baseline equilibrium, taking the difference between these values as a measure of the contribution of the source of variation v to each inequality outcome. For the production network, we do this separately for heterogeneity across suppliers and customers. For example, to eliminate heterogeneity across suppliers, we replace the observed network at the buyer group (b) and seller group (s) level, m_{bs}, with a counterfactual network $\hat{m}^S_{bs} = \frac{\sum_{s'} m_{bs} N_{s'}}{\sum_{s'} N_{s'}}$ that is randomized across suppliers while holding constant the total supplier count for each firm group, where $N_{s'}$ denotes the number of firms in seller group s'. We follow an analogous procedure to eliminate heterogeneity in customer matching.

However, note that eliminating a source of variation v not only removes variation in equilibrium outcomes arising from v but also from the covariance between v and all other sources of variation. Therefore, any changes in inequality outcomes that arise from eliminating v cannot be attributed to v alone. To address this, we adopt a Shapley-based approach: we simulate counterfactuals by eliminating *all* possible combinations of the sources of variation listed above and then compute the Shapley value for each source of variation in terms of its effect on inequality outcomes. Intuitively, this provides an *average* measure of the change in each inequality outcome when a source of variation is eliminated, under all possible combinations of the remaining sources of variation. This approach has two advantages. First, it accounts for interdependencies between sources of variation. Second, it has the appealing feature that each inequality outcome is apportioned *exactly* to each source of variation.

To illustrate, consider two sources of variation, Θ_A and Θ_B, and suppose that an inequality outcome such as the variance of earnings can be expressed in the baseline as $\mathrm{var}\,(\Theta_A)+\mathrm{var}\,(\Theta_B)+2\mathrm{cov}\,(\Theta_A,\Theta_B)$. The change in earnings variance from eliminating Θ_A relative to the baseline is $\delta_{A1}=\mathrm{var}\,(\Theta_A)+2\mathrm{cov}\,(\Theta_A,\Theta_B)$. The change in earnings variance from eliminating Θ_A relative to the equilibrium in which Θ_B has already been eliminated is $\delta_{A2}=\mathrm{var}\,(\Theta_A)$. The Shapley contribution of Θ_A to earnings variance is then $\frac{\delta_{A1}+\delta_{A2}}{2}=\mathrm{var}\,(\Theta_A)+\mathrm{cov}\,(\Theta_A,\Theta_B)$. The Shapley approach is therefore equivalent to splitting the covariance equally between Θ_A and Θ_B in this linear case, but is also applicable in the more general case where inequality outcomes cannot be expressed as a linear combination of the variances and covariances of model primitives. See Appendix I for a formal definition of the Shapley value.

7.3 Results

Table 6 presents our findings. Each panel shows results for a different inequality outcome, with the values in each panel reporting the shares of the inequality outcome accounted for by the indicated sources of variation.

First, consider the role of production network heterogeneity in explaining differences in earnings. In panel (a), around one-third (30.2%) of the variation in log earnings effects across firms (weighted by employment) is accounted for by network heterogeneity, with supplier heterogeneity (23.6%) playing a more important role than customer heterogeneity (6.6%). Therefore, network heterogeneity is a key driver of differences in employer-specific earnings premia. In panel (b), we see that production network heterogeneity also explains 12.9% of the variance of log earnings across workers, with supplier heterogeneity (9.7%) again playing a larger role than customer heterogeneity (3.2%). In comparison, own-firm primitives (productivities, production complementarities, and amenities) explain 19.3% of log earnings variance, while worker abilities explain the remainder (67.8%). Hence, network heterogeneity accounts for around two-fifths ($\frac{12.9}{12.9+19.3}$) of the variance of log worker earnings that is unexplained by worker characteristics.

Table 6: Decomposition of inequality outcomes

(a) variance, log firm earnings effect (baseline = 0.18)

supplier network:	23.6%	customer network:	6.6%	firm productivities:	40.7%
prod. complementarities:	26.7%	firm amenities:	13.3%	worker abilities:	-10.8%

(b) variance, log worker earnings (baseline = 0.64)

supplier network:	9.7%	customer network:	3.2%	firm productivities:	16.6%
prod. complementarities:	1.7%	firm amenities:	1.1%	worker abilities:	67.8%

(c) covariance, log firm earnings effect and log sales (baseline = 0.57)

supplier network:	35.3%	customer network:	15.8%	firm productivities:	44.8%
prod. complementarities:	8.0%	firm amenities:	1.7%	worker abilities:	-5.5%

(d) variance, labor share of value-added (baseline = 0.02)

supplier network:	21.9%	customer network:	4.2%	firm productivities:	76.2%
prod. complementarities:	-1.7%	firm amenities:	-1.1%	worker abilities:	0.5%

(e) covariance, labor share of value-added and log sales (baseline = -0.09)

supplier network:	-3.8%	customer network:	71.9%	firm productivities:	30.9%
prod. complementarities:	4.5%	firm amenities:	0.2%	worker abilities:	-3.8%

Notes: Each panel shows results for a different inequality outcome, with the value of the outcome in the baseline equilibrium reported in the panel headers. The variance in (a) and covariance in (c) are weighted by firm employment, while the variance in (d) and covariance in (e) are weighted by firm value-added. In each panel, the values reported are the shares of each inequality outcome accounted for by the corresponding source of variation.

Second, consider the role of the production network in explaining the positive firm size wage premium. In panel (c), 51.1% of the positive covariance between log earnings effects and log sales across firms (weighted by employment) is explained by network heterogeneity, 35.3% from supplier heterogeneity and 15.8% from customer heterogeneity. In other words, about half of the firm size wage premium is attributable to differences in production network linkages. Intuitively, this occurs because better access to suppliers and customers in the production network leads to increases in both firm size and firm earnings effects. Hence, network heterogeneity amplifies the positive relationship between these two outcomes.

Third, consider the role of the production network in explaining differences in labor shares across firms. In panel (d), 26.1% of the variation in labor shares of value-added across firms (weighted by value-added) is explained by network heterogeneity. As discussed in section 4.3, network heterogeneity affects the labor share of value-added through the relative cost of labor to materials, W_{it}/Z_{it}. Higher upstream access directly lowers Z_{it} while also having a positive effect on W_{it}. On the other hand, higher downstream access only affects this statistic through a positive effect on W_{it}. Hence, the role of production network in driving differences in labor shares of value-added across firms accrues mostly from supplier heterogeneity (21.9%) rather than customer heterogeneity (4.2%), as one might expect.

Finally, consider the role of the production network in explaining why larger firms tend to have lower labor shares of value-added. In panel (e), we see that 68.1% of the negative covariance between labor shares of value-added and log sales across firms (weighted by value-added) is explained by network heterogeneity. As shown in Fact 2 of the motivation, larger firms tend to have lower labor shares of value-added and better upstream and downstream network access in the baseline equilibrium. Furthermore, as discussed above, improvements in upstream and downstream access both tend to increase labor shares. Hence, the advantage that large firms tend to have in terms of network access amplifies the negative relationship between firm size and labor shares of value-added.

In sum, we find that production network heterogeneity plays a key role in explaining all four inequality outcomes.

7.4 Sensitivity of results to value-added production functions

How important is relaxing the assumption of value-added production functions for our main quantitative findings? To assess this, we proceed as follows.

First, we set the labor-materials substitution elasticity to $\epsilon = 1$ instead of our preferred estimate of $\epsilon = 1.5$. This imposes a Cobb-Douglas technology on the aggregation of labor and materials, which, as discussed in section 4.2.1, is necessary for a value-added representation of the production function to be valid. Second, since labor productivity ω_{it} is not separately identified from TFP when $\epsilon = 1$, we set $\omega_{it} = 1$ for all firms without loss of generality. Third,

Table 7: Decomposition of inequality outcomes under Cobb-Douglas technology

(a) variance, log firm earnings effect (baseline = 0.18)					
supplier network:	26.7%	customer network:	24.4%	firm productivities:	14.7%
prod. complementarities:	32.4%	firm amenities:	16.2%	worker abilities:	-14.4%
(b) variance, log worker earnings (baseline = 0.64)					
supplier network:	10.7%	customer network:	10.4%	firm productivities:	6.9%
prod. complementarities:	2.4%	firm amenities:	1.9%	worker abilities:	67.8%
(c) covariance, log firm earnings effect and log sales (baseline = 0.63)					
supplier network:	39.3%	customer network:	35.9%	firm productivities:	17.1%
prod. complementarities:	13.2%	firm amenities:	2.3%	worker abilities:	-7.7%

Notes: This table shows the same results as in panels (a)-(c) of Table 6, but with the elasticity of substitution between labor and materials set to $\epsilon = 1$ instead of our estimated value of $\epsilon = 1.5$ and with all other model primitives re-estimated as described in section 7.4.

note that differences in ω_{it} are what allow our baseline model to fit labor shares of cost at the firm-level, while the weight on labor in the production function λ is not separately identified from the mean of ω_{it} across firms. With $\epsilon = 1$, labor shares of cost are instead determined by λ, which we hence choose to match the aggregate labor share of cost. Fourth, since our TFP estimates depend on the value of ϵ, we re-estimate TFP under $\epsilon = 1$.

Finally, we recompute decompositions of inequality outcomes shown in Table 6. It is immediately obvious that the model with $\epsilon = 1$ cannot speak to labor share heterogeneity, since the Cobb-Douglas technology imposes common labor shares of cost and therefore common labor shares of value-added across firms. Hence, we focus here only on panels (a)-(c) of Table 6, which decompose the variance of log firm earnings effects, the variance of log worker earnings, and the covariance between log firm earnings effects and log sales. Table 7 shows the results of these decompositions with $\epsilon = 1$. We highlight two key insights.

First, the share of each inequality outcome accounted for by production network heterogeneity is substantially different under $\epsilon = 1$ compared with our baseline case of $\epsilon = 1.5$. Specifically, under Cobb-Douglas technology, we find that production network heterogeneity explains 51.2% of the variance in log firm earnings effects (baseline: 30.2%), 21.1% of the variance in log worker earnings (baseline: 12.9%), and 75.2% of the covariance between log firm earnings effects and log sales (baseline: 51.1%). In other words, in terms of our outcomes of interest, our estimated production function with $\epsilon = 1.5$ is by no means "close" to being approximated by a value-added production function.

Second, it is evident that the importance of production network heterogeneity for the three inequality outcomes above is *greater* under Cobb-Douglas technology than under our estimate of $\epsilon = 1.5$. The core intuition for this is the following. As described in section 2, firms with higher earnings effects tend to have greater upstream network access (Fact 1). Consequently, eliminating heterogeneity in the production network tends to *reduce* upstream access for firms with higher earnings effects and tends to *increase* upstream access for firms with lower earnings effects. Then, as discussed in section 4.2.2, changes in upstream access have two effects on wages. First, a reduction in upstream access lowers the scale of a firm, which tends to reduce wages as firms choose to hire fewer workers. This scale effect hence tends to reduce (increase) wages at firms with high (low) earnings effects in the baseline, thereby reducing earnings inequality. However, when $\epsilon > 1$, there is an additional substitution effect, as a reduction in upstream access induces a firm to substitute away from materials and towards labor, thereby increasing wages. Consequently, when $\epsilon > 1$, the substitution effect partially offsets the scale effect, so that the importance of heterogeneity in upstream access for earnings inequality is dampened. In contrast, the offsetting substitution effect is not operative under Cobb-Douglas technology and hence heterogeneity in upstream access matters more for earnings inequality in this case.[41]

8 Conclusion

We have developed in this paper a unifying framework with firm heterogeneity in both earnings premia and labor shares of value-added. Central to our framework are firm labor market power and heterogeneous firm-to-firm production network linkages with CES production technologies. These features allow the model to reconcile stylized facts about earnings inequality and labor shares in the cross-section, as well as empirical correlations between measures of production network access and labor market outcomes. The key mechanisms linking production network access to earning premia and labor shares in our model are well-supported by reduced-form evidence of the passthrough of demand and material cost shocks into changes in worker earnings.

Using linked employer-employee and firm-to-firm transactions data from Chile, we structurally estimate the model and show how these data can be used to identify the elasticity of substitution between labor and materials when these inputs are heterogeneous within firms. We estimate that labor and materials are gross substitutes and reject the hypothesis of value-added production functions. Counterfactual simulations of our estimated model indicate that production network heterogeneity is an important driver of key labor market outcomes, in particular the variances of worker earnings, firm earnings premia, and labor shares of value-added, as well

[41] The value of ϵ also matters for how heterogeneity in downstream network access affects earnings inequality, although this channel is less important quantitatively than the channel described above. One can see this from the fact that changing the value of ϵ mainly affects the shares for heterogeneity in the *customer* network as opposed to heterogeneity in the supplier network in panels (a)-(c) of Tables 6 and 7. Eliminating heterogeneity across customers holds constant the number of customers for each firm but assumes that each seller is equally likely to sell to any given buyer, which hence leads to identical upstream network access measures for all firms.

the covariances between firm earnings premia, labor shares of value-added, and firm size.

We conclude with two potential directions for future research on the interaction between workers and production networks. First, there is growing evidence that worker outsourcing is a key driver of increases in earnings inequality (Goldschmidt and Schmieder (2017)). However, in these settings, it is typically not possible to directly observe and hence measure outsourcing. The growing availability of linked employer-employee and firm-to-firm datasets provides a unique opportunity to measure flows of both goods and workers between firms. This will allow researchers to more accurately measure outsourcing at the firm and to understand its incidence on both workers and firms. However, there are as yet no studies of outsourcing that have leveraged these data.[42]

Second, there is growing interest among both policymakers and researchers in understanding the effects of automation on worker outcomes. It is natural to view these effects as arising from the substitution of labor by inputs such as industrial robots. For example, Acemoglu and Restrepo (2020) estimate the effects of increased robot usage on employment and wages in US labor markets, finding robust negative effects. More recent theoretical work by Jackson and Kanik (2020) develops a model of robot-labor substitution that accounts for production network linkages between firms. A quantitative study of the mechanisms highlighted by this literature using matched employer-employee and firm-to-firm transactions data is therefore likely to yield important insights.

[42] Cardoza et al. (2022) provide evidence that workers are more likely to move to a customer or supplier of their original employer than to an unrelated firm, but do not consider outsourcing specifically.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. Econometrica 67(2), 251–333.

Acemoglu, D. and P. Restrepo (2020). Robots and jobs: Evidence from us labor markets. Journal of Political Economy 128(6), 2188–2244.

Adao, R., P. Carrillo, A. Costinot, D. Donaldson, and D. Pomeranz (2020). International trade and earnings inequality: A new factor content approach. Working paper.

Alfaro-Ureña, A., I. Manelici, and J. P. Vasquez (2020). The Effects of Multinationals on Workers: Evidence from Costa Rica. Working paper.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. V. Reenen (2020). The fall of the labor share and the rise of superstar firms. Quarterly Journal of Economics 135(2), 645–709.

Azar, J. A., S. T. Berry, and I. Marinescu (2022). Estimating labor market power. Working paper.

Barth, E., A. Brysona, J. C. Davis, and R. Freeman (2016). It's where you work: Increases in earnings dispersion across establishments and individuals in the us. Journal of Labor Economics 34(S2), S67 – S97.

Berger, D., K. Herkenhoff, and S. Mongey (2019). Labor market power. Working paper.

Bernard, A., E. Dhyne, G. Magerman, A. Moxnes, and K. Manova (2022). The origins of firm heterogeneity: A production network approach. Technical Report 7.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2020). How Much Should we Trust Estimates of Firm Effects and Worker Sorting? Working paper.

Bonhomme, S., T. Lamadon, and E. Manresa (2019). A distributional framework for matched employer employee data. Econometrica 87(3), 699–739.

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

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), S13–S70.

Cardoza, M., F. Grigoli, N. Pierri, and C. Ruane (2022). Worker mobility and domestic production networks. Working paper.

Chan, M., K. Kroft, E. Mattana, and I. Mourifié (2022). An empirical framework for matching with imperfect competition. Working paper.

Chan, M., S. Salgado, and M. Xu (2021). Heterogeneous passthrough from tfp to wages. Working paper.

Demir, B., A. C. Fieler, D. Y. Xu, and K. K. Yang (2018). O-Ring Production Network. Working paper.

Dhyne, E., A. K. Kikkawa, M. Mogstad, and F. Tintelnot (2021). Trade and domestic production networks. Review of Economic Studies 88(2), 643–668.

Dhyne, E., K. Kikkawa, T. Komatsu, M. Mogstad, and F. Tintelnot (2022). How do firms grow in response to demand shocks? a network perspective. Working paper.

Doraszelski, U. and J. Jaumandreu (2018). Measuring the bias of technological change. Journal of Political Economy 126(3), 1027–1084.

Dunne, T., L. Foster, J. Haltiwanger, and K. R. Troske (2004). Wage and productivity dispersion in united states manufacturing: The role of computer investment. Journal of Labor Economics 22(2), 397–429.

Faggio, G., K. Salvanes, and J. V. Reenen (2010). The evolution of inequality in productivity and wages: Panel data evidence. Industrial and Corporate Change 19(6), 1919–1951.

Goldschmidt, D. and J. F. Schmieder (2017). The rise of domestic outsourcing and the evolution of the german wage structure. Quarterly Journal of Economics 132(3), 1165–1217.

Guiso, L., L. Pistaferri, and F. Schivardi (2005). Insurance within the firm. Journal of Political Economy 113(5), 1054–1087.

Huneeus, F. (2019). Production Network Dynamics and the Propagation of Shocks. Working paper.

Jackson, M. O. and Z. Kanik (2020). How automation that substitutes for labor affects production networks, growth, and income inequality. Working paper.

Jarosch, G., J. Nimczik, and I. Sorkin (2019). Granular search, market structure, and wages. Working paper.

Karabarbounis, L. and B. Neiman (2014). The global decline of the labor share. Quarterly Journal of Economics 129(1), 61–103.

Kehrig, M. and N. Vincent (2021). The micro-level anatomy of the labor share decline. Quarterly Journal of Economics 136(2), 1031–1087.

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), 1343–1404.

Kroft, K., Y. Luo, M. Mogstad, and B. Setzler (2022). Imperfect competition and rents in labor and product markets: The case of the construction industry. Working paper.

Lamadon, T., M. Mogstad, and B. Setzler (2022). Imperfect Competition, Compensating Differentials, and Rent Sharing in the U.S. Labor Market. American Economic Review 112(1), 169–212.

Lim, K. (2019). Production Networks and the Business Cycle. Working paper.

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

Oberfield, E. and D. Raval (2019). Micro Data and Macro Technology. Working paper.

Oi, W. and T. Idson (1999). Firm size and wages. In O. Ashenfelter and D. Card (Eds.), Handbook of Labor Economics, Vol. 3, pp. 2165–2214. Amsterdam: Elsevier.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. Econometrica 64(6), 1263–1297.

Postel-Vinay, F. and J.-M. Robin (2002). Equilibrium wage dispersion with worker and employer heterogeneity. Econometrica 70(6), 2295–2350.

Rosen, S. (1986). The theory of equalizing differences. Handbook of Labor Economics 1, 641–692.

Song, J., D. J. Price, N. Bloom, and T. von Watcher (2019). Firming up inequality. Quarterly Journal of Economics 134(1), 1–50.

Sorkin, I. (2018). Ranking firms using revealed preference. Quarterly Journal of Economics 133(3), 1331–1393.

Taber, C. and R. Vejlin (2018). Estimation of a roy/search/compensating differentials model of the labor market. Working paper.

Van Reenen, J. (1996). The creation and capture of rents: Wages and innovation in a panel of u.k. companies. Quarterly Journal of Economics 111(1), 195–226.

A Decompositions of the variance of worker earnings

Following Lamadon et al. (2022), we first utilize equation (2.1) to decompose the variance of log earnings as:

$$\text{var}\left(\log w_{imt}\right) = \underbrace{\text{var}\left(\tilde{x}_m\right)}_{57\%} + \underbrace{\text{var}\left(\log \tilde{f}_{it}\right)}_{10.8\%} + \underbrace{2\text{cov}\left(\tilde{x}_m, \log \tilde{f}_{it}\right)}_{19.8\%} + \underbrace{int}_{-2.0\%} + \underbrace{\text{var}\left(\hat{x}_{mt}\right)}_{14.4\%} \tag{A.1}$$

where $\tilde{x}_m \equiv (x_m - \bar{x})\,\bar{\theta}$ is the worker effect when employed at the average firm, $\log \tilde{f}_{it} \equiv \log f_{it} + \theta_i \bar{x}$ is the firm effect when matched with the average worker, $\left\{\bar{x}, \bar{\theta}\right\}$ denote the averages of $\{x_m, \theta_i\}$ across workers, *int* collects terms arising from non-linear interactions between the worker and firm effects, and all variances and covariances are computed at the worker-level. Unsurprisingly, we find that the variance of the (transformed) worker effect accounts for the majority of earnings variance. However, we also find that firms play an important role: the variance of the (transformed) firm effect accounts for 10.8% of log earnings variance, while the sorting covariance between worker and firm effects explains 19.8%.

Following Song et al. (2019), we also decompose the variance in log earnings across workers into a between-firm and a within-firm component:

$$\text{var}\left(\log w_{mit}\right) = \underbrace{\text{var}_i\left(\overline{\log w}_{it}\right)}_{\text{between-firm: } 46\%} + \underbrace{\Sigma_i \varrho_{it} \text{var}_{m \in M_{it}}\left(\log w_{mit}\right)}_{\text{within-firm: } 54\%} \tag{A.2}$$

where for each year t, $\overline{\log w}_{it}$ denotes the average log wage at firm i and ϱ_{it} is the share of workers employed at firm i. The operator var_i denotes the variance across firms weighted by employment and $\text{var}_{m\in M_{it}}$ denotes the variance across the set of workers M_{it} employed at firm i. Computing this decomposition for each year in our sample, we find that the between-firm component of log earnings variance explains 46% of total log earnings variance in the average year.

B Model Extension with Capital Inputs

Suppose that firms produce output using capital in addition to labor and materials with a production function of the following form:

$$X_{it} = T_{it} K_{it}^{\alpha} F\left[\left\{\phi_{it}(a)\, L_{it}(a), M_{it}(a)\right\}_{a\in A}\right]^{1-\alpha} \tag{B.1}$$

where α is the capital share of cost. Suppose also that capital is available at a price r_{it} that may vary across firms due to differences in access to capital markets. The firm's profit maximization problem can now be written as:

$$\max_{K_{it}, \{w_{it}(a), M_{it}(a)\}_{a\in A}} \left\{ D_{it}^{\frac{1}{\sigma}} X_{it}^{\frac{\sigma-1}{\sigma}} - \sum_{a\in A} w_{it}(a)\, L_{it}(a) - Z_{it} \sum_{a\in A} M_{it}(a) - r_{it} K_{it} \right\} \tag{B.2}$$

subject to the production function (B.1) and labor supply curves (3.2). The first-order condition for this problem with respect to the capital input is:

$$\alpha D_{it}^{\frac{1}{\sigma}} X_{it}^{\frac{\sigma-1}{\sigma}} = r_{it} K_{it} \tag{B.3}$$

Using this to substitute for the choice of capital, we can rewrite the profit maximization problem as a choice over wages and material inputs alone, as in the original problem (3.14):

$$\max_{\{w_{it}(a), M_{it}(a)\}_{a \in A}} \left\{ D_{it}^{\frac{1}{\tilde{\sigma}}} \tilde{X}_{it}^{\frac{\tilde{\sigma}-1}{\tilde{\sigma}}} - \sum_{a \in A} w_{it}(a) L_{it}(a) - Z_{it} \sum_{a \in A} M_{it}(a) \right\} \tag{B.4}$$

$$\text{s.t. } \tilde{X}_{it} = \tilde{T}_{it} F \left[\{ \phi_{it}(a) L_{it}(a), M_{it}(a) \}_{a \in A} \right] \tag{B.5}$$

where $\tilde{\sigma} \equiv \sigma(1-\alpha) + \alpha$ and $\tilde{T}_{it} \equiv (1-\alpha)^{\frac{\tilde{\sigma}}{\tilde{\sigma}-1}} \left(\frac{\alpha}{r_{it}} \right)^{\frac{\alpha}{1-\alpha}} T_{it}^{\frac{1}{1-\alpha}}$. Hence, the firm's problem with capital is isomorphic to the problem without capital if one replaces σ with $\tilde{\sigma}$ and T_{it} with $\tilde{T}_{it}$. Note that the introduction of capital lowers the effective price elasticity of demand (since $\tilde{\sigma} < \sigma$ for any $\alpha \in (0,1)$), while differences in capital prices r_{it} can be viewed as differences in effective productivity.

C Value-added per worker and wages

We discuss here the implications of the model for the passthrough of changes in value-added per worker to changes in worker earnings. For brevity, consider a simplified version of our model with no heterogeneity in worker ability. The production function (3.5) can then generally be represented as:

$$X_{it} = T_{it} \omega_{it} L_{it} F(1, \nu_{it}) \tag{C.1}$$

where we have utilized the property that F is homogeneous of degree one. Wages are given by:

$$w_{it} = \eta W_{it} \tag{C.2}$$

while value-added is equal to the difference between sales and material costs:

$$VA_{it} = p_{it} X_{it} - Z_{it} M_{it} \tag{C.3}$$

Since $\nu_{it} \equiv \frac{M_{it}}{\omega_{it} L_{it}}$, value-added per worker is then:

$$VAPW_{it} = p_{it} T_{it} \omega_{it} F(1, \nu_{it}) - Z_{it} \nu_{it} \omega_{it} \tag{C.4}$$

Eliminating Z_{it} using the firm's first-order condition for materials (3.15) and again utilizing the property of F being homogeneous of degree one, we can rewrite this as:

$$VAPW_{it} = \frac{\sigma - 1}{\sigma} p_{it} T_{it} \omega_{it} F_L(1, \nu_{it}) + \frac{1}{\sigma} p_{it} T_{it} \omega_{it} F(1, \nu_{it}) \tag{C.5}$$

Finally, substituting the firm's first-order condition with respect to labor (3.17), we obtain:

$$VAPW_{it} = W_{it}\left[1 + \left(\frac{1}{\sigma - 1}\right)\frac{F\left(1, \nu_{it}\right)}{F_L\left(1, \nu_{it}\right)}\right] \tag{C.6}$$

Hence, value-added per worker is not proportional to the firm effect W_{it} except in three special cases of the model: (i) no output market power ($\sigma \to \infty$); (ii) no materials in production ($\nu_{it} = 0$); and (iii) Cobb-Douglas technology (so that $\frac{F(1,\nu_{it})}{F_L(1,\nu_{it})}$ is independent of ν_{it}). An immediate corollary is that the assumption of Cobb-Douglas technology in our model implies complete passthrough of changes in value-added per worker to changes in worker earnings, which is strongly rejected by the empirical literature (as in Berger et al. (2019) and Kline et al. (2019), for example).

D Proofs of Claims and Propositions

D.1 Proof of Claim 1

Omitting time subscripts for brevity, the profit-maximization problem for a firm i can be written generally as:

$$\max_{\{p_{ji}\}_{j \in \Omega_i^C \cup \{F\}}} \left\{ \sum_{j \in \Omega_i^C \cup \{F\}_i} p_{ji} x_{ji} - C\left[X_i | l_i\left(\cdot\right), Z_i\right] \right\} \tag{D.1}$$

$$\text{s.t. } x_{ji} = \Delta_j \psi_{ji} p_{ji}^{-\sigma} \tag{D.2}$$

$$X_i = \sum_{j \in \Omega_i^C \cup \{F\}} x_{ji} \tag{D.3}$$

where $\psi_{Fi} = 1$. Here, $C\left[X_i | l_i\left(\cdot\right), Z_i\right]$ denotes the total cost of producing X_i units of output given the labor supply functions $l_i\left(\cdot\right)$ and material input cost Z_i. The latter depends on the prices charged by suppliers of firm i, which firm i takes as given in the problem above. Importantly, the total production cost for firm i depends only on total output of the firm X_i and not on how this output is allocated to each customer.

The first-order condition for the profit-maximization problem with respect to p_{ji} is then:

$$\left(1 - \sigma\right)\Delta_j \psi_{ji} p_{ji}^{-\sigma} = -\sigma C'\left[X_i | l_i\left(\cdot\right), Z_i\right]\Delta_j \psi_{ji} p_{ji}^{-\sigma - 1} \tag{D.4}$$

Solving for the optimal price yields:

$$p_{ji} = \frac{\sigma}{\sigma - 1} C'\left[X_i | l_i\left(\cdot\right), Z_i\right] \tag{D.5}$$

Note that the right-hand side of (D.5) does not vary by customer j. Hence, the optimal prices set by firm i do not vary by customer and are equal to the standard CES markup over the firm's marginal cost. The existence of imperfect competition in the labor market implies that marginal cost is not constant, but this does not break the standard CES markup result.

D.2 Proof of Proposition 1

First, we establish comparative statics of a firm's earnings effect W_{it} with respect to the firm's network characteristics $\{D_{it}, Z_{it}\}$, technological primitives $\{T_{it}, \omega_{it}\}$, and sorting composite $\bar{\phi}_{it}$. In what follows, we omit firm and time subscripts for brevity and all derivatives of the production function f are evaluated at $\{\phi L, M\} = \{1, \nu\}$. Totally differentiating (3.15), (3.17), and (3.18) for a given firm, we obtain:

$$\hat{W} + \frac{1}{\sigma}\hat{X} - \left(\frac{f_{LM}\nu}{f_L}\right)\hat{\nu} = \frac{1}{\sigma}\hat{D} + \hat{T} + \hat{\omega} \tag{D.6}$$

$$\frac{1}{\sigma}\hat{X} - \left(\frac{f_{MM}\nu}{f_M}\right)\hat{\nu} = \frac{1}{\sigma}\hat{D} + \hat{T} - \hat{Z} \tag{D.7}$$

$$-\gamma\hat{W} + \hat{X} - \left(\frac{f_M \upsilon}{f}\right)\hat{\nu} = \hat{T} + \hat{\omega} + \hat{\bar{\phi}} \tag{D.8}$$

Solving for $\left\{\hat{W}, \hat{X}, \hat{\nu}\right\}$, we obtain:

$$\begin{aligned}
\hat{W} &= \Gamma\hat{D} + (\sigma-1)\,\Gamma\hat{T} - (\sigma-\epsilon)\,\varepsilon^M\Gamma\hat{Z} + \left[\sigma - 1 - (\sigma-\epsilon)\,\varepsilon^M\right]\Gamma\hat{\omega} && \text{(D.9)}\\
&\quad - \Gamma\hat{\bar{\phi}}\\
\hat{X} &= \left(\gamma + \epsilon\varepsilon^M\right)\Gamma\hat{D} + \sigma\left(\gamma + \epsilon\varepsilon^M + 1 - \varepsilon^M\right)\Gamma\hat{T} - \sigma\,(\gamma+\epsilon)\,\varepsilon^M\Gamma\hat{Z} && \text{(D.10)}\\
&\quad + \sigma\left(1-\varepsilon^M\right)(1+\gamma)\,\Gamma\hat{\omega} + \sigma\left(1-\varepsilon^M\right)\Gamma\hat{\bar{\phi}}\\
\hat{\upsilon} &= \epsilon\Gamma\hat{D} + \epsilon\,(\sigma-1)\,\Gamma\hat{T} - \epsilon\,(\gamma+\sigma)\,\Gamma\hat{Z} - \epsilon\,(1+\gamma)\,\Gamma\hat{\omega} && \text{(D.11)}\\
&\quad - \epsilon\Gamma\hat{\bar{\phi}}
\end{aligned}$$

where $\varepsilon^M \equiv \frac{f_M\nu}{f}$ denotes the elasticity of f with respect to materials and $\Gamma \equiv \frac{1}{\gamma+\sigma(1-\varepsilon^M)+\epsilon\varepsilon^M}$ is the scale elasticity for the firm.

Now from equations (3.21) and (3.22), we can express the material share of cost (adjusted for markdowns on wage) for the firm as:

$$s^M \equiv \frac{E^M}{\frac{1}{\eta}E^L + E^M} = \frac{Z\omega\nu}{W + Z\omega\nu} \tag{D.12}$$

Then, from the first-order conditions (3.17) and (3.15), relative factor prices can be expressed as:

$$\frac{Z}{W/\omega} = \frac{f_M}{f_L} \tag{D.13}$$

Using the result that $f = f_M\nu + f_L$ for a homogeneous of degree one function f then implies:

$$\varepsilon^M = \frac{Z\omega\nu}{W + Z\omega\nu} \tag{D.14}$$

Finally, comparing equations (D.12) and (D.14) implies $\varepsilon^M = s^M$, so that the elasticity of f with respect to materials is equal to the material share of cost in equilibrium. From the coefficients on the right-hand side of equation (D.9), we observe the own-firm effects of demand

and material cost shocks as stated in Proposition 1. Furthermore, we see that a firm's earnings effect is increasing in its TFP, increasing in its labor productivity if and only if $s^M < \frac{\sigma-1}{\sigma-\epsilon}$, and decreasing in its sorting composite.

Next, we characterize the passthrough of shocks in the production network into firm earnings effects more generally, including the passthrough effects highlighted in Proposition 1. We begin by deriving an expression for marginal changes in demand shifters, $\hat{D}$. Totally differentiating equation (3.10) gives:

$$\hat{D} = S^{sales}\hat{\Delta} \tag{D.15}$$

where we have used the result that the share of firm i's sales accounted for by firm j can be expressed as:

$$s_{jit}^{sales} \equiv \frac{R_{jit}}{\sum_{k\in\Omega_{it}^{C}\cup\{F\}} R_{kit}} = \frac{\Delta_j}{D_i} \tag{D.16}$$

Totally differentiating (3.11) and using (3.22), we obtain:

$$\hat{\Delta} = \gamma\hat{W} + \sigma\hat{Z} + \hat{\nu} + \hat{\omega} \tag{D.17}$$

Then, taking the ratio of the first-order conditions for the profit-maximization problem (3.15) and (3.17) and totally differentiating gives:

$$\hat{W} - \hat{Z} = \epsilon^{-1}\hat{\nu} + \hat{\omega} \tag{D.18}$$

Combining (D.15), (D.17), and (D.18), we then obtain the following expression for marginal changes in demand shifters:

$$\hat{D} = S^{sales}\left[(\gamma+\epsilon)\hat{W} + (\sigma-\epsilon)\hat{Z} + (1-\epsilon)\hat{\omega}\right] \tag{D.19}$$

Next, we derive an expression for marginal changes in material costs, $\hat{Z}$. Totally differentiating equation (3.12) gives:

$$\hat{Z} = -\frac{1}{\sigma-1}S^{mat}\hat{\Phi} \tag{D.20}$$

where we have used the result that the share of firm i's input expenditures accounted for by firm j can be expressed as:

$$s_{ijt}^{mat} \equiv \frac{R_{ijt}}{\sum_{k\in\Omega_{it}^{S}} R_{ikt}} = \frac{\Phi_{jt}\psi_{ijt}}{Z_{it}^{1-\sigma}} \tag{D.21}$$

Then, from (3.9) and (3.13), we can express marginal changes in network productivities as:

$$\hat{\Phi} = \frac{\sigma-1}{\sigma}\left(\hat{X} - \hat{D}\right) \tag{D.22}$$

Hence, combining (D.20) and (D.22), we obtain the following expression for marginal changes in material costs:

$$\hat{Z} = \frac{1}{\sigma}S^{mat}\left(\hat{D} - \hat{X}\right) \tag{D.23}$$

Now equations (D.9)-(D.11), (D.19), and (D.23) define a linear system in $\left\{\hat{W}, \hat{X}, \hat{\nu}, \hat{D}, \hat{Z}\right\}$, given changes in TFP $\hat{T}$ and labor productivity $\hat{\omega}$. Eliminating $\hat{X}$ and $\hat{\nu}$ from this system, we can write the remaining equations as:

$$\hat{W} = H^{WT}\hat{T} + H^{W\omega}\hat{\omega} + H^{WD}\hat{D} + H^{WZ}\hat{Z} \tag{D.24}$$

$$\hat{D} = S^{sales}\left[H^{DT}\hat{T} + H^{D\omega}\hat{\omega} + H^{DD}\hat{D} + H^{DZ}\hat{Z}\right] \tag{D.25}$$

$$\hat{Z} = S^{mat}\left[H^{ZT}\hat{T} + H^{Z\omega}\hat{\omega} + H^{ZZ}\hat{Z} + H^{ZD}\hat{D}\right] \tag{D.26}$$

where the H matrices are all $\left|\Omega^F\right| \times \left|\Omega^F\right|$ diagonal matrices. The matrices summarizing the dependence of $\left\{\hat{W}, \hat{D}, \hat{Z}\right\}$ on productivity shocks $\left\{\hat{T}, \hat{\omega}\right\}$ have i^{th}-diagonal elements given by:

$$\begin{array}{ll} H_i^{WT} = (\sigma-1)\,\Gamma_i & H_i^{W\omega} = \left[(\sigma-1) - (\sigma-\epsilon)\,s_i^M\right]\Gamma_i \\ H_i^{DT} = (\gamma+\epsilon)\,(\sigma-1)\,\Gamma_i & H_i^{D\omega} = (1+\gamma)\,(\sigma-\epsilon)\left(1-s_i^M\right)\Gamma_i \\ H_i^{ZT} = -\left[\gamma+1-s_i^M+\epsilon s_i^M\right]\Gamma_i & H_i^{Z\omega} = -\,(1+\gamma)\left(1-s_i^M\right)\Gamma_i \end{array} \tag{D.27}$$

while the matrices summarizing the interrelation between $\left\{\hat{W}, \hat{D}, \hat{Z}\right\}$ have i^{th}-diagonal elements given by:

$$\begin{array}{lll} H_i^{WT} = (\sigma-1)\,\Gamma_i & H_i^{WD} = \Gamma_i & H_i^{WZ} = -\,(\sigma-\epsilon)\,s_i^M\Gamma_i \\ H_i^{DT} = (\gamma+\epsilon)\,(\sigma-1)\,\Gamma_i & H_i^{DD} = (\gamma+\epsilon)\,\Gamma_i & H_i^{DZ} = (\sigma-\epsilon)\,(\gamma+\sigma)\left(1-s_i^M\right)\Gamma_i \\ H_i^{ZT} = -\left[\gamma+1-s_i^M+\epsilon s_i^M\right]\Gamma_i & H_i^{ZD} = \left(1-s_i^M\right)\Gamma_i & H_i^{ZZ} = (\gamma+\epsilon)\,s_i^M\Gamma_i \end{array} \tag{D.28}$$

Now, first note the existence of *feedback effects* arising from the fact that marginal costs are increasing with scale due to the upward-sloping labor supply curves faced by each firm. These feedback effects go in two directions. To illustrate, consider a simple supply chain $j_s \to i \to j_c$, where arrows indicate the flow of goods. First, consider a positive demand shock to customer j_c. This leads to an increase in demand D_i for firm i (H^{DD}), which not only has an effect on firm i's earnings effect as in Proposition 1 (H^{WD}), but also leads to an increase in marginal cost and hence in the output price for firm i, thus raising the cost of materials for customer j_c (H^{ZD}). This in turn has a feedback effect on the demand from customer j_c (H^{DZ}). Second, consider an increase in material cost for supplier j_s. This raises the material cost for firm i (H^{ZZ}), which not only has an effect on firm i's earnings effect as in Proposition 1 (H^{WZ}), but also affects the demand for materials by firm i and hence the demand faced by supplier j_s (H^{DZ}). This in turn has a feedback effect on the marginal cost and output price of the supplier, and hence on the cost of materials faced by firm i (H^{ZD}).

In sum, feedback effects stemming from scale-dependent marginal costs are captured by elements of the product $H^{ZD}H^{DZ}$ (which is symmetric, given that the H matrices are diagonal). Given our estimates of $\{\gamma, \sigma, \epsilon\}$ and evaluating firm-specific material cost shares at the median value in the average year in our sample, the magnitudes of the elements of these matrices are approximately $H_i^{DZ} \approx 0.29$ and $H_i^{ZD} \approx 0.02$, so that the feedback elasticity is approximately 0.6%. Hence, feedback effects are likely to be small empirically. Ignoring these feedback effects by setting H^{DZ} and H^{ZD} to zero in equations (D.25) and (D.26) and solving for $\hat{W}$ as a function of $\hat{D}$ and $\hat{Z}$ then gives the passthrough expressions in Proposition 1. More generally, however, it is straightforward to solve the linear system (D.24)-(D.26) for $\hat{W}$ including feedback effects, with coefficients that can be fully determined given estimates of $\{\gamma, \sigma, \epsilon\}$, network shares

$\left\{S^{sales}, S^{mat}\right\}$, and material shares of cost s_i^M.

E Data Details

E.1 Data cleaning

To clean the firm-to-firm trade dataset, we drop relationships involving firms that do not report value-added or employment, or firms that report negative value-added, sales, or materials. We also follow Bernard et al. (2022) and iteratively drop firms that have only one relationship, which is required for a decomposition of firm-to-firm transaction values into buyer and seller effects that we describe below.

To clean the employer-employee dataset, we impose sample restrictions following the criteria outlined in Lamadon et al. (2022). In each year, we start with all individuals aged 25-60 who are linked to at least one employer. We identify links using only information on labor contracts (tax affidavit 1887). Next, we drop firms that have missing or negative value-added, sales, or materials in the balance sheet data (tax form 29). Then, we keep for each worker the firm that pays the highest earnings in a given year. Since we do not have hours worked or a direct measure of full-time employment, we follow the literature by including workers for whom annual earnings are above a minimum threshold (Song et al., 2019). We set the threshold equal to 32.5% of the national average of earnings in order to make our estimates comparable to the cross-country study of earnings inequality in Bonhomme et al. (2020).

E.2 Sample sizes and descriptive statistics

Table A.1 provides sample size information for our baseline firm-to-firm dataset, employer-employee dataset (including the movers and stayers subsamples), and firm-level dataset, which are defined in section 6.1. Table A.2 provides basic descriptive statistics about these datasets.

Table A.1: Overview of Sample Sizes

Panel A: Firm-to-Firm Dataset	**Links**		**Suppliers**		**Buyers**	
Sample	**Unique**	**Observation-Years**	**Unique**	**Observation-Years**	**Unique**	**Observation-Years**
Baseline	16,831,546	31,743,495	194,615	592,622	289,344	923,155
Panel B: Employer-Employee Dataset	**Workers**		**Firms**			
Sample	**Unique**	**Observation-Years**	**Unique**	**Observations-Years**		
Baseline	6,496,849	41,954,008	487,504	2,315,927		
Movers	6,183,692	40,130,960	200,592	1,378,554		
Stayers: Complete Spells	953,865	8,472,302	64,670	602,622		
Stayers: 10 Stayers per Firm	724,957	6,571,483	5,726	61,823		
Panel C: Firm Dataset	**Firms**					
Sample	**Unique**	**Observations-Years**				
Baseline	47,685	125,726				

Notes: This table provides an overview of the samples used throughout the paper.

F Estimation Details

F.1 Labor supply elasticity

We first formally describe the identification of γ in the presence of measurement error in wage bills. To this end, suppose that wage bills in the data $\ddot{E}_{it}^L$ are related to wage bills in the model

Table A.2: Descriptive Statistics of Datasets

Dataset	**Employer-Employee**			**Firm**	**Firm-to-Firm**
Panel A: Worker Characteristics	**Baseline**	**Movers**	**Stayers**	**Baseline**	**Baseline**
Mean Log Worker Earnings (Log US $)	9.36	9.38	9.74	9.17	9.22
Median Log Worker Earnings (Log US $)	9.25	9.27	9.66	9.02	9.10
Mean Worker Age	40.2	40.1	42.6	39.3	39.8
Median Worker Age	39.4	39.4	42.6	38.5	39.0
Panel B: Firm Characteristics	**Baseline**	**Movers**	**Stayers**	**Baseline**	**Baseline**
Mean Number of Workers	9	20	281	27	12
Median Number of Workers	2	4	94	7	2
Mean Wage Bill per Worker (US $)	10,199	11,145	7,833	9,440	8,306
Median Wage Bill per Worker (US $)	6,943	8,323	6,672	7,103	5,490
Mean Value Added per Worker (US $)	56,315	58,610	50,077	49,604	50,091
Median Value Added per Worker (US $)	23,424	25,659	26,583	23,389	18,771
Mean Log Value Added (Log US $)	11.0	11.8	14.6	12.2	10.9
Median Log Value Added (Log US $)	11.0	11.7	14.8	12.1	10.9
Mean Labor Share	0.49	0.45	0.70	0.42	0.49
Median Labor Share	0.32	0.34	0.21	0.34	0.32
Panel C: Production Network Characteristics	**Baseline**	**Movers**	**Stayers**	**Baseline**	**Baseline**
Mean Number of Suppliers	67	67	306	67	35
Median Number of Suppliers	36	36	208	36	19
Mean Number of Buyers	80	80	580	80	34
Median Number of Buyers	8	8	59	8	4
Mean Materials Share of Sales	0.58	0.58	0.55	0.58	0.57
Median Materials Share of Sales	0.61	0.61	0.60	0.61	0.61
Mean Intermediate Share of Sales	0.40	0.40	0.45	0.40	0.38
Median Intermediate Share of Sales	0.38	0.38	0.50	0.38	0.33

Notes: This table provides descriptive statistics of all the samples used in the paper.

E_{it}^L as follows:

$$\log E_{it}^L = \log \ddot{E}_{it}^L + e_{it}^L \tag{F.1}$$

where e_{it}^L denotes an MA(k) measurement error given by $e_{it}^L = \sum_{s=0}^{k} \delta^{L,s} u_{i,t-s}^L$ for some weights $\delta^{L,s}$ and mean-zero shocks u_{it}^L that are iid across firms and time. In this case, equation (6.4) becomes:

$$\Delta \log w_{imt} = \frac{1}{1+\gamma} \Delta \log \ddot{E}_{it}^L + \Delta \log \hat{a}_{mt} + \frac{1}{1+\gamma} \Delta e_{it}^L \tag{F.2}$$

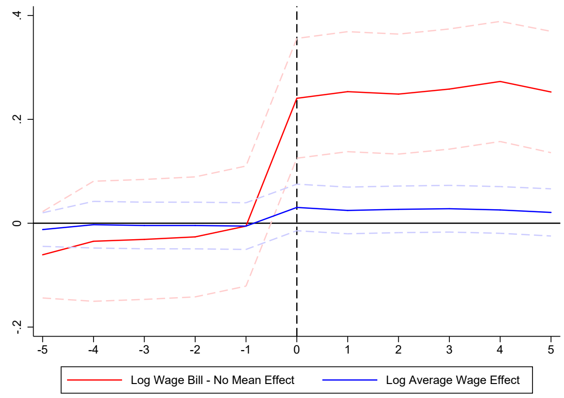
In the absence of measurement error, the residual in equation (F.2) contains only worker-level shocks ($\Delta \log \hat{a}_{mt}$), which are orthogonal to changes in firm wage bills under Assumption 6.5. However, with measurement error in wage bills, the unobserved error term in equation (F.2) contains a component that is potentially correlated with observed changes in the wage bill.

To address this, note that Δe_{it}^L depends only on measurement error shocks u_{it}^L in periods $\{t-k-1, \cdots, t\}$. Hence, as long as u_{it}^L is orthogonal to all *lagged* innovations in the Markov processes for time-varying firm primitives, lagged changes in wage bills $\log \Delta \ddot{E}_{is}^L$ for any $s < t-k-1$ are valid instruments for $\log \Delta \ddot{E}_{it}^L$ in identifying γ from equation (F.2). The relevance of these instruments requires serial correlation in $\Delta \ddot{E}_{it}^L$ to be non-zero with at least $k+2$ lags, which is also consistent with the Markov processes for firm productivities specified in Assumption 6.4.

For robustness, we also follow Lamadon et al. (2022) and estimate γ using a difference-in-difference approach (DiD). For this, we follow a three step procedure. First, for each year, we order firms according to log changes of the wage bill of the firm. Second, we identify the treatment when firms have log changes of their wage bill above the median of log changes of wage bill across firms each year. Finally, we plot difference in wage bill of treated and control firms both at each year ($t = 0$) and years before ($t < 0$) and after ($t > 0$). We perform this step for each calendar year and weight firms by the number of workers.

Results are presented Figure A.1. By construction, the treatment and control groups differ in the wage bill from period $t = -1$ to $t = 0$. On average, firms in the treatment group face an increase of 21 log points growth in their wage bill relative to firms in the control group. The effect of the treatment appears to be permanent in levels up to 5 years after the treatment. Figure A.1 also shows the effect on the average earnings of firms. On average, firms in the treatment group face an increase of 3.25 log points of their average earnings relative to firms in the control group. Once again, the effect of the treatment appears to be permanent in levels up to 5 years after the treatment. Finally, firms in the treatment and control group do not experience statistically significant differences up to 5 years before the treatment, for both the wage bill and the average earnings. Through the lens of a DiD design, these results imply a passthrough rate of firms shocks of around 0.155 ($= 0.0325/0.21$). From equation (6.3), this implies a labor supply elasticity of $\hat{\gamma} = 5.5$, which is the same as our preferred estimate documented in the main text.

Figure A.1: Difference-in-difference Estimate of passthrough of Firm Shocks to Worker Earnings



Notes: This figure presents the results from the Lamadon et al. (2022) difference-in-difference approach to estimating passthrough of wage bill shocks to worker wages.

F.2 Worker and firm wage effects

To estimate the Bonhomme et al. (2019) decomposition of worker earnings from equation (6.5), we first cluster firms using a k-means clustering algorithm into $K = 10$ groups. We use a weighted K-means algorithm with 100 randomly generated starting values. We use firms' empirical distributions of log earnings on a grid of 10 percentiles of the overall log-earnings distribution. Second, we use these K groups as the relevant firm identifier in the Bonhomme et al. (2019) estimation approach. This procedure yields estimates of the firm fixed effect $\bar{W}_i$ and the worker-firm production complementarity θ_i for every firm $i \in \Omega^F$, as well as the permanent and transient components of ability for every worker.

To assess robustness of our results to the number of clusters used, Table A.3 documents the share of variance of wages accounted for by the firm fixed effect $\bar{W}_i$. We implement this for the basic model of Abowd et al. (1999) and also the basic version of the model of Bonhomme et al. (2019) with only firm and worker fixed effects for different levels of K (thus, excluding interactions and time-varying firm effects). First, one can see that the basic version of the model of Bonhomme et al. (2019) implies a role for the firm fixed effect that is significantly lower than the model of Abowd et al. (1999), consistent with previous literature that has found that addressing the limited mobility bias inherent in estimates of Abowd et al. (1999) decreases the share of the variance accounted for by the firm fixed effect (Bonhomme et al., 2020). Second, as one increases K from 10 to 50, the share of the variance of wages accounted for the firm fixed effects increases only 0.7 percentage points from 7.8 to 8.5%. At least with this piece of evidence, this implies that the limited mobility bias does not represent a substantially bigger problem for $K = 50$ than what it represents for $K = 10$.

Table A.3: Share of Log Earnings Variance Accounted for by the Firm Fixed Effect

Estimation Strategy	Number of Clusters	Firm Fixed EffectShare
AKM		12.3
BLM	10	7.8
BLM	50	8.5

Notes: This table documents the share of the log of earnings variance accounted for by the firm fixed effect. It is documented for the estimation strategy of Abowd et al. (1999) (row 1), for the estimation strategy of Bonhomme et al. (2019) with $K = 10$ clusters (row 2) and the estimation strategy of Bonhomme et al. (2019) with $K = 50$ clusters (row 3).

To further assess whether clustering with $K = 10$ or $K = 50$ makes a difference, we document how much clusters account for the variance of firm-level characteristics. Tables A.4-A.5 document the share of the variance of variables accounted for by within-cluster variation. Table A.4 shows the within-cluster share of variance of variables in levels, whereas Table A.5 shows the same evidence for variables in ratios. Although there is substantial heterogeneity across firms that the clustering procedure of Bonhomme et al. (2019) does not account for, this result does not vary significantly if one uses $K = 10$ or $K = 50$ clusters.

Table A.4: Within Clusters Share of Total Variance of Variables in Levels

Number Clusters	Total Sales	Materials	Wage Bill	Employment	Number of Buyers	Number of Suppliers	Firm-to-Firm Sales
10	79	90	67	88	90	85	95
50	74	86	62	84	88	81	92

Notes: This table documents the share of the variance of each variable accounted for by the within cluster variance. It is implemented for $K = 10$ and $K = 50$ and for variables in levels.

Table A.5: Within Clusters Share of Total Variance of Variables in Ratios

Number Clusters	Wage Bill/Sales	Materials/Sales	Materials/Wage Bill	Sales/Employment	Wage Bill/Employment	Materials/Employment
10	96	97	95	92	26	99
50	95	97	95	90	21	98

Notes: This table documents the share of the variance of each variable accounted for by the within cluster variance. It is implemented for $K = 10$ and $K = 50$ and for variables in ratios.

F.3 Amenities

To estimate firm amenities, we begin with the labor supply equation (3.2). It will be useful for the exposition to write this explicitly in terms of permanent and transient worker abilities:

$$\frac{L_{it}\left(\bar{a}, \hat{a}\right)}{L\left(\bar{a}, \hat{a}\right)} = \frac{\left[g_i\left(\bar{a}\right) w_{it}\left(\bar{a}, \hat{a}\right)\right]^{\gamma}}{\sum_{j \in \Omega^F}\left[g_j\left(\bar{a}\right) w_{jt}\left(\bar{a}, \hat{a}\right)\right]^{\gamma}} \tag{F.3}$$

where note that under Assumption 6.1, amenity values only vary across workers in relation to permanent ability $\bar{a}$. Next, consider the equilibrium wage equation (3.16). Under assumption

6.1, we can write this as:

$$w_{it}(\bar{a},\hat{a}) = \eta\bar{a}^{\theta_i}\hat{a}W_{it} \tag{F.4}$$

The average wage paid by firm i to workers with permanent ability $\bar{a}$ is hence:

$$\bar{w}_{it}(\bar{a}) = \eta\bar{a}^{\theta_i}\mathbb{E}[\hat{a}]W_{it} \tag{F.5}$$

where $\mathbb{E}[\hat{a}]$ denotes the average value of transient ability. Under Assumptions 6.1 and 6.5, this mean does not depend on permanent ability of the worker or the identity of the firm. Combining (F.4) and (F.5), we then have:

$$w_{it}(\bar{a},\hat{a}) = \bar{w}_{it}(\bar{a})\frac{\hat{a}}{\mathbb{E}[\hat{a}]} \tag{F.6}$$

Substituting this into (F.3) and using the decomposition of amenities in equation (6.11), we obtain:

$$\frac{L_{it}(\bar{a},\hat{a})}{L(\bar{a},\hat{a})} = \frac{\left[\tilde{g}_i\bar{g}_{k(i)}(\bar{a})\,\bar{a}^{\theta_{k(i)}}W_{it}\right]^{\gamma}}{\sum_j\left[\tilde{g}_j\bar{g}_{k(j)}(\bar{a})\,\bar{a}^{\theta_{k(j)}}W_{jt}\right]^{\gamma}} \tag{F.7}$$

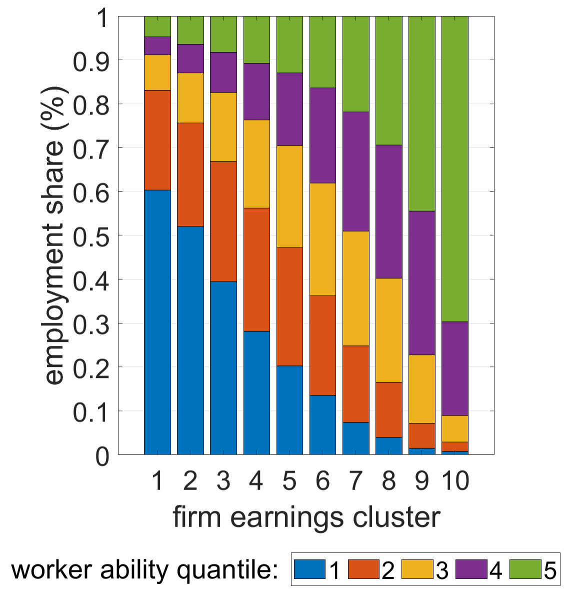
Now notice that the employment share of workers of ability $\{\bar{a},\hat{a}\}$ varies across firms only in relation to permanent ability $\bar{a}$. This is a direct implication of Assumption 6.1, which implies that workers do not sort to firms based on transient ability $\hat{a}$. Therefore, the share of workers of permanent ability $\bar{a}$ employed by firm i is also given by equation (F.7). Summing this (F.7) across all firms within cluster k, we can similarly express the share of workers of permanent ability $\bar{a}$ that are employed by firms in cluster k as:

$$\Lambda_{kt}(\bar{a}) = \frac{\sum_{i\in k}\left[\tilde{g}_i\bar{g}_k(\bar{a})\,\bar{a}^{\theta_k}W_{it}\right]^{\gamma}}{\sum_j\left[\tilde{g}_j\bar{g}_{k(j)}(\bar{a})\,\bar{a}^{\theta_{k(j)}}W_{jt}\right]^{\gamma}} \tag{F.8}$$

Next, note that for each value of permanent ability $\bar{a}$, equilibrium outcomes are invariant to scaling $g_i(\bar{a})$ by a constant for all firms i. Therefore, we are allowed to choose one normalization of amenity values for each permanent worker ability type $\bar{a}$. For this, we choose $\sum_j\left[\tilde{g}_j\bar{g}_{k(j)}(\bar{a})\,\bar{a}^{\theta_{k(j)}}W_{jt}\right]^{\gamma} = 1$. Furthermore, mean differences in amenity values can be loaded onto either $\tilde{g}_i$ or $\bar{g}_{k(i)}(\bar{a})$. Hence, we are allowed to choose one normalization of the values for $\tilde{g}_i$ for each firm cluster. For this, we choose $\sum_{i\in k}[\tilde{g}_iW_{it}]^{\gamma} = 1$. With these normalizations, equations (6.12) and (6.13) follow immediately.

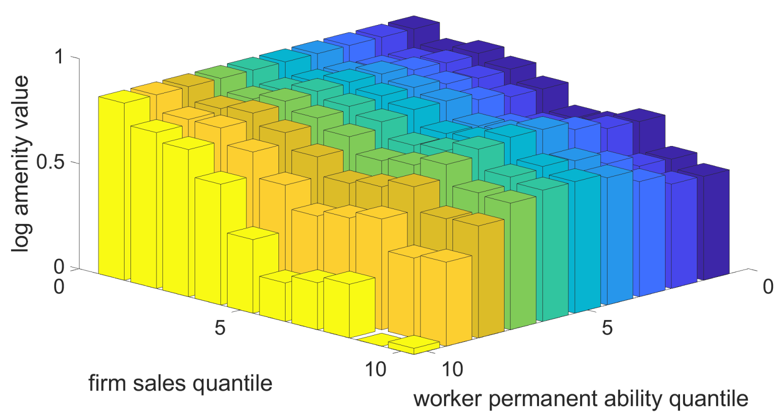
Our results are summarized in Figure A.3, which shows average log amenity values by deciles of firm sales and worker permanent ability. Evidently, larger firms tend to offer lower amenity values to workers of each ability type, with this relationship being more pronounced for workers of higher permanent ability. Furthermore, as shown in Figure A.2, our estimates of amenities and production complementarities imply positive sorting of workers to firms. Note that by construction, the model provides an exact fit to the cluster-level employment shares shown in the figure.

Figure A.2: Employment shares by firm earnings cluster and worker ability



Notes: Firm earnings clusters are sorted in ascending order of the time-invariant firm earnings effect, $\bar{W}_{k(i)}$.

Figure A.3: Distribution of Amenities



Notes: This figure shows the joint distribution of amenity estimates $\log g_i(\bar{a})$ by deciles of firm sales and worker permanent ability. Values are normalized for presentation purposes such that: (i) average log amenities within the smallest decile of firm sales are equal across deciles of worker permanent ability, and (ii) the smallest value of mean log amenities across sales-ability quantiles is equal to zero.

F.4 Firm relationship capability and relationship-specific productivity

To estimate equation (6.7), firms must have multiple connections. To identify seller fixed effects, each seller needs to have at least two buyers. Similarly, to identify buyer fixed effects, each buyer needs to have at least two sellers. In the data, some firms have either one supplier or one seller. Hence, we implement the aforementioned restriction using an iterative approach known as "avalanching". Specifically, we first drop firms with one supplier or seller. Doing this may result in additional firms that have one supplier or seller, hence in the next step, we drop these firms as well. We continue this process until firms are no longer dropped from the sample. The algorithm takes three iterations to converge in practice and reduces the sample size of firm-to-firm linkages from a total of 32 million transactions to 31.7 million transactions, that is, a reduction of 1% of transactions. Hence, the avalanching algorithm has little impact on our sample size. Bernard et al. (2022) report that avalanching also eliminates around 1% of firm-to-firm links in the production network for Belgium.

F.5 Product substitution elasticity

To derive equation (6.9), first note that the share of firm profits in total sales can be expressed as:

$$\frac{\pi_{it}}{R_{it}} = \frac{1}{\sigma}\left[1 + \frac{(\sigma-1)(1-\eta)}{1+\eta\frac{E_{it}^M}{E_{it}^L}}\right] \tag{F.9}$$

Solving for σ and using the fact that $\pi_{it} = R_{it} - E_{it}^L - E_{it}^M$ gives equation (6.9). Hence, we estimate σ using the sample average of the right-hand side of (6.9), which is observable given our estimate of the labor supply elasticity γ and data on firm sales, labor costs, and material costs.

F.6 Labor-materials substitution elasticity and labor productivity

For estimation of ϵ using equation (6.10), we follow the approach in Doraszelski and Jaumandreu (2018). To control for $F^{\omega}(\omega_{i,t-1})$, we first rearrange the $t-1$ version of equation (6.10) to write:

$$\log\omega_{i,t-1} = \frac{1}{\epsilon-1}\log\left[\frac{1}{\eta}\left(\frac{1-\lambda}{\lambda}\right)\right] - \frac{1}{\epsilon-1}\log\frac{E_{i,t-1}^M}{E_{i,t-1}^L} + \log\frac{W_{i,t-1}}{Z_{i,t-1}} \tag{F.10}$$

$$\equiv G\left(\log\frac{E_{i,t-1}^M}{E_{i,t-1}^L}, \log\frac{W_{i,t-1}}{Z_{i,t-1}}\right) \tag{F.11}$$

Substituting this into (6.10), we obtain:

$$\log\frac{E_{it}^M}{E_{it}^L} = \log\left[\frac{1}{\eta}\left(\frac{1-\lambda}{\lambda}\right)\right] + (\epsilon-1)\log\frac{W_{it}}{Z_{it}} + H\left(\log\frac{E_{i,t-1}^M}{E_{i,t-1}^L}, \log\frac{W_{i,t-1}}{Z_{i,t-1}}\right) \tag{F.12}$$

$$+ (1-\epsilon)\,\xi_{it}^{\omega} \tag{F.13}$$

where $H(\cdot,\cdot) \equiv (1-\epsilon) F^{\omega}[G(\cdot,\cdot)]$. Hence, we control for the term H using polynomials in lagged relative expenditures $\log \frac{\tilde{E}^M_{i,t-1}}{\tilde{E}^L_{i,t-1}}$ and lagged relative input prices $\log \frac{\tilde{W}_{i,t-1}}{Z_{i,t-1}}$. In addition, we follow Doraszelski and Jaumandreu (2018) and instrument for relative input prices at date t using polynomials in one-period lags of logged input expenditures and factor prices.

F.7 Firm TFP

We choose values for TFP T_{it} to fit the estimated firm-level wages W_{it} as specified in equation (6.14). We do this using an iterative numerical procedure that is similar in spirit to the equilibrium solution algorithm described in section H:

1. Compute $\left\{\bar{\phi}_{it}\right\}_{i\in\Omega^F}$ from (3.20), using (3.4), (3.3), and the estimated firm-level wages $\{W_{it}\}_{i\in\Omega^F}$.
2. Guess E_t.
   (a) Guess $\{D_{it}, Z_{it}\}_{i\in\Omega^F}$.
   (b) Compute the values of $\{T_{it}\}_{i\in\Omega^F}$ implied by equation (H.1), given the estimated firm-level wages $\{W_{it}\}_{i\in\Omega^F}$.
   (c) Compute new guesses of $\{D_{it}\}_{i\in\Omega^F}$ from (3.10) and $\{Z_{it}\}_{i\in\Omega^F}$ from (3.12).
   (d) Iterate on steps (a)-(c) until convergence.
3. Compute a new guess of E_t from (3.23), using (3.2), (3.14), and (3.16).
4. Iterate on steps 1-2 until convergence.

G Construction of export demand and import cost shocks

To construct export demand shocks, suppose that Chilean exporter i sells to a set of export markets $\Omega^{M,X}_{it}$, with each market comprised of a representative customer with exogenous buyer effect Δ^F_{mt}. Then, we can write the firm's demand as:

$$D_{it} = E_t + \sum_{j\in\Omega^{C,D}_{it}} \Delta_{jt}\psi_{jit} + \sum_{m\in\Omega^{M,X}_{it}} \Delta^F_{mt}\psi^F_{mit} \tag{G.1}$$

where $\Omega^{C,D}_{it}$ now denotes the set of firm i's *domestic* customers and ψ^F_{mit} accounts for firm heterogeneity in export demand from each export market m. Then, differentiating (G.1) with respect to $\left\{\Delta^F_{mt}\right\}_{m\in\Omega^{M,X}_{it}}$ allows us to write:

$$\hat{D}_{it} = s^{sales}_{Xit} \sum_{m\in\Omega^{M,X}_{it}} s^X_{mit}\hat{\Delta}^F_{mt} \tag{G.2}$$

Now let E^I_{mt} denote the total value of imports by market m from all countries excluding Chile in year t. Under the assumption that foreign customers have the same CES preferences as consumers in Chile, imports by market m are given by:

$$\log E_{mt}^{I} = (1-\sigma)\log Z_{mt}^{F} + \log \Delta_{mt}^{F} \tag{G.3}$$

where Z_{mt}^{F} is the CES price index that market m faces for its non-Chilean imports. We further suppose that these price indices can be decomposed as $Z_{mt}^{F} = \tau_m z_{h(m)t}^{F}$, where τ_m captures static differences in import costs across markets, while $z_{h(m)t}^{F}$ captures time-varying differences in import costs that are equal for all importing countries within the product category $h(m)$ to which market m belongs. Hence, taking first differences of equation (G.3), we obtain:

$$\hat{E}_{mt}^{I} = \delta_{h(m)t} + \hat{\Delta}_{mt}^{F} \tag{G.4}$$

where $\delta_{h(m)t} \equiv (1-\sigma)\,\hat{z}_{h(m)t}$ is a product-year fixed effect that is orthogonal to $\hat{\Delta}_{mt}^{F}$ by assumption. Now we assume that the expectation of $\hat{\Delta}_{mt}^{F}$ across countries is zero, so that there is no aggregate growth in demand within a product category. The log change in market m's buyer effect is then equal to the log change in E_{mt}^{I} relative to the corresponding change in world imports of product category $h(m)$. Given this, the first expression in (5.2) follows immediately from (G.2).

Similarly, to construct import cost shocks, suppose that Chilean importer i purchases imported materials from a set of markets $\Omega_{it}^{M,I}$, with each market comprised of a representative supplier that charges price p_{mt}^{X}. Then, we can write the firm's material input cost as:

$$Z_{it} = \left[\sum_{s\in\Omega_{it}^{S,D}} (p_{st})^{1-\sigma} + \sum_{m\in\Omega_{it}^{M,I}} \left(p_{mt}^{X}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{G.5}$$

Differentiating this with respect to $\left\{p_{mt}^{X}\right\}_{m\in\Omega_{it}^{M,I}}$ and replacing contemporaneous import shares with initial import shares as before then gives the second expression in (5.2).

H Solution Algorithm

We solve numerically for an equilibrium of the model using the following solution algorithm.

1. Guess E_t.
   (a) Guess $\left\{\Delta_{it}, \Phi_{it}, \bar{\phi}_{it}\right\}_{i\in\Omega^F}$.
   (b) Compute $\{D_{it}\}_{i\in\Omega^F}$ from (3.10) and $\{Z_{it}\}_{i\in\Omega^F}$ from (3.12).
   (c) Solve for $\{W_{it}, \nu_{it}, X_{it}\}_{i\in\Omega^F}$ from (3.15), (3.17), and (3.18).
   (d) Compute new guesses of $\{\Delta_{it}\}_{i\in\Omega^F}$ from (3.11), $\{\Phi_{it}\}_{i\in\Omega^F}$ from (3.13), and $\left\{\bar{\phi}_{it}\right\}_{i\in\Omega^F}$ from (3.20).
   (e) Iterate on steps (a)-(d) until convergence.
2. Compute a new guess of E_t from (3.23), using (3.2), (3.14), and (3.16).
3. Iterate on steps 1-2 until convergence.

Note that step 1(c) involves numerical solution of a system in $\{W_{it}, \nu_{it}, X_{it}\}$. This system can be reduced to one in firm-level wages alone:

$$W_{it}^{\gamma+\epsilon}\left[\lambda\left(W_{it}/\omega_{it}\right)^{1-\epsilon}+(1-\lambda)\,Z_{it}^{1-\epsilon}\right]^{\frac{\sigma-\epsilon}{1-\epsilon}}\bar{\phi}_{it}=\frac{\lambda}{\mu^{\sigma}\eta^{\gamma}}D_{it}T_{it}^{\sigma-1}\omega_{it}^{\epsilon-1} \tag{H.1}$$

which has a unique solution for W_{it} given $\left\{D_{it}, Z_{it}, \bar{\phi}_{it}\right\}$. Solutions for ν_{it} and X_{it} are then easy to recover given W_{it}.

I A Shapley value approach for model counterfactuals

In the counterfactual exercises studied in section 7, we deal with interdependencies between sources of variation in shaping inequality outcomes using the following approach. Let Θ denote the estimated vector of values for all model primitives and let $X(\Theta)$ denote the value of some equilibrium outcome X under this parameter vector. Now, define some N subsets of the parameter vector $\{\theta_n\}_{n=1}^{N}$ such that $\Theta=\cup_{n=1}^{N}\theta_n$ and denote $\mathcal{N}\equiv\{1,\cdots,N\}$. We are interested in computing values of outcome X under known counterfactual values $\hat{\theta}_n$ for each subset of the parameter vector. Therefore, let $\hat{\Theta}_S\equiv\left\{\cup_{n\in S}\hat{\theta}_n\right\}\cup\{\cup_{n\notin S}\theta_n\}$ denote the parameter vector under counterfactual values for parameter subsets in S for some $S\subseteq\mathcal{N}$. We define the Shapley value X_n for parameter subset n in relation to outcome X as follows:

$$X_n=\sum_{S\subseteq\mathcal{N}\setminus\{n\}}\frac{|S|!\,(N!-|S|!-1)}{N!}\left[X\left(\hat{\Theta}_{S\cup\{n\}}\right)-X\left(\hat{\Theta}_S\right)\right] \tag{I.1}$$

For example, suppose that X is the variance of log earnings across all workers, θ_n is the estimated vector of firm TFPs, and $\hat{\theta}_n$ is a counterfactual vector of firm TFPs with each value equal to the mean of θ_n across firms. Then, we measure the contribution of TFP heterogeneity to earnings variance as $-\frac{X_n}{X(\Theta)}$. By construction of the Shapley value, these measures sum to one across all $n\in\mathcal{N}$.