Who gains from scale?

Trade and Wage Inequality Within and Between Firms

Fabian Trottner*

UC, San Diego

October 2022

Abstract

This paper provides new theory and evidence on the importance of firm and worker heterogeneity for aggregate wage inequality and welfare. Using German micro-data, I show that relative wages and employment of skilled compared to less-skilled workers are higher at larger firms, suggesting that decisions regarding scale, the demand and wages for skill are interconnected within firms. I develop a model in which firms operate a non-homothetic production technology and hire heterogeneous workers in monopsonistic labor markets. The model provides a unified framework to study the joint determination of the firm size distribution, wage and skill distributions within firms, and aggregate wage inequality. I structurally estimate the model using a new method that separately identifies the elasticities of labor demand and supply. Quantitatively, I utilize trade liberalization to study how a shock that initiates changes in the firm size distribution impacts aggregate inequality by changing wage distributions within and between firms. I find that trade raises inequality by 20 percent with within-firm effects accounting for 30 percent of the overall change. Turning to welfare, I show that a tax reform that corrects misallocations due to labor market frictions raises the gains from trade for all workers by improving worker-to-firm sorting and redistributing income from firm profits to wages.

*Contact: ftrottner@ucsd.edu. I am very grateful to my advisors for their support, encouragement, and guidance: Steve Redding, Gene Grossman, and Ezra Oberfield. I have also learned and benefitted a lot from discussions and feedback on earlier drafts from Fabian Eckert, Cecilia Fieler, Oleg Itskhoki, Eduardo Morales, Fernando Parro, Richard Rogerson, and seminar participants at numerous venues. This study uses the Linked-Employer-Employee Data (LIAB) longitudinal model 1993-2014 (LIAB LM 9314) from the IAB. Data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and subsequently remote data access. I thank the FDZ and the IRS at Princeton for data access.

1 Introduction

The distributional effects of international trade are closely tied to the importance of firm and worker heterogeneity for aggregate wage inequality. First, when firms differ in their demand for different worker types, e.g., skill, trade-induced reallocations of resources between firms may shift the aggregate demand and returns to skill. Second, when labor markets are not perfectly competitive, trade liberalization may change cross-firm wage dispersion for similar workers. Finally, reallocations between firms may impact firm-worker sorting patterns; particularly which types of workers are more likely to be matched to firms offering high wage premiums.

The channels linking firm and worker heterogeneity to the distributional effects of trade - firm-level differences in exposure to trade shocks, skill composition, and wages - are not mutually exclusive. However, the literature lacks a unified quantitative framework accounting for these three channels. In this paper, I propose a tractable model that captures the equilibrium effects of shocks to the firm size on the wage distribution featuring realistic firm-worker sorting patterns, firm wage premia, and within-firm wage distributions. The first half of the paper provides motivating facts on the joint distribution of wages and skill across firms from German matched employer-employee data, describes the model and discusses its theoretical properties. The second half presents a quantitative assessment of the distributional and welfare implications of trade liberalization.

I begin by collecting three stylized facts from German matched employer-employee data about relative wages and relative employment for different types of workers across firms. Specifically, the data show that while larger firms pay higher wages to workers of all skill levels,[1] wages of more skilled workers are more strongly associated with firm sales, labor productivity, and value added. Within firms, larger firms pay higher relative wages to more skilled employees. Further, the relative employment shares of more skilled workers are positively related to firm size: High skill workers are more likely to be employed by larger firms. These features suggest that a firm's decision to scale production, e.g., the decision to export,[2] is interconnected to within-firm distributions of skill and wages.

I develop a model that accounts for these features of the data by combining non-homothetic technologies (Hanoch (1975), Sato (1975)) with an upward-sloping labor supply function for each type of labor. Non-homotheticity in technology implies that a firm's level of output explicitly affects the relative productivity of different types of workers.[3] As a result, firms jointly decide the skill composition of their workforce and

[1] A large empirical literature documents that larger firms pay higher wages (e.g. Brown & Medoff (1989), Oi & Idson (1999)), and that firm-size wage gaps are potentially important for aggregate inequality (e.g. Card *et al.* (2013), Alvarez *et al.* (2018), Helpman *et al.* (2017), Song *et al.* (2018)).

[2] It is a well-established empirical fact that exporters are more skill intensive, and pay higher wages (e.g. Bernard & Jensen (1995), Bernard *et al.* (2007), Verhoogen (2008)).

[3] I provide several microfoundations for this technology. In particular, I show that the non-homothetic CES

their scale,[4] relative labor demands for skill types differ across firms with different levels of output, and firm productivity is jointly determined with its scale. In the labor market, firms face upward-sloping labor supply function for each skill type. Workers receive heterogeneous preference shocks over non-wage job characteristics that are unobservable to firms. A firm seeking to expand employment posts higher wages so that more employees choose to accept the job. As firms cannot price-discriminate, all employees benefit from an increase in the wage, and workers earn rents from employment relationships. Consistent with the reduced form facts, the model predicts that skill premia and skill intensity vary endogenously with firm size and labor productivity.

In the model, the relationship between firm size, skill composition, and the skill wage premium depends on both the degree of non-homotheticity in production and on the relative curvature of type-specific labor supply curves. Because the elasticity of labor supply determines marginal hiring costs and worker rents, the model predicts that workers in less elastic supply receive disproportionately higher wages at larger employers. Further, firm size determines the relative productivity of skill groups, and thereby impacts the relative wages that skill types earn across employers - skill types whose relative productivity increases in firm size receive relatively higher wages at larger firms. Similar forces determine the relationship between size and workforce skill composition: Ceteris paribus, larger firms hire relatively more skill types who are in more elastic supply, and whose relative productivity is increasing in firm size.

To assess the aggregate distributional consequences of an exogenous shock to the firm size distribution, it is critical to separately identify the parameters that govern non-homotheticity in labor demand and the elasticity of the labor supply for different skill types. I achieve this by developing an estimation strategy that identifies the elasticities of labor demand independently of the curvature of labor supply. To address the simultaneity bias that arises from upward-sloping labor supply curves and to separately identify non-homotheticities from exogenous differences in labor demand, I build on approaches to partial identification from Leamer (1981) and Feenstra (1994). I show that an exogenous shock to firm demand is sufficient to provide partial identification of the production function's output elasticities and the elasticity of substitution, which in turn fully characterize the elasticities of labor demand. To the extent that regional labor markets (i.e., East and West Germany) are subject to different shocks, the additional cross-regional variation provides full identification of the labor demand elasticities. The estimates of the production technology are then used to infer labor supply elasticities by matching

captures a continuous generalization of the binary technology choice model in (Bustos (2011b)).

[4] The cross-sectional predictions are similar to models based on skill-bias in technology (Harrigan & Reshef (2015), Burstein & Vogel (2017)). In contrast to these models, here market size is itself a determinant of skill composition. Thus, a firm's production efficiency is jointly determined with its skill composition and output, similar to knowledge-based models of production hierarchies Lucas (1978), Rosen (1981), Garicano (2000), Garicano & Rossi-Hansberg (2015), Eckert *et al.* (2019)), and models of technology adoption (Yeaple (2005), Atkeson & Burstein (2010), Bustos (2011b)).

simulated moments of wage distributions to their empirical counterparts.

The resulting parameter estimates imply that more skilled workers earn disproportionately higher wages at larger firms for two reasons. First, their productivity is more complementary to scale. Second, their labor supply curve is less elastic, implying that finding skilled workers is more costly to firms, and thus, these workers earn higher rents from employment relationships. By allowing for firm-varying returns to skill, the model quantitatively matches many dimensions of earnings inequality not typically featured in frameworks of international trade: Wage inequality between- and within-firms, as well as empirical patterns in within-group wage dispersion. Even though labor demand is estimated on firm-level outcomes and not explicitly targeted in the structural simulations, the model matches aggregate patterns of worker-to-firm sorting in the data well.

To assess the model's aggregate implications, I consider counterfactual changes in trade costs implied by the observed changes in the share of exporting firms, export shares, and total import shares between the periods 1993-2002 and 2003-2014. I find that the associated decrease in trade costs increased wage inequality by 4.2 percent, which captures 22 percent of the total increase in wage inequality observed in the data. I find that trade accounts for 30 percent of the observed changes in aggregate skill premia. In part, this reflects that within-firm effects, which account for 30 percent of counterfactual changes in earnings inequality, amplify the change in aggregate skill premia. Between firms, wages of skilled workers employed at exporting firms see the most significant increase. As a consequence, counterfactual changes in inequality are largest within high skill types and occur predominantly at the top of the wage distribution.

This theoretical mechanism also implies new welfare effects of trade. The labor market power of firms differs across skill types, implying distortions in the equilibrium allocation in both labor and product markets and hence influences the gains from trade. To quantify the associated welfare effects, I analytically derive a set of correcting proportional income taxes. The tax rate varies by skill type and the tax revenue funds wage-bill subsidies that induce firms to ignore rent-sharing concerns when making hiring decisions, to produce at a more efficient scale, and to reduce their profit margins. As a result of the tax reform, relatively more high skill workers sort into firms that expand upon trade liberalization. As a consequence of the reallocation in the labor market, the tax decreases wage inequality within skill groups by 20 percent, relative to the effect of trade without intervention. Worker welfare rises for all skill types and on average, by 6 percent. These findings highlight the interaction between trade frictions and domestic distortions in influencing the distributional consequences of trade liberalization and the welfare gains from trade.

Related Literature Traditionally, the literature on trade and inequality emphasizes between-sector reallocation and studies changes in mean wages of broadly defined worker groups. While quantitative frameworks based on this approach (Parro (2013), Caron *et al.*

(2017, 2014), Cravino & Sotelo (2017), Burstein *et al.* (2019), Caliendo *et al.* (2019), Eckert (2019)) have yielded valuable insights, I depart from this literature by focusing on reallocation between firms, and by allowing for imperfect competition in labor markets.

As part of the recent literature on heterogeneous firms and trade (Melitz (2003)), one line of research highlights the importance of trade liberalization in changing the relative demand for skilled workers within and between firms (Yeaple (2005), Verhoogen (2008), Bustos (2011b), Davis & Harrigan (2011), Harrigan & Reshef (2015), Burstein & Vogel (2017), Fieler *et al.* (2018)). I contribute to this literature by providing a model of firm-varying labor demand that explicitly accounts for the role of firm size, and remains tractable even with arbitrarily many worker types. Contrary to these papers, I highlight wage dispersion within skill groups, firm-varying skill wage premia, and imperfect competition in labor markets.

Labor market frictions were introduced into models with firm heterogeneity in the form of efficiency wages (Davis & Harrigan (2011)), fair wages (Egger & Kreickemeier (2009, 2012),Amiti & Davis (2012), Egger *et al.* (2013)) and search frictions (Davidson & Matusz (2006), Felbermayr *et al.* (2013, 2011), Helpman & Itskhoki (2010), Ritter (2015), Helpman *et al.* (2017, 2010)). All of these papers focus on wage dispersion across firms for ex-ante similar workers. In contrast, I provide an integrated framework that also accounts for within-firm wage dispersion and worker-to-firm sorting.

This paper also contributes to the literature studying the gains from trade in heterogeneous firm models (e.g., Arkolakis *et al.* (2012), Arkolakis *et al.* (2019)) by quantifying the welfare effects of distortions in worker-to-firm sorting.[5] While previous research has emphasized this channel theoretically (Davidson *et al.* (2008), Helpman *et al.* (2010)), and via reduced-form evidence (Davidson *et al.* (2012, 2014)), this paper is the first to quantify its relevance.

A mechanism not explored in this paper is competitive assortative matching (Ohnsorge & Trefler (2007), Costinot & Vogel (2010), Sampson (2014)). Recent papers in this literature featuring endogenous firm sizes predict that firms match with a single worker type (e.g., Grossman *et al.* (2017)), and that within-firm wage dispersion is a sign of inefficiencies, i.e., search frictions in the labor market (Eeckhout & Kircher (2018)).[6] Accounting for matching based on unobserved worker ability might constitute an interesting extension of my model.[7]

[5] To the best of my knowledge, this paper is also the first one to quantify the welfare effects of a correcting tax in the context of imperfectly competitive labor markets and international trade.

[6] Eeckhout & Pinheiro (2014) provide a model with many-to-one matching, where, interestingly, departures from a homothetic CES are necessary to generate different skill mixes across firms.

[7] My paper relates more broadly to the literature on labor market sorting (e.g., Shimer (2005), Gautier *et al.* (2010), Gautier & Teulings (2015), Lise *et al.* (2016)). While much of this literature describes worker-to-jobs sorting, I develop a model that determines equilibrium firm size and generates wage inequality, both between firms for a given worker type, and within firms across different worker types.

The non-homothetic CES was first studied by Sato (1975) and Hanoch (1975), and recently re-introduced by Comin *et al.* (2017) to analyze income effects in preferences in a growth model. Matsuyama (2015) uses the same model of utility in the context of a home market effect model of trade. This paper is the first to use it in production to study wage inequality and labor market sorting. In recent work, Bauer *et al.* (2019) analyze non-homotheticity in firms' demand for IT, and Blaum *et al.* (2019) study non-homotheticity in firms' import demand for quality.

This paper also relates to a broad literature arguing that employer heterogeneity is important for aggregate wage inequality. The model of the labor supply side is related to Card *et al.* (2018), Haanwinckel (2021), Lamadon *et al.* (2022) and Berger *et al.* (2021), although I consider a different specification for labor demand and output markets. Since my framework features reduced-form wage expressions featuring log-linear interactions between skill and firm effects, it also relates to a large reduced form literature built around two-way fixed effects models of wages (Abowd *et al.* (1999), Card *et al.* (2013), Alvarez *et al.* (2018), Song *et al.* (2018), Bonhomme *et al.* (2019), Borovickova & Shimer (2019)).

2 Motivating Facts

In this section, I establish three key stylized facts about the relationship between relative wages and relative employment across firms that guide the theoretical model developed below.

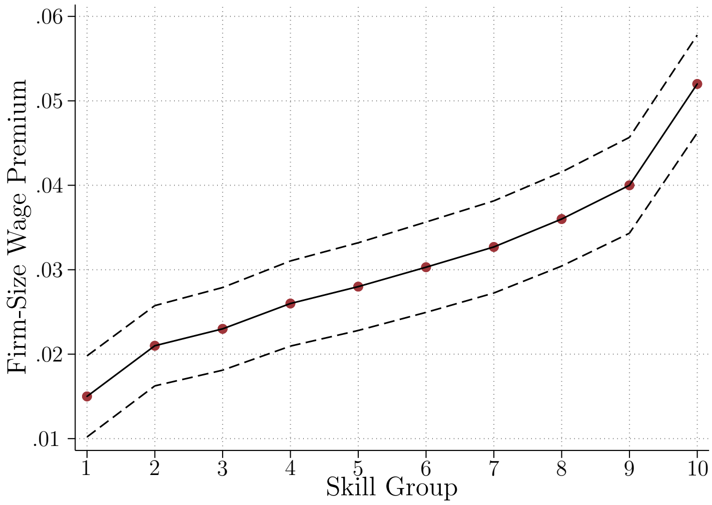
2.1 Data Description and Definition of Skill Types

This study uses the Linked-Employer-Employee Data (LIAB) (longitudinal model 1993-2014 (LIAB LM 9314)) provided by the Institute for Employment Research (IAB) in Germany.[8] The data contains information on the complete workforce of a subset of German establishments. The sample establishments are the ones selected - at least once - in an annually conducted survey between 2000 and 2008. The employee information contains employment biographies from 1993 to 2014 of all individuals, which were, at least one day, employed at one of the sample establishments. I organize the resulting dataset as an annual panel.

The data provides detailed information on individual workers - daily wages, days worked, age, gender, nationality, tenure at the firm, education, and occupation. The information on employers includes revenues, the sector of economic activity, spending

[8] Data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and subsequently remote data access. A detailed documentation of the data can be found in Heining *et al.* (2016),Fischer *et al.* (2009) and Heining *et al.* (2014). Data access was provided under project FDZ1440.

Figure 2.1 The Relationship Between Firm Sales and Worker Wages by Skill Type



Notes: This figure displays the coefficient estimates β_s in the empirical model in (2.1), which measure the relationship between worker wages and the sales of their employer for different skill groups. Standard errors are clustered at the establishment level. All regressions include fixed effects for a worker's age-group, gender, nationality, occupation, sector of occupation, the federal state of residence and years. Firm controls include the revenue share of intermediate inputs, and average skill intensity of all full-time employees.

on intermediate inputs, and exports. Throughout, I use the term firms to refer to establishments. A more detailed overview of the data is presented in Appendix A.

Definition of Skill This paper associates the skill of a worker with the *average* wage that workers of her type earn. A worker's type corresponds to her sector of employment, occupation, and educational attainment. For each sector, I calculate age- and gender-adjusted average wages across all workers of a given type and rank types according to their average wages. Skill group 1 consists of workers who belong to a type that falls into the first decile, and henceforth until skill group 10.

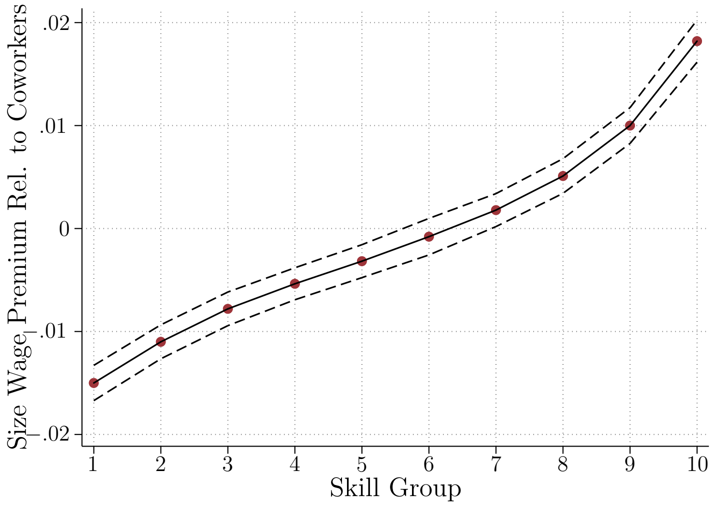
2.2 Facts

Fact 1 *Compared to less skilled workers, the wages of more skilled workers are more strongly related to firm size.*

A canonical feature of labor markets is that larger firms pay higher wages to their employees. Here, I document that firm-size wage premia differ across skill types. I regress the wage of an individual worker i in year t, $\log W_{i,t}$, on a vector of worker controls $\boldsymbol{x}_{i,t}$, a vector of employer controls $\boldsymbol{y}_{f(i,t)}$, fixed effects $\boldsymbol{\omega}(i,t)$ for sector of employment, occupation, education, federal state of employment and year, as well as log revenues, $\log REV_{f(i,t)}$:

$$\log W_{i,t} = \beta_{s(i,t)} \log REV_{f(i,t)} + \boldsymbol{\delta}' \mathbf{x}_{i,t} + \boldsymbol{\gamma}' \boldsymbol{y}_{f(i,t)} + \boldsymbol{\omega}(i,t) + \epsilon_{i,t}, \tag{2.1}$$

Figure 2.2 The Relationship Between Firm Size and a Worker's Relative Wage Within the Firm by Skill Type



Notes: This figure displays the coefficient estimates β_s in the empirical model in (2.2), which measures the relationship between a worker's pay relative to the mean pay at the establishment and the establishment sales by skill group. Standard errors are clustered at the establishment level. All regressions include fixed effects for a worker's age-group, gender, nationality, occupation, sector of occupation, the federal state of residence, years and control for the revenue share of intermediate inputs.

where $s(i,t)$ and $f(i,t)$ denote the skill level and employer of worker i.

Figure 2.1 displays the coefficient estimates for β_s. The results indicate that the earnings of high skill workers are more strongly related to employer revenues. On average, high skill workers are paid five percent more when working for an employer with twice the sales volume. In contrast, among the lowest skill group, a doubling of employer sales is associated with a one percent increase in wages earned.

Table C.1 shows this pattern is robust to using value-added or sales to total worker ratios to measure firm size and performance, as well as to the inclusion of worker and firm fixed effects.

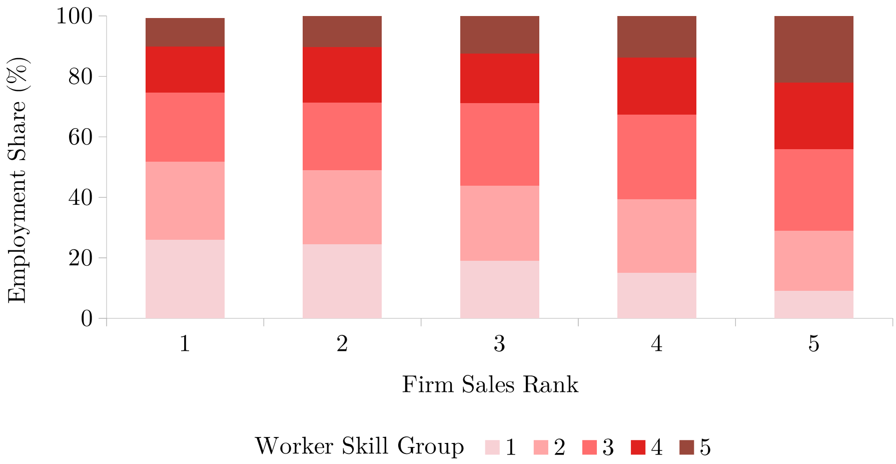
Fact 2 *Within-firm wage inequality between skill groups is systematically related to firm size.*

To show how size relates to the earnings within the firm, I consider the same explanatory variables as before but instead use a worker's wage relative to the average wage $\overline{W}_{f(i,t)}$ paid at her firm as a dependent variable:

$$\log W_{i,t} - \log \overline{W}_{f(i,t)} = \beta_{s(i,t)} \log REV_{f(i,t)} + \boldsymbol{\delta}'\mathbf{x}_{i,t} + \boldsymbol{\gamma}'\boldsymbol{y}_{f(i,t)} + \boldsymbol{\omega}(i,t) + \epsilon_{i,t}. \tag{2.2}$$

Differences in the coefficient β across skill types indicate that within-firm inequality between worker groups is systematically related to firm revenues. Figure 2.2 displays the coefficient estimates and indicates that within firm-pay inequality is increasing in firm sales. High skill workers are paid 2 percent more than their coworkers when working

Figure 2.3 Employment Distribution of Skill Groups Across The Establishment Sales Distribution



Notes: This figure displays the distribution of skills across the firm size distribution. Firm sales ranks denote quintiles of sector-specific sales distributions.

for an employer with double the sales volume. In contrast, low skill workers are paid 1.5 percent less than their coworkers when working for an employer with double the sales volume.

Table C.2 provides detailed regression results and shows that the patterns hold for alternative measures of firm scale and are qualitatively robust to the inclusion of individual firm and worker fixed effects.

Fact 3 *Larger firms employ disproportionately more skilled workers.*

Figure 2.3 compares the skill distribution of workers employed across different parts of the firm sales distribution and shows that more skilled workers tend to work for larger firms.

To supplement the graphical evidence, I regress an index measuring a firm's skill intensity, $comp_{f,t} = \overline{s(i,t)}_{i\in f(t)}$, on firm revenues, a vector of firm controls $\mathbf{y}_{f,t}$, and fixed effects $\boldsymbol{\omega}_{f,t}$ for industry, geographic location and years:

$$comp_{f,t} = \beta \log REV_{f,t} + \gamma' \mathbf{y}_{f,t} + \boldsymbol{\omega}_{f,t} + \epsilon_{f,t}. \tag{2.3}$$

Table C.3 displays the coefficient estimates and indicates that firm size is positively related to skill intensity, suggesting that the skill composition of labor demand is heterogeneous across firms and systematically related to firm size.

3 Framework

This section develops a theoretical framework that rationalizes the reduced form facts in the previous section and highlights a new theoretical mechanism through which international trade affects wage inequality within and between firms.

3.1 Economic Environment

A small open economy H (Germany in the application) trades with an aggregate rest of the world F. Foreign variables, denoted with an asterisk, are exogenous. Workers and firms populate the economy, and all agents have CES preferences over differentiated goods with an elasticity of substitution $\eta > 1$. Product markets are monopolistically competitive.

Workers differ in skill $s \in \mathcal{S} = \{1, ..., S\}$, and there is a continuum of workers of each skill type of measure L_s. Workers derive utility from consumption and job-specific amenities. Section 3.3 describes the role of amenities and the problem of workers in detail.

As in Melitz (2003), there is a large, unbounded pool of prospective entrants into the industry. Entrants pay a fixed cost F_E to enter and gain monopoly rights over a single differentiated variety. Active Firms are indexed by f, and $\mathcal{F}$ denotes the set of active firms. Upon entry, firms draw the demand for their product φ, as well as non-wage amenities A that they can offer to employees, from a joint cumulative distribution function $G(\varphi, A)$. The draw φ determines the demand schedule faced by the firm. Firms incur a fixed cost F_X to export, and exports are subject to an iceberg trade cost τ.

3.2 Technology

Each firm f produces output Q_f by combining labor inputs $l_{s,f}$ according to a non-homothetic CES production function (Hanoch (1975), Sato (1975)):

$$1 = \sum_{s \in \mathcal{S}} \Omega_s^{\frac{1}{\sigma}} l_{s,f}^{\frac{\sigma-1}{\sigma}} Q_f^{\frac{\varepsilon_s - \sigma}{\sigma}}. \tag{3.1}$$

Ω_s denote skill-specific productivity shifters that are common to all firms. σ governs the elasticity of substitution between worker types. The parameter ε_s controls the complementarity of skill type s with firm output Q. Intuitively, as the index Q rises, the productivity of skill type s varies at a rate controlled by parameter ε_s. To see this, consider the relative marginal products of labor MPL_s for two skill types:

$$\log\left(\frac{MPL_{s,f}}{MPL_{s',f}}\right) = \frac{1}{\sigma}\log\left(\frac{\Omega_s}{\Omega_{s'}}\right) - \frac{1}{\sigma}\log\left(\frac{l_{s,f}}{l_{s',f}}\right) + \left(\frac{\varepsilon_s - \varepsilon_{s'}}{\sigma}\right)\log Q_f. \tag{3.2}$$

It is evident that the non-homothetic CES nests both the homothetic CES and Cobb-Douglas functional forms. If $\varepsilon_s = \varepsilon_{s'} = 1$, the production function is equal to the standard CES. Additionally, if $\sigma = 1$, then the production function is Cobb-Douglas. In the general non-homothetic case, the relative productivity of worker types depends explicitly on the size of the firm.

To provide a rationale for this feature, Appendix B.2 provides two microfoundations for the technology. The link between output and relative productivities of workers is implied by any model where, in order to expand production, firms make costly investment decisions in new technologies with differing relative productivities of skill types. In particular, I show that a generalization of the technology-choice framework in Bustos (2011a) provides a microfoundation for the non-homothetic CES functional form.[9] Further, non-homotheticity in firms' labor demands for skilled workers arises in the context of a task-assignment model, where it captures the extent to which scale affects the relative comparative advantages of skill types in performing a given task.[10]

The following parametric restrictions ensure that the production quantity index defined by equation (3.1) is strictly increasing and quasi-concave in all labor inputs. Appendix B.1 provides a detailed overview of the properties of the non-homothetic CES production function.

Assumption 1 *For all* $s \in \mathcal{S}$, $(\sigma - \varepsilon_s)(\sigma - 1) > 0$.

3.3 Labor Supply to Individual Firms

Workers derive utility from consumption and non-wage amenities offered by jobs. They view jobs as imperfect substitutes and have random utility over employment opportunities, which generates upward-sloping labor supply functions to firms because a higher wage induces more workers to accept a job (Card *et al.* (2018)).

The utility of worker i working at job f is given by:

$$U_i(f,s) = \log C_i + \log A_f + \frac{1}{\beta_s}\epsilon_{i,f}, \tag{3.3}$$

where $C_i = \left(\sum_{f\in\mathcal{F}} \left(\varphi_f c_{f,i}\right)^{(\eta-1)/\eta} + \left(c_i^*\right)^{(\eta-1)/\eta}\right)^{\eta/(\eta-1)}$ is a CES consumption index defined over domestic varitieties, c_f, and foreign varieties, c^*. A_f denotes non-wage amenities

[9] Specifically, I demonstrate that the technology-choice model implies the same cost function as the general class of non-homothetic CES functions (Sato (1975)).

[10] Beyond these particular microfoundations, several models in the literature provide qualitatively similar predictions. Models of organizational hierarchy (Lucas (1978),Rosen (1981), Garicano (2000), Garicano & Rossi-Hansberg (2015)), or CEO-compensation (e.g. Gabaix & Landier (2008)) feature environments where the relative productivity of different worker groups varies with the scale of a firm.

that are enjoyed by all employees of the firm. $\epsilon_{i,f}$ is an extreme-value distributed taste shock with variance 1.

The idiosyncratic taste shock gives rise to horizontal employer differentiation as workers differ in their preferences over the same set of job opportunities. Individual workplace preferences $\epsilon_{i,f}$ are private information to workers. Thus, firms cannot price-discriminate and offer wages $W_{s,f}$ conditional on worker type s, but not on realizations of $\epsilon_{i,f}$.

To characterize the labor supply to a firm offering wage $W_{s,f}$, note that (3.3) implies the following indirect utility:[11]

$$U_i(f,s) = \log\left(W_{s,f}/P\right) + \log A_f + \frac{1}{\beta_s}\epsilon_{i,f}, \tag{3.4}$$

where P is the standard CES price index $P = \left(\sum_{f\in\mathcal{F}} \varphi_f^{\eta-1} p_f^{1-\eta} + (P^*)^{1-\eta}\right)^{1/(1-\eta)}$.

Firms post wages, and each worker i chooses a job opportunity f_i^* to maximize her indirect utility in (3.4). The labor supply to an individual firm is proportional to the probability $\lambda_{s,f}$ that workers of type s chooses to work for f. Provided that $\beta_s > 1$, $\lambda_{s,f}$ depends on the wage offered by the job, on wages offered by all other jobs as well as firm-level amenities:[12]

$$\lambda_{s,f} \equiv \mathbb{P}\left(f_{i,s}^* = f \mid \left\{W_{s,f'}, A_{f'}\right\}_{f'\in\mathcal{F}}\right) = \frac{\left(W_{s,f}A_f\right)^{\beta_s}}{\sum_{f'\in\mathcal{F}}\left(W_{s,f'}A_{f'}\right)^{\beta_s}}. \tag{3.5}$$

The labor supply to a firm equals the mass of workers that choose to accept its job offer:

$$\log \mathcal{L}_s(W;A) = \log\left(\lambda_{s,f}L_s\right) = \beta_s \log W + \beta_s \log A + \log \Lambda_s, \tag{3.6}$$

where $\Lambda_s \equiv \left(\sum_{f\in\mathcal{F}}\left(A_f W_{s,f}\right)^{\beta_s}\right)^{-1} L_s$ captures aggregate labor demand.

Firms view themselves as infinitesimal within the market and compete monopsonistically for workers. The elasticity of the labor supply to an individual firm thus equals $\frac{d\log\mathcal{L}_s}{d\log W_s} = \beta_s$. β_s captures the degree of competitiveness of the labor market. If $\beta_s \to \infty$, the labor market for skill type s is perfectly competitive, and firms can hire any number of workers at a fixed wage. Lower values of β_s, in turn, capture less competitive labor markets.[13]

The information asymmetry between workers and firms gives rise to rent-sharing between

[11] See Section B.3.1 for the derivation.

[12] See Appendix B.3.2 for the derivation.

[13] Manning (2003) notes that upward-sloping labor supply curves to the firm arise within models of directed search or convex vacancy posting cost. Appendix B.4.1 illustrates that upward-sloping labor supply curves can be microfounded through a model with search frictions (e.g. Burdett & Mortensen (1998)). This illustrates that β_s is conceptually related to the job-finding to job-destruction ratio in search models of the labor market. Further, Appendix B.4.3 shows that a model with random matching in labor markets, convex vacancy posting cost, and multilateral bargaining between workers and employers (Stole & Zwiebel (1996)) gives rise to similar implications for wages.

workers and firms, and rents $R_{s,f}$ measure workers' average willingness to pay to remain employed at job f (Rosen (1987)):

$$R_{s,f} \equiv \mathbb{E}\left\{R : U_i(f,s) - \log R/P = \max_{f' \in \mathcal{F} \backslash f} U_i(f',s) \,|i{:}f^*_{i,s} = f\right\}, \tag{3.7}$$

The following proposition illustrates that worker rents are inversely related to the elasticity of labor supply.

Proposition 1 *The average rents earned by employees of type s at firm f equal* $R_{s,f} = (1 - \frac{\beta_s}{1+\beta_s})W_{s,f} = \frac{1}{1+\beta_s} \cdot W_{s,f}$.

For intuition, consider a firm wanting to expand production. The firm cannot observe the idiosyncratic workplace preferences of its current employees and, therefore, has to raise wages for all workers to increase hiring. An increase in wages raises the value of the job for workers that already enjoyed working at the firm. If labor supply is less elastic, then to increase production, firms have to raise wages, and hence the value of jobs to workers, by more.

3.4 The Relationship Between Firm Scale, Wages and Employment Composition

3.4.1 Cost-Minimization

To derive the relationship between a firm's scale, its wages, and workforce composition, I solve for the (variable) cost function $C\left(Q_f, A_f\right)$, which minimizes labor cost at any given target level of output Q :

$$C\left(Q_f, A_f\right) \equiv \min_{\{l_{s,f}\}} \; \sum_s W\left(l_{s,f}, A_f\right) l_{s,f} \quad \text{s.t. } 1 = \sum_{s \in \mathcal{S}} \Omega_s^{\frac{1}{\sigma}} l_{s,f}^{\frac{\sigma-1}{\sigma}} Q_f^{\frac{\varepsilon_S - \sigma}{\sigma}} \;, \tag{3.8}$$

where $W(l, A)$ denotes the inverse of the labor supply curve in equation (3.6). Under Assumption 1, this problem has a well-defined interior solution.

Cost-minimization implies that wages equal a mark-up over the marginal product of labor, MPL_s, and marginal cost:

$$\log\left(W_{s,f}\right) = \log\left(\frac{\beta_s}{\beta_s + 1}\right) + \log\left(MC\left(Q_f, A_f\right) MPL_s\left(Q_f\right)\right). \tag{3.9}$$

In equilibrium, a monopolist maximizes profits by equating marginal cost and marginal revenues. Hence, wages correspond to a mark-down over the marginal revenue product of labor, as in classical Monopsony Theory (e.g., Robinson (1933)). The mark-down $\beta_s / (\beta_s + 1)$ is decreasing in the elasticity of the labor supply faced by the firm. Further, Proposition 1 and the solution for wages in (3.9) jointly imply that skill types who are in less elastic labor supply earn higher rents, relative to their marginal product of labor.[14]

3.4.2 The Determinants of Wage Outcomes

A key outcome of the analysis is how firm size determines the wages paid to different skill types. The following proposition illustrates the channels that determine this relationship.

PROPOSITION 2 *Denote logged variables by lower-case letters.*

1. *Wages are given by:*

$$w_{s,f} = \chi_s + \frac{\varepsilon_s}{\sigma + \beta_s} \times q_f + \frac{\sigma}{\sigma + \beta_s} \times \psi_f - \frac{\beta_s}{\sigma + \beta_s} \times a_f,$$

where $\chi_s \equiv \frac{1}{\sigma+\beta_s} \log\left(\Omega_s / \left(L_s \Lambda_s\right)\right)$ *and* $\psi_f \equiv \log\left(MC_f / \left(\sum_s \Omega_s^{1/\sigma} l_{s,f}^{(\sigma-1)/\sigma} Q_f^{(\varepsilon_s-\sigma)/\sigma}\right)^{\frac{\varepsilon_s - \sigma}{1-\sigma}}\right)$.

2. *If* $\beta_s = \beta$, *and* $\varepsilon_s = \varepsilon$, *then wages are given by:*

$$w_{s,f} = \chi_s + \frac{\sigma}{\sigma + \beta} \times \widetilde{\psi}_f - \frac{\beta}{\sigma + \beta} \times a_f,$$

where $\widetilde{\psi}_f = \log\left(MC_f \cdot Q / \left(\sum_s \Omega_s^{1/\sigma} l_{s,f}^{(\sigma-1)/\sigma}\right)^{\frac{\varepsilon - \sigma}{1-\sigma}}\right)$.

3. *If* $\beta_s \to \infty$, *then wages are given by:*

$$w_{s,f} = \chi_s - a_f.$$

Proof. The result follows from inserting equation (3.6) into equation (3.9). ∎

Statement 1 shows that two intuitive channels determine the effect of firm size on relative wage outcomes within skill groups and firms. First, firm size directly impacts relative productivity and, therefore, relative wage outcomes of skill groups. Workers whose relative productivity is increasing in firm size receive larger firm-size wage premia and receive relatively earnings than their coworkers in larger firms. Second, Proposition 2 shows that larger firms pay disproportionately higher wages to workers that earn higher rents. Skill groups that are in less elastic labor supply receive higher firm-size wage premia and receive relatively higher wages than their coworkers when working for larger firms.

Proposition 2 further highlights that the model accommodates the standard model with a firm-size wage premium common to all skill types and homogeneous skill wage premia across firms as a special case (statement 2). If the labor supply curve for each skill type has the same elasticity, and if skill types are equally complementary to output, then

[14] Denoting marginal revenue products by $MRPL_s$, worker rents equal: $R_{s,f} = \frac{\beta_s}{(1+\beta_s)^2} \cdot MRPL_s\left(Q_f, A_f\right)$.

firm wage effects are common to all skill groups, and relative wages between skill types are equal across firms.[15] If labor markets are perfectly competitive (statement 3), then wage differentials between firms are only driven by compensating differentials and in particular, invariant to firm size.

3.4.3 Firm Size and the Skill Composition of the Workforce

In a perfectly competitive labor market, firms can hire any number of workers at a given wage. In this case, differences in workforce skill composition indicate relative productivity differences in worker skill types across firms. In the presence of upward-sloping labor supply curves, however, workers command rents, and hiring decisions by firms are subject to two concerns. First, firms wish to hire the most productive workers, given their scale. Second, firms prefer to hire workers that demand lower rents. The relationship between skill composition and output depends on how firm size changes the relative productivity of workers, as well as on the rents that skill types command.

To illustrate the productivity channel, Figure 3.1a assumes that skill types only differ in their complementarity to output. In this case, the *relative* employment for different skill types is not affected by rent sharing concerns. Due to their relative technological advantage, the relative labor demand for skill types more complementary to scale increases in firm size. Relative wages paid by firms reflect differences in relative worker productivity only and, therefore, also increase in firm size.

Figure 3.1b focuses on the alternative case when the relationship between skill composition and firm size is driven only by the rent-sharing channel. Workers that command higher rents become increasingly expensive to larger firms, which reduces their employment shares.

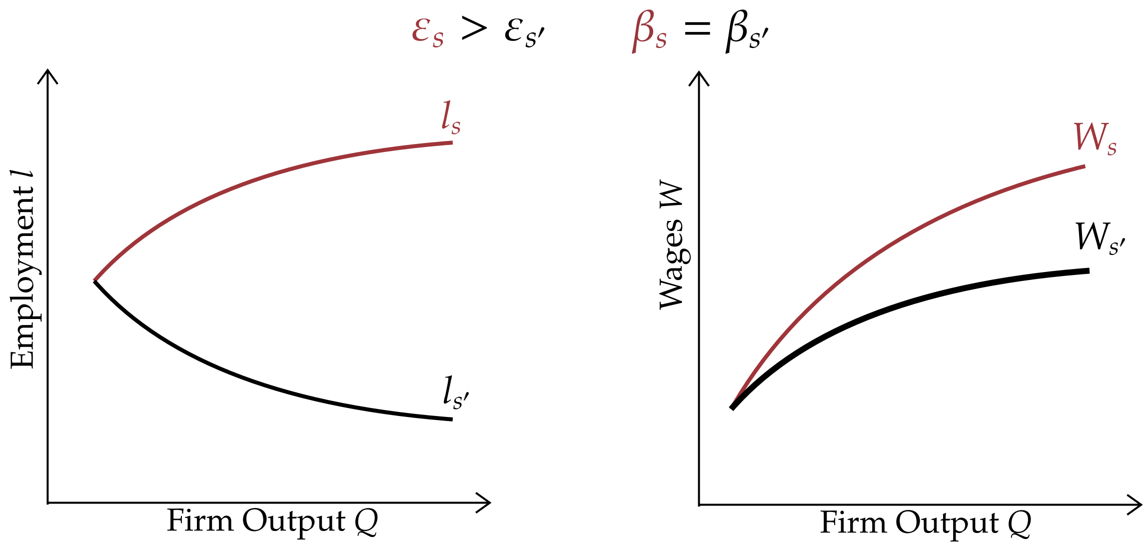
3.5 Closing the Model

Thus far, the analysis took a firm's output as given. This subsection closes the model by characterizing firms' profit-maximizing choices of output, as well as the equilibrium conditions for market-clearing.

[15] The model does not feature wage effects on the individual worker level and, therefore, does not directly relate to common reduced-form empirical models that are log-additive in worker and firm fixed effects (e.g. Abowd *et al.* (1999)). In Appendix B.4.2, I consider an extension of the model that introduces heterogeneous efficiency units for a given type of worker. Armed with this simple extension, the model rationalizes a log-additive wage model with *individual* and firm effects under the parametric restrictions provided in statement 3. of Proposition 2.

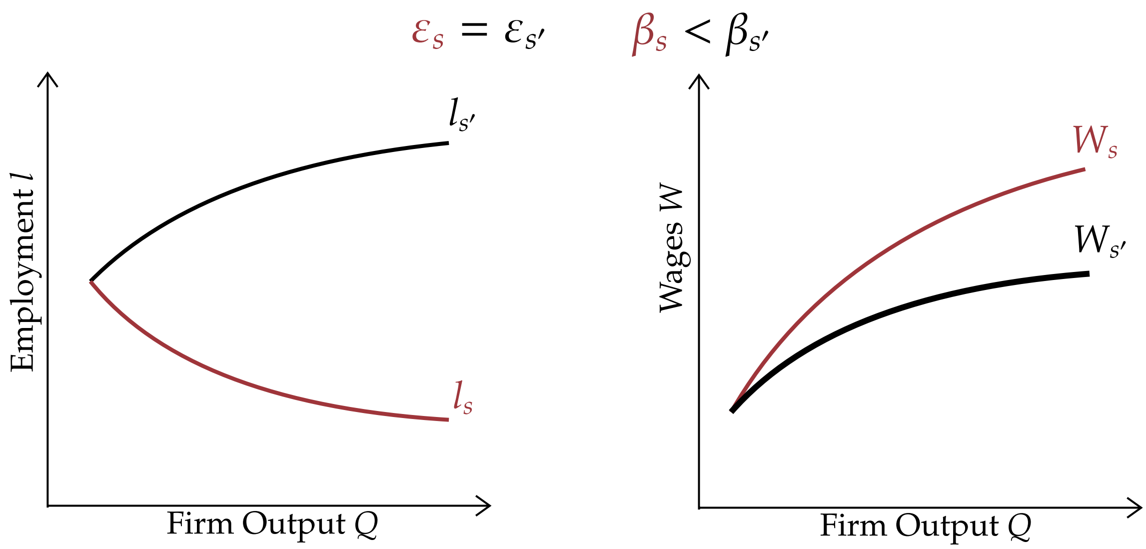
Figure 3.1 Workforce Composition and Firm Size

(a) The Role of Technology



Notes: This figure illustrates the model-implied relationship between employment, wages, and firm size when labor supply elasticities are equal across skill types. Worker types that are more complementary with firm scale are hired relatively more by larger firms and receive relatively higher wages at larger firms.

(b) The Role of Labor Supply



Notes: This figure illustrates the model-implied relationship between employment, wages, and firm size when labor supply elasticities differ across skill types. Workers that command larger rents (lower labor supply elasticities) are hired relatively less by larger firms, but receive higher relative wages at larger firms.

3.5.1 Profit-Maximization

Domestic and foreign demands, $Q_H(p;\varphi)$ and $Q_F(p;\varphi)$, at price p, are given by:

$$Q_H(p;\varphi) = \varphi^{\eta-1}p^{-\eta}P^{\eta-1}Y, \qquad Q_F(p;\varphi) = \varphi^{\eta-1}p^{-\eta}Y^{*}, \tag{3.10}$$

where Y denotes domestic consumption expenditures, P is the CES price index, and Y^{*} is the foreign demand shifter.

Firms choose whether to enter foreign export markets and how much to produce in order to maximize operating profits π:

$$\begin{aligned} \pi\left(\varphi_f, A_f\right) &= \max_{p_H, p_F, 1_X}\left\{\{Q_H(p_H;\varphi)\} + 1_x Q_F(p_F;\varphi) - C\left(Q_f, A_f\right) - 1_X F_X\right\} \\ &\qquad \text{s.t.} \quad Q_f = Q_H(p_H;\varphi) + \tau Q_F(p_F;\varphi) \end{aligned} \tag{3.11}$$

where p_H and p_F denote prices charged in the domestic and foreign markets. $C\left(Q_f, A_f\right)$ denotes the total operating cost that solves the cost-minimization problem described in equation (3.8).

Optimal pricing implies that firms charge a constant mark-up over their marginal cost. In the domestic market, a firm therefore charges:

$$p_{H,f} = \frac{\eta}{\eta-1}MC\left(Q_f, A_f\right). \tag{3.12}$$

The price of a firm exporting to market F is given by:

$$p_{F,f} = \tau\frac{\eta}{\eta-1}MC\left(Q_f, A_f\right) \tag{3.13}$$

Marginal costs are non-constant and depend on the overall quantity produced. Hence, the decision to export affects the cost of production of the goods sold in the domestic market. In contrast to the standard model, an exporter chooses its overall output to maximize profits in all markets jointly.[16] Appendix B.3.3 shows that marginal costs depend directly on output and are equal to a wedge $\mathcal{W}_f$ over average cost:

$$MC\left(Q_f, A_f\right) = \frac{C\left(Q_f, A_f\right)}{Q_f}\left\{\sum_s \omega(s,f)\frac{\beta_s+1}{\beta_s}\frac{\varepsilon_s-\sigma}{1-\sigma}\right\} \equiv \frac{C\left(Q_f, A_f\right)}{Q_f}\mathcal{W}_f, \tag{3.14}$$

where $\omega(s,f) = W_{s,f}l_{s,f}/C\left(Q_f, A_f\right)$ denotes the wage bill share of skill group s. Appendix

[16] This is the case in any model featuring non-constant marginal costs of production (e.g. Caliendo & Rossi-Hansberg (2012), Almunia *et al.* (2018)).

B.3.4 derives sufficient parametric restrictions that ensure convexity of costs in output. Operating profits $\pi\left(\varphi_f, A_f\right)$ are increasing in the demand shifter φ_f and increasing in firm amenities A_f. Exporting increases operating profits, and as a result, there is selection into exporting. The following proposition summarizes these properties.

Proposition 3 *Operating profits $\pi(\varphi, A)$ are (i) strictly concave in output Q, (ii) increasing in a firm's demand φ, and (iii) increasing in amenities A. Given expenditures and aggregate labor demands for all skill types, the partial equilibrium in product markets is unique.*

Proof. See Appendix B.3.5. ■

3.5.2 Domestic expenditures

Fixed costs use a separate factor of production that is in perfectly elastic supply at a price that is normalized 1. Domestic expenditures Y equal the aggregate income from wages, profits, and fixed cost payments:[17]

$$Y = \sum_{s \in \mathcal{S}} \left[\int_{f \in \mathcal{F}} W_{s,f} l_{s,f} df \right] + \mathbf{F}_X + \int_{f \in \mathcal{F}} \pi\left(\varphi_f, A_F\right) df, \tag{3.15}$$

where $\mathbf{F}_X = \int_{f \in \mathcal{F}} 1_{X,f} F_X df$ denotes aggregate spending on fixed cost.[18]

3.5.3 Free Entry

Free entry implies that equilibrium expected profits for entrants equal the fixed cost of entry F_E. The free entry condition determining the equilibrium number of active firms is given by:

$$F_E = \mathbb{E}_G \left[\pi(\varphi, A)\right]. \tag{3.16}$$

3.6 Pass-Through of Foreign Demand Shocks into Wages

Not all firms export, and therefore an aggregate change in foreign demand constitutes a heterogeneous demand shock across firms. To accommodate an idiosyncratic demand

[17] I assume that the aggregate factor used for the production of fixed cost is owned by an anonymous agent that has the same CES preferences. An alternative assumption that would leave the distributional implications of the model for wages unchanged is to distribute ownership of this factor equally among skill types according to their relative measures (e.g., as in Fieler *et al.* (2018)). As the counterfactual trade-liberalization considered later leads to an increase in entry and exports, this alternative assumption implies larger, yet proportional, gains from trade for skill groups than the ones presented in the main text.

[18] Labor market clearing, $L_s = \int_{\mathcal{F}} l_{s,f} df$, is ensured by the fact that firms' optimal choices of wages need to be consistent with labor supply.

shock, a firm increases employment and therefore raises wages for all its employees. However, the rate at which a demand shock affects wages differs across worker types. In general, a change in trade costs has different effects on within-firm inequality for employers that expand, and for employers that contract, upon trade liberalization.

The following proposition considers the effect of a change in the variable trade cost τ to characterize the forces that shape the response of within-firm inequality to a change in a firm's output demand.

Proposition 4 *Consider a reduction in variable trade cost, $d \log \tau < 0$.*

1. *If $\varepsilon_s = \varepsilon$ and $\beta_s = \beta$, then within-firm inequality changes proportionally across all firms:*
2. *If $\beta_s = \beta$ and $\varepsilon_s > \varepsilon_{s'}$, then $\frac{d \log W_{s,f}}{d \log \tau} - \frac{d \log W_{s',f}}{d \log \tau}$ is increasing in $\frac{d \log Q_f}{d \log \tau}$.*
3. *If $\varepsilon_s = \varepsilon$ and $\beta_s < \beta_{s'}$ then $\frac{d \log W_{s,f}}{d \log \tau} - \frac{d \log W_{s',f}}{d \log \tau}$ is increasing in $\frac{d \log Q_f}{d \log \tau}$.*

Proof. See Appendix B.3.6. ■

The first part of Proposition 4 implies that idiosyncratic shocks to firm product demand have no direct effect on within-firm inequality between skill groups in the particular case where labor demand is homothetic, and labor supply functions are equally elastic for all skill types.

Statements 2. and 3. show that, in general, skill types whose relative productivity increases in output, or whose labor supply is less elastic receive relatively higher wages at firms that expand after trade liberalization.[19] As a result, a change in trade costs has heterogeneous effects on within-firm wage distributions across firms that trade, and those that do not.

3.7 Misallocation and Policy

Differences in scale create relative productivity differences for worker skill types across employers. In a competitive labor market, differences in skill composition across firms would only reflect differences in productivity. However, if labor supply curves are upward-sloping, and if their elasticities differ across types, the relative costs posed by rent sharing factor into firms' hiring decisions. Instead of hiring the most productive workers, given their scale, firms have an additional incentive to hire workers that demand lower rents. Therefore, hiring at the firm-level is distorted, relative to a competitive allocation, and there is misallocation in labor markets.

[19] Through the lens of my model, heterogeneous pass-throughs of firm-specific demand and productivity shocks into wages across worker types, as documented by Cho & Krueger (2019) and Chan *et al.* (2019), thus can be rationalized through both labor demand or labor supply related channels.

Simultaneously, upward-sloping labor supply curves raise marginal costs, consumer prices and therefore reduce the equilibrium scale at which firms operate:

$$MC\left(Q_f, A_f\right) = \frac{C\left(Q_f, A_f\right)}{Q_f}\left\{\sum_s \omega\left(s, f\right) \frac{\beta_s + 1}{\beta_s}\frac{\varepsilon_s - \sigma}{1 - \sigma}\right\} \equiv \frac{C\left(Q_f, A_f\right)}{Q_f}\mathcal{W}_f.$$

Upward-sloping labor supply curves increase the wedge between average and marginal cost, which, relative to an economy with competitive labor markets, reduces the average firm size, and increases the average share of profits in national income.[20] If labor supply elasticities differ across skill types, monopsonistic competition in labor markets also affects the relative sizes of firms and, therefore, the shape of the firm size distribution.

Therefore, the model leaves room for policy to improve upon the allocations in product and labor markets. The following proposition shows that employer wage subsidies, funded by proportional income taxes, restore the competitive allocation.

Proposition 5 *An employer payroll subsidy $\tau_s = \frac{1}{1+\beta_s}$, funded by proportional wage income taxes, restores the competitive allocation in labor markets.*

Proof. See Appendix B.3.7. ■

An employer wage subsidy incentivizes employers to hire more workers and undoes the effect of rent-sharing concerns on relative hiring decisions between worker groups. To fund the subsidy, the government levies higher proportional taxes on skill groups that command larger rents. While the subsidy is not firm-specific by design, since firms differ in workforce composition, and allocate different wage bill shares to different workers, the subsidy is de-facto firm-specific.[21]

Assessing the net welfare consequences of such a tax reform from the perspective of workers requires weighing the potentially adverse welfare effects from labor income taxes against the welfare gains from improved allocations in product and labor markets.

4 Estimation

In this section, I discuss my approach to estimating the relevant structural parameters of the model. I do so in two steps. First, I develop and implement a new strategy to identify

[20] This effect is similar to the distortion caused by imperfect competition in product markets (e.g., Arkolakis *et al.* (2019)). Indeed, in both cases, firms operate at a higher profit to revenue ratio than in the respective perfectly competitive benchmark.

[21] The logic is reminiscent of the literature in economic geography that argues that place-based hiring subsidies can help alleviate the adverse effects of hiring cost on unemployment (e.g., Kline & Moretti (2013)).

non-homothetic CES production functions. I then calibrate the remaining parameters to match empirical moments in German manufacturing in the period 1993 to 2002.

4.1 Production Function Estimation

4.1.1 Challenges to Identification

The identification of the non-homothetic CES production function is subject to two challenges. A common approach to estimating a CES production function (e.g., Goldin & Katz (1996), Oberfield & Raval (2014)) is to estimate relative factor demand directly. Changes in relative labor demand between two skill types s and s' can be written as:

$$d\log\left(l_{s,f}/l_{s',f}\right) = -\sigma d\log\left(W_{s,f}/W_{s',f}\right) + (\varepsilon_s - \varepsilon_{s'})\, d\log Q_f + \gamma_{s,s'f}, \tag{4.1}$$

where $\gamma_{s,s',f} \equiv d\log\left(\Omega_{s,f}/\Omega_{s',f} \cdot \beta_{s,f}/\left(\beta_{s,f}+1\right)\cdot\left(\beta_{s',f}/\beta_{s',f}+1\right)\right)$ is an unobserved error. To illustrate potential sources of bias, I allow the elasticities of labor supply β_s and skill-specific technology shifters Ω_s to be firm-specific.[22]

The identification of the parameters ε_s and σ is subject to two independent challenges. First, output Q_f is endogenous to changes in unobserved changes in labor demand $\gamma_{s,s'f}$, irrespective of the assumptions placed on labor markets.[23] Second, upward-sloping labor supply curves introduce simultaneity bias, and shifts in relative labor demand need to be separately identified from shifts in relative labor supply. Thus, both simultaneity, as well as potentially unobserved changes in labor demand, pose a threat to identification.

To deal with these challenges, I propose an estimation strategy that combines one instrumental variable for partial identification with variation across labor markets for full identification of the structural parameters of interest. I proceed by illustrating the idea of partial identification by way of a graphical argument, followed by a formal derivation.

4.1.2 Identification Strategy

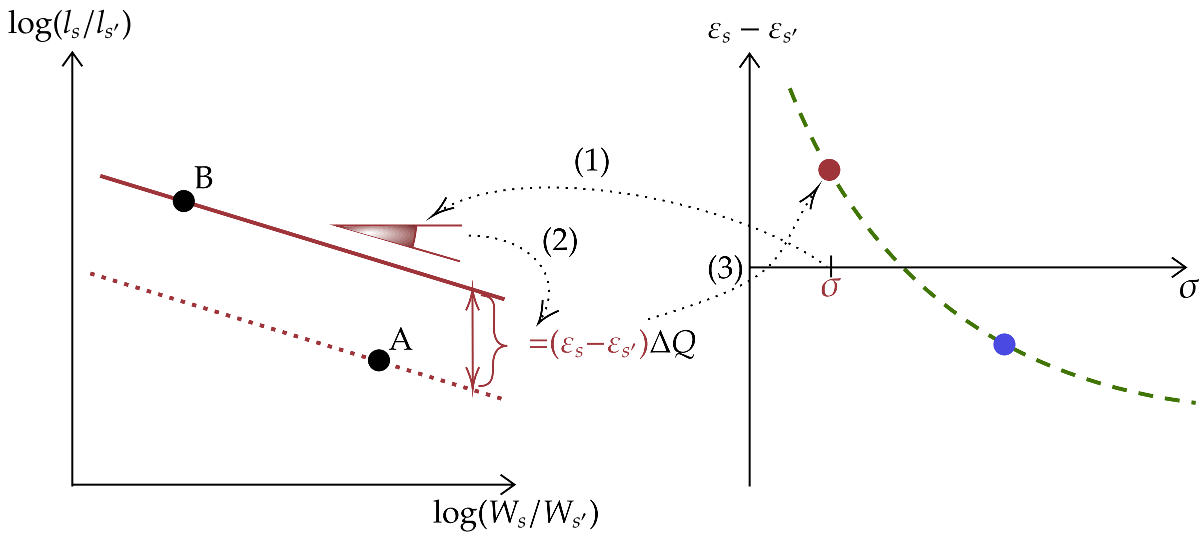
Informal Graphical Argument The central insight of the argument is that changes in wages, employment, and output in response to a shock that is exogenous to unobserved changes in labor demand can provide partial identification of the elasticities of labor demand. Figure 4.1 illustrates the underlying logic of the argument.

The left graph in Figure 4.1 shows changes in wages and employment from point A to point B. Assuming a value for the elasticity of substitution σ, highlighted in Figure 4.1,

[22] In addition, this illustrates that the approach to estimating the production function does not depend on the particular microfoundation adopted to generate upward-sloping labor supply functions, so long as the elasticity of labor supply faced by an individual firm is exogenous.

[23] This threat arises for any production function that is not separable in all inputs.

FIGURE 4.1 IDENTIFICATION THROUGH EXOGENOUS CHANGES IN WAGES, EMPLOYMENT AND OUTPUT



*Notes:*This figure illustrates the logic of the identification argument. The left figure considers a change in observed labor demand and wages from point A to point B. The right figure illustrates the space spanned by the elasticity of substitution σ and relative output scale elasticities $\varepsilon_s - \varepsilon_{s'}$.

two lines with slope $-\sigma$ through points A and B recover the old and new labor demand. This, in turn, allows computing by how much labor demand has shifted vertically. Given that the initial shift from A to B is assumed to be independent of unobservable shocks to labor demand, equation (4.1) implies that the vertical shift equals the non-homothetic effects of output on labor demand, $(\varepsilon_s - \varepsilon_{s'})\Delta Q$. Since one observes changes in output, and given an assumed elasticity of substitution σ, one can compute a unique $(\varepsilon_s - \varepsilon_{s'})$ that rationalizes the observed changes.

Repeating this procedure over alternative guesses for the elasticity of substitution, one can compute a set of partially identified parameters, which is illustrated as a hyperbola in Figure 4.1. If one can observe another change similar to the one from A to B depicted in Figure 4.1 (for example, in another regional labor market or another period in time), one can construct another set of partially identified parameters. If these hyperbolae, in turn, do neither lie on top of each other nor intersect - and production technologies are invariant across regions/time - their intersection recovers both σ and $(\varepsilon_s - \varepsilon_{s'})$.

Formal Argument The formal argument follows the same logic as the graphical one. First, I use changes in wages, employment, and output in response to an identifying shock for partial identification of the structural parameters. I then formalize the conditions, under which cross-regional variation provides full identification of the labor demand elasticities.

Denote $l_{s,s',f,t} \equiv \log l_{s,f,t}/l_{s',f,t}$ and $w_{s,s',f,t} = \log\left(W_{s,f,t}/W_{s',f,t}\right)$ and $q_{f,t} = \log Q_{f,t}$. Letting $dx_t = x_{t+1} - x_t$, changes in labor demand between skill types s and s' between periods t

and $t+1$ are given by:

$$dl_{s,s',f,t} = -\sigma d\left(w_{s,s',f,t}\right) + \left(\varepsilon_s - \varepsilon_{s'}\right) dq_{f,t} + \gamma_{s,s',f,t}, \tag{4.2}$$

where $\gamma_{s,s',f,t}$ is the unobserved structural shock to labor demand. The covariance between observed changes in relative labor demand and an empirical shock $\xi_{f,t}$ - that is assumed to be orthogonal to changes in the error $\omega_{s,s',f,t}$ - across firms f equals:

$$\text{Cov}\left(dl_{s,s',f,t}, d\xi_{f,t}\right) = -\sigma\text{Cov}\left(dw_{s,s',f,t}, d\xi_{f,t}\right) + \left(\varepsilon_s - \varepsilon_{s'}\right)\text{Cov}\left(dq_{f,t}, d\xi_{f,t}\right) + \text{Cov}\left(d\gamma_{s,s',f,t}, d\xi_{f,t}\right). \tag{4.3}$$

Provided that $\xi_{f,t}$ is correlated with changes in relative employment levels, one can derive the following restriction on the technological parameters:

$$1 = \phi\frac{\text{Cov}\left(dw_{s,s',f,t}, d\xi_{f,t}\right)}{\text{Cov}\left(dl_{s,s',f,t}, d\xi_{f,t}\right)} + \theta_s\frac{\text{Cov}\left(dq_{f,t}, d\log\xi_{f,t}\right)}{\text{Cov}\left(dl_{s,s',f,t}, d\log\xi_{f,t}\right)} + \nu_{s,s',t}, \tag{4.4}$$

where $\phi \equiv -\sigma$, $\theta_s \equiv \varepsilon_s - \varepsilon_{s'}$ and $\nu_{s,s',t}$ is an error given by:

$$\nu_{s,s',t} \equiv \frac{\text{Cov}\left(d\gamma_{s,s',f,t}, d\xi_{f,t}\right)}{\text{Cov}\left(dl_{s,s',f,t}, d\xi_{f,t}\right)}.$$

Stacked for all skill groups, equation (4.4) defines a set of moment conditions that provide partial identification of the structural parameters:

$$\mathcal{G}\left(\Theta\right) = \mathbf{0}, \tag{4.5}$$

where $\mathcal{G}\left(\Theta\right) \equiv \left(\phi\frac{\text{Cov}\left(dw_{s,s',f,t}, d\xi_{f,t}\right)}{\text{Cov}\left(dl_{s,s',f,t}, d\xi_{f,t}\right)} + \theta_s\frac{\text{Cov}\left(dq_{f,t}, d\xi_{f,t}\right)}{\text{Cov}\left(dl_{s,s',f,t}, d\xi_{f,t}\right)} + \nu_{s,s',t} - 1\right)_{s\in\mathcal{S}\backslash s'}$.

To see why this provides partial identification, first, note that the covariance ratios are computable by way of two-stage regressions. $\text{Cov}\left(dq_{f,t}, d\xi_{f,t}\right)/\text{Cov}\left(dl_{s,s',f,t}, d\xi_{f,t}\right)$, for example, is the coefficient estimate of a two-stage regression that first regresses changes in employment on changes in $\xi_{f,t}$, and then regresses the fitted values for employment on changes in output. Given an empirical shock $\xi_{f,t}$, these quantities can, therefore, be computed using observed changes in wages, employment, and output of firms. As changes in $\xi_{f,t}$ are assumed to be uncorrelated with the unobserved error, the term $\nu_{s,s',t}$ is equal to zero.[24] Therefore, equation (4.4) provides partial identification of the parameters.[25]

Given observed changes in separate regional labor markets r, one can compute sets of analogous moment conditions $\mathcal{G}_r\left(\Theta\right)$ for each labor market r. The estimator of the technological parameters is given by:

[24] Formally, the probability limit taken with respect to firms f equals zero: $p\lim_{f\to\infty}\nu_{s,s',t} = 0$.

[25] Note that these steps are conceptually equivalent to the logic underlying Figure 4.1.

$$\Theta^{*} = \arg\min_{\Theta} \mathcal{G}(\Theta)' W \mathcal{G}(\Theta), \tag{4.6}$$

where $G(\Theta) \equiv [\mathcal{G}_r(\Theta)]_r$ collects moment conditions across regions, and W is a weighting matrix.

Θ^* identifies Θ under three conditions. First, changes in $\xi_{f,t}$ need to be correlated with firm-level wages, employment, and output (relevance). Second, changes in $\xi_{f,t}$ need to be uncorrelated with changes in unobserved sources of heterogeneity in labor demand across firms (orthogonality). These conditions ensure that for each region, $\mathcal{G}_r(\Theta)$ provides partial identification of Θ.[26] Further, full identification requires that the hyperbolae computed across different regions do not lie on top of each other. A sufficient requirement for full identification is that for worker groups s and s' and regions r and r' :

$$\frac{\mathrm{Cov}_r\left(dl_{s,s',f,t}, d\xi_{f,t}\right)}{\mathrm{Cov}_r\left(dq_{f,t}, d\xi_{f,t}\right)} \neq \frac{\mathrm{Cov}_{r'}\left(dl_{s,s',f,t}, d\xi_{f,t}\right)}{\mathrm{Cov}_{r'}\left(dq_{f,t}, d\xi_{f,t}\right)}. \tag{4.7}$$

The condition stated in equation (4.7) is reminiscent of heteroskedasticity based conditions for partial identification in Leamer (1981) and Feenstra (1994). Appendix D.1.2 provides a more detailed discussion of this identifying assumption and argues that it holds in the empirical application if for example the share of exporting firms differs across regions.

The estimator defined in equation (4.6) fits in the class of Classical Minimum Distance estimators. Appendix D derives its asymptotic distribution that is later used to construct standard errors.

4.1.3 Implementation

To obtain identifying shocks that affect firm-level wages, output and employment, and are not systematically correlated with unobserved shocks to labor demand, I construct Bartik-type (Bartik (1991)) firm-level shocks to output demand that are a variant of firm-level export shocks constructed in earlier papers (Hummels *et al.* (2014), Berman *et al.* (2015), Caliendo *et al.* (2017), Garin & Silvero (2018)).

The Bartik instrument generates exogenous variation in product demand between otherwise similar firms that is based on *where* rather than *which* products a firm is selling. To define the instrument, let $EXP_{s,t}^{GER\to c}$ denote the aggregate exports of the German manufacturing sector s to country c. For each firm observed in the data, let t_0 denote the year where the firm was first observed exporting, if at all. Denoting the sales share of firm f

[26] For example, idiosyncratic shocks to firms' product demands satisfy these conditions. However, these shocks would not satisfy this condition if the elasticity of the labor supply faced by individual firms were endogenous to, say, the size of the firm. Appendix D.1.2 gives more details on how to augment the approach to partial identification in this case.

in region r at time t_0 by ω_{f,r,t_0}, the firm-level demand shock for time $t > t_0$ is defined as:

$$D_{f,t} = \sum_r \omega_{f,r,t_0} EXP_{s,t}^{GER \to c}. \tag{4.8}$$

As pointed out in Garin & Silvero (2018), even though the shock $D_{f,t}$ is firm-level in construction, it may nonetheless contain demand variation that also affects other firms that sell similar products to firm f. However, in this case, so long as $D_{f,t}$ does vary across firms and is exogenous to unobserved changes in labor demand, it identifies the parameters of interest.

Identification requires that $D_{f,t}$ is not systematically correlated with unobserved firm-level changes that affect labor demand between worker types - such as factor-biased technological change. Recent literature points out that exogeneity in Bartik instruments can stem from two sources. Borusyak *et al.* (2022) or Adão *et al.* (2019) argue that variation in changes in export demand across sufficiently many regions alone can be sufficient to generate asymptotically consistent estimates. Goldsmith-Pinkham *et al.* (2020), on the other hand, point out that exogeneity in the "shift shares" alone can also sufficient to yield consistent parameter estimates.

I observe firm-level export shares of firms across most three regions.[27] Therefore, I augment the standard Bartik instrument with an estimated exogenous probability of exporting. I fit a probit model that projects a firm's current exporter status on its lagged exporter status, its past position in the domestic sales distribution as well as a dummy for whether it is located in a formerly East German state. Denoting $p_{f,t}^X$ the resulting estimated probability that firm f exports in year t, the empirical demand shifter $\xi_{f,t}$ is defined as:

$$\xi_{f,t} = p_{f,t}^X D_{f,t}. \tag{4.9}$$

I implement the estimation strategy for three skill groups. I refer to these groups as high skill, medium skill, and low skill, respectively, for the remainder of the paper. High skill workers are those in jobs that pay on average more than the 70th percentile of jobs in their sector, while the jobs of low skill workers fall into the bottom three percentiles. Firm size Q is measured by value-added. To estimate a medium-term production function, I use changes across all available lags.

4.1.4 Results

Individual moments used in the minimum distance estimator are shown in Table C.4. The coefficient estimates and standard errors are displayed in Table 4.1. I estimate that worker

[27] The firm survey, in most years, provides only export shares to the rest of the world. For selected years, export shares to Eastern Europe and the Euro area are also featured in the data.

Table 4.1 Estimates of Technology Parameters

Description	Parameter	Estimate	Standard Error
Elasticity of Substitution	σ	1.92	0.45
Scale Complementarity	$\varepsilon_H - \varepsilon_M$	0.31	0.21
	$\varepsilon_L - \varepsilon_M$	−0.40	0.09
Observations	122,842		

Notes: This table presents the parameter estimates of the non-homothetic CES production function. H, M, L denote high, medium, and low skill workers, respectively. Estimates are obtained by implementing the estimation strategy outlined in Section 4.1.2. The construction of the standard errors is detailed in Appendix D. The sample includes all manufacturing firms in the Linked-Employer-Employee Data longitudinal model 1993 - 2014 from the Institute for Employment Research (IAB), and differences are estimated across all available lags.

groups are gross substitutes with an elasticity of substitution equal to $\sigma = 1.92$. High skill types H have the highest scale elasticity, while low skill workers are least complementary with firm scale.

These estimates qualitatively rationalize the sorting patterns displayed in Section 2 as being driven in part by non-homotheticities in production. Thus, the estimates imply that trade liberalization will tend to shift labor demand towards more high skill worker types.

4.2 Calibration of the Remaining Structural Parameters

4.2.1 Parameters & Simulation

Table 4.2 summarizes the parameters that require estimation. I normalize the total labor force of the Home economy and choose the total stock for each skill group s to reflect the distribution of types in the data. The elasticity of substitution across goods - η - is set to 5 from Broda & Weinstein (2006). I parameterize the underlying distribution of firm heterogeneity in demand shifters and amenities by a multivariate log-normal distribution.

To solve the model for a given set of parameters, I discretize the joint distribution of demand shifters φ and amenities A. Here, I sketch the solution algorithm, and Appendix D provides further details.

The initial step is to guess the equilibrium number of firms that choose to enter. Upon entering, firms learn about φ and A and maximize profits taking aggregate equilibrium outcomes - the price index, aggregate labor demand, and expenditure - as given. Hiring and wage decisions depend on a firm's output, which, in turn, affects marginal cost and, therefore, prices and demand. Taken aggregates as given, firms' profit-maximizing choices of scale, wages, and hiring can be solved for as a fixed point.

Table 4.2 List of Parameters

Description	Parameterization	Parameter
Firm Heterogeneity	$(\varphi, A) \sim MV\log - \text{normal}$	$\boldsymbol{\mu}, \boldsymbol{\Sigma}$
Labor Supply Elasticity		β_s
Shifters in Productions		Ω_s
Size of the Foreign Market		Y^*
Domestic Price of Foreign Varities		P^*
Fixed Cost		F_X, F_E

Note: Parameters not estimated: $\eta = 5$, L_s for all worker groups s, $\varepsilon_2 = 0.6$, $\tau = 1.3$.

The optimal pricing decisions of firms yield marginal costs, which aggregate to the domestic price index P. The individual employment decisions of firms on the other labor demand aggregate to labor demand Λ_s for each worker type. Therefore, both the price index and aggregate labor demand solve fixed points.

Optimal firm choices that are consistent with aggregate prices and labor market clearing provide firm-level wages and profits. Wages and profits across firms aggregate to domestic expenditure Y. The algorithm iterates over updated guesses for expenditure Y until convergence.

The final step is to derive expected profits, which updates the mass of entering firms. For a new guess for the number of entering firms, I return to the initial step and re-calculate optimal firm choices that are consistent with product market clearing and labor market clearing. This procedure is repeated until convergence.

4.2.2 Targeted Moments

Table 4.3 presents the list of moments targeted in the estimation.[28] For guidance, the second column of Table 4.3 lists parameters and their associated target moments. I briefly elaborate on the underlying logic for key parameters.

β_s affects the firm-level elasticity of wages with respect to sales, as shown in Proposition 2. For a given distribution of firm sales, the level of β_s affects the aggregate dispersion of wages within worker group s. Therefore differences in wage dispersion between worker groups as well as the relation between firm sales and wages are directly informative about β_s.

The parameters of the variance-covariance matrix $\boldsymbol{\Sigma}$ are associated with the distribution of sales, the level of wage dispersion, and the covariance between wages and employment at the firm level. A higher (lower) covariance between demand shifters and amenities

[28] Appendix A.2 details the construction of the empirical moments.

Table 4.3 Targeted Moments and Associated Parameters

Description	Parameters	Moments
Firm Heterogeneity	$(\varphi, a) \sim$ MV LogNormal$(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Sales distribution
		$\text{CoV}\left(\log W_f, \log l_f\right)$
		Residual Inequality (levels)
Labor Supply Elasticitiy	β_s	Residual Inequality (relative)
		Firm Wage Premia
Shifters in Productions	Ω_s	Mean Skill Wage Premium
Foreign Demand Shifter	Y^*	Export share
Price of Foreign Varities	P^*	Import Share in GDP
Fixed Cost	F_E	Mass of firms,
	F_X	Share of exporters

Notes: This table provides an overview of the empirical moments used for the calibration of the model.

weakens (strengthens) the correlation between wages and employment along the firm sales distribution in the model. I include the correlation between wages and employment across all firms, as well as those in the top quintile of the sales distribution, as targets in the estimation.

The productivity shifters Ω_s affect the relative technological advantage and therefore mean differences in log-wages between worker types. For the estimation, I normalize $\Omega_2 = 1$.

The identification of the parameters related to international trade is standard. An increase in the fixed cost of exporting F_X weakly decreases the share of firms that find it profitable to export. Exporting fixed cost F_X thus govern the share of exporting firms. The (relative) size of the foreign market Y^* governs the export share of exporting plants. Export shares in the structural model are constant and given by $\frac{P_F Q_F}{P_H Q_H} = \tau^{1-\eta} Y^* / \left(P^{\eta-1} Y\right)$. The export share in sales across exporting firms informs Y^*. The structural expression for domestic spending on foreign relative to home goods is given by $\left(\frac{P^*}{P}\right)^{1-\eta}$. P^* directly affects domestic consumption shares. Reductions in P^* increase domestic competition from abroad and tend to increase consumer spending on foreign varieties.

4.2.3 Results

Parameters and Model Fit

Table 4.4 presents the calibrated parameters. Labor supply elasticities β_s are estimated to vary across skill groups, and the model assigns lower labor supply elasticities to skill types that display higher levels of within-group wage inequality in the data. The estimates

Table 4.4 Calibrated Parameters

Parameter	Estimate	Parameter	Estimate
β_{L}	6.9	Ω_{L}	0.14
β_{M}	5.5	Ω_{M}	1
β_{H}	3.9	Ω_{H}	1.75
μ_{φ}	0.03	F_E	0.002
μ_A	0	F_X	0.004
σ^2_{φ}	2.3	P^*	0.005
σ^2_A	0.10	Y^* ($\times 10^{-6}$)	1.3
$\sigma_{A\varphi}$	0.08		

Notes: This table presents the calibrated parameters. β_s is the elasticity of labor supply for skill group s. H, M, L denotes high, medium, and low skill, respectively. μ_{φ} and μ_A are means of the multivariate log Normal that parameterizes the underlying distribution of demand shifters φ and amenities A. σ^2_{φ} is the variance of demand shifters, σ^2_a is the variance of amenities and $\sigma_{a\varphi}$ is the covariance between amenities and demand shifters. Parameters Ω are technological productivity shifters. F_X and F_E denote the fixed cost of exporting and entry, respectively. Gray coloring indicates normalization.

imply that low skill workers receive a mark-down of 13 percent on their marginal revenue product. High skill workers, in turn, receive a mark-down of 20 percent. Equivalently, the estimates imply that workers in group 1 would be willing to sacrifice 13 percent of their earnings to remain employed at their current firm. Conversely, high wage worker types would be willing to sacrifice 20 percent of their earnings to keep their current job. The estimated labor supply elasticities are similar in magnitude to estimates from studies that rely on similar models of the labor market (Lamadon *et al.* (2022), Haanwinckel (2021)).

The estimated parameters of the distribution of firm heterogeneity implies that amenities are moderately positively correlated with demand shifters. The parameters related to international trade - fixed cost of exporting, foreign prices, and foreign demand shifters - compare closely to similar models.[29]

To assess the model fit, Table 4.5 displays empirical targets and the associated simulated moments. Overall, the model manages to fit the targeted moments well. The model adequately captures relative wages at different percentiles of worker-group specific distributions and relative mean wages between groups, as well as the targeted moments related to international trade.

The largest mismatch between empirical and simulated moments occurs in the domestic sales distribution. This, in part, may reflect that a log-normal distribution is only partially suited to fit the entire distribution of firm sales (see, for example, Nigai (2017)).

To check for over-identification, I now assess the model's implications for moments that

[29] Note that Y^* corresponds to the foreign demand shifter and is not equal to foreign relative income, but also includes information on the foreign price index.

Table 4.5 Model Fit

Moment	Model	Data	Moment	Model	Data
Wage Inequallity Within Skill Types			Sales Distr.		
sd (log W)			p25-p10 (logs)	0.76	0.63
Low	0.27	0.27	p50-p10	1.51	1.39
Medium	0.28	0.29	p75-p10	2.51	2.47
High	0.34	0.34	p90-p10	3.52	3.7
$50-10$ ratio (log)			Trade		
Low	0.34	0.37	Exporter Share	0.18	0.18
Medium	0.37	0.41	Export Share Sales	0.22	0.22
High	0.48	0.50	Import Share	0.26	0.26
$90-10$ ratio (logs)			$\log W = \alpha + \gamma_s \log Rev$		
Low	0.70	0.69	γ_L	0.02	0.02
Medium	0.72	0.73	γ_M	0.04	0.03
High	0.88	0.87	γ_H	0.08	0.05
Mean Wage Differences			Corr($\log W$, $\log L$)		
Low to Medium	-0.14	-0.13	Sample	0.16	0.19
High to Medium	0.26	0.27	p80 Sales	0.04	0.07

Notes: This table presents simulated targeted moments and their empirical counterparts. H, M, and L denotes high, medium, and low skill, respectively. Empirical moments capture manufacturing firms and are calculated for the period 1993-2002.

were not directly targeted by the calibration procedure.

Model Fit for Non-Targeted Moments

Employment Distribution The parameters of the production function are directly estimated from firm-level responses in employment and wages to demand shocks. Therefore, the overall distribution of workers along the firm size distribution was not directly targeted. Figure C.1 plots the simulated and empirical distribution of workers along the quintiles of the firm sales distribution. The model adequately captures that the employment share of high skilled workers is increasing, while that of low skill workers is decreasing along the firm size distribution.

Within-Firm Inequality The estimation procedure did not explicitly target the relationship between firm sales and within-firm inequality. I test the model's ability to account for this fact by comparing the relationship between within-firm wage dispersion to firm revenues in the simulated model to the empirical relationship in the data. Further, high skill worker types are paid relatively more than their coworkers in larger firms. I also test this prediction using the data simulated by the model.

Table C.5 compares the model generated regression coefficients to their empirical counterparts. The model adequately accounts for the empirical relationship between within-firm wage dispersion and revenues. While it qualitatively fits the relationship between wage differentials within the firm and firm sales, the effect of firm sales on the relative wages of high skilled workers is larger than in the data, albeit it is only off by a small margin.

Wage distributions for exporting and non-exporting firms Demand shifters φ are the primary predictor of exporting decisions. However, firms of different sizes pay different wages due to heterogeneity in amenities. As opposed to models with a single dimension of firm heterogeneity, some domestic firms, therefore, pay higher wages than exporting firms. This is an important feature of empirical wage distributions.

Figure C.2 displays simulated kernel density plots of log wages across domestic and exporting firms. Consistent with the data, wages paid by domestic firms generally more dispersed. While exporting firms on average pay a wage premium, high skill workers receive higher firm-exporter wage premia, mirroring the fact that firm scale has unequal effects on wages by skill group.

5 Trade and Policy Counterfactual

I now come to the primary concern of my analysis: How does trade affect the distribution of wages, within and between firms, and within and between skill groups? Wage inequality in German manufacturing, as measured by the variance of log earnings, has increased by 19% from 1993-2002 to 2002-2014. Simultaneously, the transformation of former socialist countries in Eastern Europe, and the integration of China into the world economy led to new trade opportunities.[30] In this section, I use the estimated model to study how trade shocks have contributed to the observed increase in wage inequality. I then evaluate the effect of a tax intervention that was theoretically derived in Proposition 5, and that aims to improve equilibrium allocations in trade and product markets.

5.1 The Distributional Consequences of Trade

5.1.1 Changes in German Trade Participation

Between the 1990s and the first decade of the 2000s, Germany increased its trade participation significantly, which reflects both the expansion into new Eastern markets in Europe

[30] For an analyis of the reduced-form effects of these events on German labor markets, see Dauth *et al.* (2014, 2017, 2018).

Table 5.1 Empirical Changes in Trade Related Moments

	1993-2002	2003-2014	Change
Share of Exporting Firms	0.18	0.25	+7 p.p.
Firm Export Share in Sales	0.24	0.29	+5 p.p.
Import Share in GDP	0.25	0.36	+11 p.p.

Notes: This table displays empirical changes in trade-related moments. p.p. denotes percentage points. The share of exporting firms and the export share in sales in manufacturing firms are calculated from the IAB firm panel. Average import shares are taken from the World Bank's World Development Indicators.

as well as the entry of China into the World Trade Organization. Table 5.1 shows that the share of exporters has increased by 40 percent during the two periods. The average foreign sales share among exporters in turn increased by 25 percent, while the aggregate import share has increased by 35 percent.

For the counterfactual analysis, I consider changes in variable exporting trade cost τ (-11 percent), fixed cost of exporting F_X (-9 percent), and the foreign price index P^* (-15 percent), which match the empirical changes displayed in Table 5.1.

5.1.2 Counterfactual Effect of Trade on the Distribution of Earnings

The counterfactual exercises indicate that trade had a sizeable effect on overall wage inequality. Table 5.1 shows that overall wage inequality increases by 4.2 percent, which accounts for 22 percent of the overall 19 percent increase in wage inequality in the data.

The majority of studies focusing on the effect of international trade on residual inequality abstract from within-firm inequality (e.g., Helpman *et al.* (2017), Egger *et al.* (2013) and Amiti & Davis (2012)). Table 5.1 shows that within-firm inequality accounts for one third of the total counterfactual distributional changes.[31] The last column of Table 5.1 shows that 16 percent of the overall counterfactual change in inequality is due to heterogeneous changes in firm wage premia between worker groups, indicating that both a worker's employer, as well as his type, are important determinants of his wage exposure to international trade.

Table C.6 provides a more detailed overview of the counterfactual effects on wage inequality. Counterfactual changes imply that trade explains about a third of the observed increase in average wages between groups. This reflects that aggregate labor demand shifts toward workers in high wage worker groups as firms increase their scale in response to the export demand shock.

[31] Here, I use the fact that the aggregate variation in wages can be decomposed as: $\mathrm{Var}\left(\log W_{s,f}\right) = \mathrm{Var}\left(\log W_{s,f} - \log \overline{W}_f\right) + \mathrm{Var}\left(\log \overline{W}_f\right)$,where $\overline{W}_f$ denotes the mean wage paid by employer f.

Table 5.2 Counterfactual Effect of Trade on Aggregate Earnings Inequality

Overall Change		%
Aggregate Inequality	$\text{Var}(\log W_{f,s})$	+4.2
Decomposition		% of total
Within Firm Inequality	$\text{Var}\left(\log W_{f,s} - \log \overline{W}_f\right)$	34
Between Firm Inequality	$\text{Var}\left(\log \overline{W}_f\right)$	66
Heterogeneous Skill Wage Premia	$\text{Var}\left(\log \overline{W}_{f,s} - \log \overline{W}_f - \log \overline{W}_s\right)$	16

Notes: This table displays the simulated counterfactual effect of trade on wage inequality. $\log \overline{W}_f$ denotes mean log wages of firm f. $\log \overline{W}_s$ denotes mean log wages across all workers belonging to skill group s.

Table C.6 further shows that trade has the largest effect on wage inequality within high skill workers. The standard deviation of wages within high skill workers increases by 2.8 percent, while the data shows an increase of 14 percent for the same period. This, in turn, is partly driven by the fact that between-firm inequality within high skill workers is rising, as indicated by the increase in the 90-10 pay ratio. In comparison, counterfactual wage inequality within low skill workers in group 1 increases by 1.4 percent, compared with an empirical increase of 7 percent in the data.

5.2 Optimal Policy: Imperfect Competition in Labor Markets and the Gains from Trade

The parameter estimates imply that labor market sorting is inefficient in equilibrium. As a consequence, the gains from trade might be affected by inefficiencies in the labor market and product market outcomes. Proposition 5 implies that the size of these losses can be quantified by implementing an optimal tax reform to reevaluate the same trade counterfactual. Table 5.3 displays the set of correcting proportional taxes. The correcting taxes are progressive in skill: Since the labor supply of skilled workers is less elastic, incentivizing the best firms to hire more such workers requires larger wage-bill subsidies. As discussed previously, the net welfare effects of the tax reform depends on allocative gains in labor and product markets and direct losses in labor income stemming from taxation.

The model-implied changes in welfare, relative to the initial counterfactual trade equilibrium, are displayed in Table 5.4. The results suggest that imperfect competition in labor markets creates significant distortions that undermine the gains from trade for workers. Relative to the trade counterfactual, the tax reform increases worker welfare on average by 6 percent. The gains associated with improved market allocations, therefore, outweigh potential wage losses stemming from taxation.

Table 5.3 Optimal Tax Reform

Skill Group	Payroll subsidy/Income Tax	
	Theory	Implementation
Low		13%
Medium	$\frac{1}{1+\beta_s}$	15%
High		20%

*Notes:*This table displays the set of income taxes that, according to 5, restores the Walrasian equilibrium allocation in labor and product markets. In the Walrasian equilibrium, workers earn their marginal revenue product of labor, rather than a monopsonistic mark-down on their marginal revenue product.

The remaining columns in Table 5.4 give further insights into the channels through which these gains materialize. Relative to the initial counterfactual allocation, the tax reform reduces the profit share in income by 17 percent. The tax reform redistributes national income from firms to workers, which reflects that firms produce at more efficient scale. The ratio of variable profits to total wage cost of a (domestic) firm before the tax reform is given by:

$$\frac{P_f Q_f - C_f}{C_f} = \left(1 - \frac{\eta - 1}{\eta} \cdot \frac{1}{\mathcal{W}_f}\right), \tag{5.1}$$

where $\mathcal{W}_f = \sum_s \omega_{s,f} \left(\frac{\beta_s+1}{\beta_s} \cdot \frac{\varepsilon_s - \sigma}{1-\sigma}\right)$ is the previously defined wedge between average and marginal cost. After the tax reform, this wedge reduces to $\widetilde{\mathcal{W}}_f = \sum_s \omega_{s,f} \left(\frac{\varepsilon_s - \sigma}{1-\sigma}\right)$. The tax reform, therefore, increases the gains from trade for workers by reallocating profits from workers to firms.

The tax reform also increases the share of exporting firms. The exercise thus illustrates that imperfect competition in labor markets reduces the gains from trade through an extensive margin by reducing the number of firms that reach sufficient scale to find it profitable to export.

The effects of the tax reform on labor market allocations reduce wage dispersion within groups. Within-group inequality decreases for all groups under the tax reform. Relative to the initial increase in inequality induced by trade, the tax reform decreases within-group inequality on average by 20 percent across worker groups.

Between groups, wage inequality increases. This reflects that the sorting of high skilled workers to large firms strengthens, which is illustrated in Figure 5.1 After the tax, mass in the distribution of the employment share of skilled workers across firms shifts from the left-tail into newly exporting firms.

In sum, wage inequality decreases after the tax reform, implying that the increase in

TABLE 5.4 EFFECT OF A TAX REFORM ON COUNTERFACTUAL CHANGES FROM TRADE

Outcome		Change to Trade Counterfactual (%)
Worker Welfare	$\mathbb{E} \log U$	
Low Skill		+7
Medium Skill		+6
High Skill		+5
Profit Share of Income	Π / Y	-17
Share of Exporting Firms		+3
Overall Wage Inequality	$\text{Var}(\log W_s)$	-0.3
Within Group Inequality	$sd(\log W_s)$	
Low Skill		-0.5
Medium Skill		-0.3
High Skill		-0.3
Between Group Inequality	$\log \overline{W}_s / \overline{W}_2$	
Low to Medium Skill		+0.4
High to Low Skill		+0.1

Notes: This table shows THE percentage change in the estimated effect of a tax reform relative to the trade counterfactual. $\log \overline{W}_s$ denotes mean log wages across all workers belonging to group s.

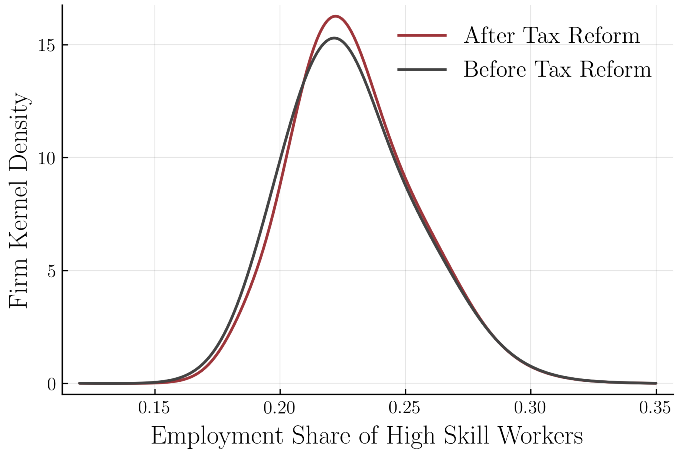
aggregate skill premia is offset by a more than proportional decrease in within-group inequality.

To summarize, the results imply that imperfect competition in labor markets affects the gains from trade for workers through at least three channels: The allocation of national income to profits and wages, the share of firms that export, and the sorting of workers to firms.

6 Conclusion

The distributional consequences of trade liberalization is one of the most hotly-debated issues of our time. In this paper, I outline a hitherto neglected mechanism through which international trade affects wage inequality. The key to this theoretical mechanism is that larger firms pay higher wages and have higher employment for more skilled workers, which induces wage inequality between firms for a given worker type, and wage inequality within firms between worker types. As international trade costs fall, more productive firms expand as they enter export markets, while less productive firms

Figure 5.1 Tax Induced Reallocation of High Skill Workers



Notes: This figure displays simulated kernel density plots of firm-level employment shares of high skilled workers after trade liberalization. Blue indicates the employment shares before the tax reform, while red indicates employment shares after the tax reform.

contract in the domestic market, which increases the dispersion in employment across firms. As contracting firms reduce their relative wage and employment for more skilled workers, and expanding firms increase their relative wages and employment for more skilled workers, this increases overall wage inequality through both higher between and within-firm wage inequality.

To provide microfoundations for the observed differences in relative wages and employment across firms of different sizes, I develop a theoretical model in which firms operate a non-homothetic production technology and face an upward-sloping labor supply function for each type of worker. The non-homotheticity in production implies that firms of different sizes have different relative productivities and relative demands for workers with different skills. Combining these differences in relative demand for more and less-skilled workers with an upward-sloping supply function for each type of worker induces the positive comovement between relative wages and relative employment across firms of different sizes.

To quantify the importance of this mechanism for the distributional consequences of trade liberalization, I structurally estimate the model using matched employer-employee data and a new method that separately identifies the elasticities of labor demand and labor supply. I show that the estimated model can account for the reduced-form patterns in the data not only qualitatively, but also quantitatively. I show that the model is successful in matching both targeted and non-targeted moments, including the observed patterns of wage inequality both between and within firms.

In counterfactuals, I show that this new mechanism is quantitatively relevant for understanding the impact of trade on wage inequality. The changes in trade costs implied by the observed data raise overall wage inequality by 4.2 percent, which corresponds to 22 percent of the overall increase in wage inequality in Germany during my sample period. I find that the imperfect competition introduced into the labor market by upward-sloping labor supply curves leads to quantitatively relevant distortions in the allocations of skilled and unskilled workers across firms of different sizes. Implementing the optimal tax scheme to eliminate these distortions, decreases wage inequality within skill groups by 20 percent, raises worker welfare by 6 percent on average, and therefore increases the magnitude of the welfare gains of trade.

References

Abowd, John M., Francis Kramarz and David N. Margolis, 'High Wage Workers and High Wage Firms.' *Econometrica*, **67** (2), pp. 251–334, *1999*.

Adão, Rodrigo, Michal Kolesár and Eduardo Morales, 'Shift-Share Designs: Theory and Inference.' *The Quarterly Journal of Economics*, **134** (4), pp. 1949–2010, *2019*.

Almunia, Miguel, Pol Antras, David Lopez-Rodriguez and Eduardo Morales, 'Venting Out: Exports during a Domestic Slump.' working paper, *2018*.

Alvarez, Jorge, Felipe Benguria, Niklas Engbom and Christian Moser, 'Firms and the Decline in Earnings Inequality in brazil.' *American Economic Journal: Macroeconomics*, **10** (1), pp. 149–189, *2018*.

Amiti, Mary and Donald R. Davis, 'Trade, Firms, and Wages: Theory and Evidence.' *Review of Economic Studies*, **79** (1), pp. 1–36, *2012*.

Arkolakis, Costas, Arnaud Costinot, Dave Donaldson and Andrés Rodrìguez-Clare, 'The Elusive Pro-Competitive Effects of Trade.' *Review of Economic Studies*, **86** (1), pp. 46–80, *2019*.

_ , _ and Andres Rodriguez-Clare, 'New Trade Models, Same Old Gains?.' *American Economic Review*, **102** (1), pp. 94–130, *2012*.

Atkeson, Andrew and Ariel T. Burstein, 'Innovation, Firm Dynamics, and International Trade.' *Journal of Political Economy*, **118** (3), pp. 433–484, *2010*.

Bartik, Timothy J., *Who Benefits from State and Local Economic Development Policies?*, Books from Upjohn Press: W.E. Upjohn Institute for Employment Research, *1991*.

Bauer, Arthur, Joceylyn Boussard and Danial Lashkari, 'Information Technology and Returns to Scale.' working paper, *2019*.

Berger, David W., Kyle F. Herkenhoff and Simon Mongey, 'Labor Market Power.' Opportunity and Inclusive Growth Institute Working Papers 48, Federal Reserve Bank of Minneapolis, *2021*.

Berman, Nicolas, Antoine Berthou and Jerome Hericourt, 'Export dynamics and sales at home.' *Journal of International Economics*, **96** (2), pp. 298–310, *2015*, DOI: http://dx.doi.org/10.1016/j.jinteco.2015.04.

Bernard, Andrew B. and J. Bradford Jensen, 'Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987.' *Brookings Papers on Economic Activity*, **26** (1995 Micr), pp. 67–119, *1995*.

—, —, **Stephen J. Redding and Peter K. Schott**, 'Firms in International Trade.' *Journal of Economic Perspectives*, **21** (3), pp. 105–130, *2007*.

Blaum, Joaquin, Claire Lelarge and Michael Peters, 'Firm size, quality bias and import demand.' *Journal of International Economics*, **120** (C), pp. 59–83, *2019*, DOI: http://dx.doi.org/10.1016/j.jinteco.2019.04.

Bonhomme, Stéphane, Thibaut Lamadon and Elena Manresa, 'A Distributional Framework for Matched Employer Employee Data.' *Econometrica*, **87** (3), pp. 699–739, *2019*.

Borovickova, Katarina and Robert Shimer, 'High Wage Workers Work for High Wage Firms.' NBER Working Papers 24074, National Bureau of Economic Research, Inc, *2019*.

Borusyak, Kirill, Peter Hull and Xavier Jaravel, 'Quasi-experimental Shift-share Designs.' *Review of Economic Studies*, **1** (89), pp. 181–213, *2022*.

Broda, Christian and David E. Weinstein, 'Globalization and the Gains From Variety.' *The Quarterly Journal of Economics*, **121** (2), pp. 541–585, *2006*.

Brown, Charles and James Medoff, 'The Employer Size-Wage Effect.' *Journal of Political Economy*, **97** (5), pp. 1027–1059, *1989*.

Burdett, Kenneth and Dale T Mortensen, 'Wage Differentials, Employer Size, and Unemployment.' *International Economic Review*, **39** (2), pp. 257–273, *1998*.

Burstein, Ariel, Eduardo Morales and Jonathan Vogel, 'Changes in Between-Group Inequality: Computers, Occupations, and International Trade.' *American Economic Journal: Macroeconomics*, **11**, pp. 348–400, *2019*.

— **and Jonathan Vogel**, 'International Trade, Technology, and the Skill Premium.' *Journal of Political Economy*, **125** (5), pp. 1356–1412, *2017*.

Bustos, Paula, 'Trade Liberalization, Exports and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian firms.' *American Economic Review*, **101**, pp. 304–340, *2011a*.

—, 'The Impact of Trade Liberalization on Skill Upgrading Evidence from Argentina.' Working Papers 559, Barcelona Graduate School of Economics, *2011b*.

Cahuc, Pierre, Francois Marque and Etienne Wasmer, 'A Theory Of Wages And Labor Demand With Intra-Firm Bargaining And Matching Frictions.' *International Economic Review*, **49** (3), pp. 943–972, *2008*.

Caliendo, Lorenzo, Maximiliano Dvorkin and Fernando Parro, 'Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock.' *Econometrica*, **87** (3), pp. 741–835, *2019*.

— **, Ferdinando Monte and Esteban Rossi-Hansberg**, 'Exporting and Organizational Change.' CEPR Discussion Papers 12177, C.E.P.R. Discussion Papers, *2017*.

— **and Esteban Rossi-Hansberg**, 'The Impact of Trade on Organization and Productivity.' *The Quarterly Journal of Economics*, **127** (3), pp. 1393–1467, *2012*.

Card, David, Ana Rute Cardoso, Joerg Heining and Patrick Kline, 'Firms and Labor Market Inequality: Evidence and Some Theory.' *Journal of Labor Economics*, **36** (S1), pp. 13–70, *2018*.

— **, Joerg Heining and Patrick Kline**, 'Workplace Heterogeneity and the Rise of West German Wage Inequality.' *The Quarterly Journal of Economics*, **128** (3), pp. 967–1015, *2013*.

Caron, Justin, Thibault Fally and James Markusen, 'Per Capita Income and the Demand for Skills.' NBER Working Papers 23482, National Bureau of Economic Research, Inc, *2017*.

— **,** — **and James R. Markusen**, 'International Trade Puzzles: A Solution Linking Production and Preferences.' *The Quarterly Journal of Economics*, **129** (3), pp. 1501–1552, *2014*.

Chan, Mons, Ming Xu and Sergio Salgado, 'Heterogeneous Passthrough from TFP to Wages.' 2019 Meeting Papers 1447, Society for Economic Dynamics, *2019*.

Chetty, Raj, 'Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods.' *Annual Review of Economics*, **1** (1), pp. 451–488, *2009*.

Cho, David and Alan Krueger, 'Rent Sharing Within Firms.'Technical report, *2019*.

Comin, Diego A, Danial Lashkari and Marti Mestieri, 'Structural Change with Long-Run Income and Price Effects.' *NBER Working Paper*, *2017*.

Costinot, Arnaud and Jonathan Vogel, 'Matching and Inequality in the World Economy.' *Journal of Political Economy*, **118** (4), pp. 747–786, *2010*.

Cravino, Javier and Sebastian Sotelo, 'Trade Induced Structural Change and the Skill Premium.' NBER Working Papers 23503, National Bureau of Economic Research, Inc, *2017*.

Dauth, Wolfgang, Sebastian Findeisen and Jens Suedekum, 'The Rise Of The East And The Far East: German Labor Markets And Trade Integration.' *Journal of the European Economic Association*, **12** (6), pp. 1643–1675, *2014*.

— , — and — , 'Trade and Manufacturing Jobs in Germany.' *American Economic Review*, **107** (5), pp. 337–342, *2017*.

— , — and — , 'Adjusting to Globalization in Germany.' IZA Discussion Papers 11299, Institute for the Study of Labor (IZA), *2018*.

Davidson, Carl, Fredrik Heyman, Steven Matusz, Fredrik Sjöholm and Susan Chun Zhu, 'Liberalized Trade and Worker-Firm Matching.' *American Economic Review*, **102** (3), pp. 429–434, *2012*.

— , — , — , — and — , 'Globalization and imperfect labor market sorting.' *Journal of International Economics*, **94** (2), pp. 177–194, *2014*.

— **and Steven J. Matusz**, 'Trade Liberalization And Compensation.' *International Economic Review*, **47** (3), pp. 723–747, *2006*.

— , — **and Andrei Shevchenko**, 'Globalization and firm level adjustment with imperfect labor markets.' *Journal of International Economics*, **75** (2), pp. 295–309, *2008*.

Davis, Donald R. and James Harrigan, 'Good jobs, bad jobs, and trade liberalization.' *Journal of International Economics*, **84** (1), pp. 26–36, *2011*.

Dustmann, Christian, Johannes Ludsteck and Uta Schonberg, 'Revisiting the German Wage Structure.' *Quarterly Journal of Economics*, **124** (2), pp. 843–881, *2009*.

Eckert, Fabian, 'Growing Apart: Tradable Services and the Fragmentation of the US Economy.' working paper, *2019*.

— **, Sharat Ganapati and Conor Walsh**, 'Skilled Tradable Services: The Transformation of US High-Skill Labor Markets.' working paper, *2019*.

Eeckhout, Jan and Philipp Kircher, 'Assortative Matching with Large Firms.' *Econometrica*, **86(1)**, pp. 85–132, *2018*.

— **and Roberto Pinheiro**, 'Diverse Organizations and the Competition for Talent.' *International Economic Review*, **55** (3), pp. 625–664, *2014*, DOI: http://dx.doi.org/10.1111/iere.12065.

Egger, Hartmut, Peter Egger and Udo Kreickemeier, 'Trade, wages, and profits.' *European Economic Review*, **64** (C), pp. 332–350, *2013*, DOI: http://dx.doi.org/10.1016/j.euroecorev.2013.

— **and Udo Kreickemeier**, 'Firm Heterogeneity and the Labor Market Effects of Trade Liberalization.' *International Economic Review*, **50** (1), pp. 187–216, *2009*.

— **and** — , 'Fairness, trade, and inequality.' *Journal of International Economics*, **86** (2), pp. 184–196, *2012*, DOI: http://dx.doi.org/10.1016/j.jinteco.2011.10.

Feenstra, Robert C, 'New Product Varieties and the Measurement of International Prices.' *American Economic Review*, **84** (1), pp. 157–177, *1994*.

Felbermayr, Gabriel J., Mario Larch and Wolfgang Lechthaler, 'Unemployment in an Interdependent World.' *American Economic Journal: Economic Policy*, **5** (1), pp. 262–301, *2013*.

Felbermayr, Gabriel, Julien Prat and Hans-Jörg Schmerer, 'Globalization and labor market outcomes: Wage bargaining, search frictions, and firm heterogeneity.' *Journal of Economic Theory*, **146** (1), pp. 39–73, *2011*.

Fieler, Ana Cecilia, Marcela Eslava and Daniel Yi Xu, 'Trade, Quality Upgrading, and Input Linkages: Theory and Evidence from Colombia.' *American Economic Review*, **108** (1), pp. 109–146, *2018*.

Fischer, Gabriele, Florian Janik, Dana Mueller and Alexandra Schmucker, 'The IAB Establishment Panel - things users should know.' In *Schmollers Jahrbuch, Zeitschrift fur Wirtschafts-and Sozialwissenschaften*, **129**, Chapter 1, pp. 133–148, *2009*.

Gabaix, Xavier and Augustin Landier, 'Why has CEO Pay Increased So Much?.' *The Quarterly Journal of Economics*, **123** (1), pp. 49–100, *2008*.

Garicano, Luis, 'Hierarchies and the Organization of Knowledge in Production.' *Journal of Political Economy*, **108** (5), pp. 874–904, *2000*.

— **and Esteban Rossi-Hansberg**, 'Knowledge-Based Hierarchies: Using Organizations to Understand the Economy.' *Annual Review of Economics*, **7** (1), pp. 1–30, *2015*.

Garin, Andrew and Filipe Silvero, 'How Responsive are Wages to Demand within the Firm? Evidence from Idiosyncratic Export Demand Shocks.' working paper, *2018*.

Gautier, Pieter A. and Coen N. Teulings, 'Sorting And The Output Loss Due To Search Frictions.' *Journal of the European Economic Association*, **13** (6), pp. 1136–1166, *2015*.

Gautier, Pieter, Coen Teulings and Aico Van Vuuren, 'On-the-job-search, mismatch and Efficiency.' *Review of Economic Studies*, **77** (1), pp. 245–272, *2010*.

Goldin, Claudia and Lawrence F Katz, 'Technology, Skill, and the Wage Structure: Insights from the Past.' *American Economic Review*, **86** (2), pp. 252–257, *1996*.

Goldsmith-Pinkham, Paul, Isaac Sorkin and Henry Swift, 'Bartik Instruments: What, When, Why, and How.' *American Economic Review*, **110** (8), pp. 2586–2624, *2020*, DOI: http://dx.doi.org/10.1257/aer.20181047.

Grossman, Gene M., Elhanan Helpman and Philipp Kircher, 'Matching, Sorting, and the Distributional Effects of International Trade.' *Journal of Political Economy*, **125** (1), pp. 224–264, *2017*, DOI: http://dx.doi.org/10.1086/689608.

Haanwinckel, Daniel, 'Supply, Demand, Institutions and FIrms: A Theory of Labor Market Sorting and the Wage Distribution.' working paper, *2021*.

Hanoch, Giora, 'Production and Demand Models with Direct or Indirect Implicit Additivity.' *Econometrica*, **43** (3), pp. 395–419, *1975*.

Harrigan, James and Ariell Reshef, 'Skill-biased heterogeneous firms, trade liberalization and the skill premium.' *Canadian Journal of Economics*, **48** (3), pp. 1024–1066, *2015*.

Heining, Joerg, Wolfram Klosterhuber, Patrick Lehnert and Stefan Seth, ' An overview on the Linked EmployerEmployee Data of the Institute for Employment Research (IAB).' In *Schmollers Jahrbuch, Zeitschrift fur Wirtschafts-and Sozialwissenschaften*, **134**, Chapter 1, pp. 141–148, *2014*.

_ , _ , _ and _ , 'Linked Employer-Employee Data from the IAB: LIAB Longitudinal Model 1993-2014 (LIAB LM 9314).' fdz data report, *2016*.

Helpman, Elhanan and Oleg Itskhoki, 'Labour Market Rigidities, Trade and Unemployment.' *Review of Economic Studies*, **77** (3), pp. 1100–1137, *2010*.

_ , _ , Marc-Andreas Muendler and Stephen J. Redding, 'Trade and Inequality: From Theory to Estimation.' *Review of Economic Studies*, **84** (1), pp. 357–405, *2017*.

_ , _ and Stephen Redding, 'Inequality and Unemployment in a Global Economy.' *Econometrica*, **78** (4), pp. 1239–1283, *2010*.

Hummels, David, Rasmus Jorgensen, Jakob Munch and Chong Xiang, 'The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data.' *American Economic Review*, **104** (6), pp. 1597–1629, *2014*.

Kennan, John, 'Uniqueness of Positive Fixed Points for Increasing Concave Functions on Rn: An Elementary Result.' *Review of Economic Dynamics*, **4** (4), pp. 893–899, *2001*, DOI: http://dx.doi.org/10.1006/redy.2001.0139.

Kline, Patrick and Enrico Moretti, 'Place Based Policies with Unemployment.' *American Economic Review*, **103** (3), pp. 238–243, *2013*.

Lamadon, Thibaut, Magne Mogstad and Bradley Setzler, 'Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market.' *American Economic Review*, **112** (1), pp. 169–212, *2022*, DOI: http://dx.doi.org/10.1257/aer.20190790.

Leamer, Edward E, 'Is It a Demand Curve, or Is It a Supply Curve? Partial Identification through Inequality Constraints.' *The Review of Economics and Statistics*, **63** (3), pp. 319–327, *1981*.

Lise, Jeremy, Costas Meghir and Jean-Marc Robin, 'Matching, Sorting and Wages.' *Review of Economic Dynamics*, **19**, pp. 63–87, *2016*, DOI: `http://dx.doi.org/10.1016/j.red.2015.11.004`.

Lucas, Robert E., 'On the Size Distribution of Business Firms.' *Bell Journal of Economics*, **9** (2), pp. 508–523, *1978*.

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

—, 'Imperfect Competition in the Labor Market.' **4B**: Elsevier, 1st edition, Chapter 11, pp. 973–1041, *2011*.

— **and Barbara Petrongolo**, 'How Local Are Labor Markets? Evidence from a Spatial Job Search Model.' *American Economic Review*, **107** (10), pp. 2877–2907, *2017*.

Matsuyama, Kiminori, 'The Home Market Effect and Patterns of Trade of Rich and Poor Countries.' working paper, *2015*.

McFadden, D., *Conditional Logit Analysis of Qualitative Choice Behaviour*: Academic Press New York, pp.105-142, *1973*.

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

Nigai, Sergey, 'A tale of two tails: Productivity distribution and the gains from trade.' *Journal of International Economics*, **104** (C), pp. 44–62, *2017*, DOI: `http://dx.doi.org/10.1016/j.jinteco.2016.10`.

Oberfield, Ezra and Devesh Raval, 'Micro Data and Macro Technology.' NBER Working Papers 20452, National Bureau of Economic Research, Inc, *2014*.

Ohnsorge, Franziska and Daniel Trefler, 'Sorting It Out: International Trade with Heterogeneous Workers.' *Journal of Political Economy*, **115** (5), pp. 868–892, *2007*.

Oi, Walter Y. and Todd L. Idson, 'Firm size and wages.' In O. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics*, **3** of Handbook of Labor Economics: Elsevier, Chapter 33, pp. 2165–2214, *1999*.

Parro, Ferando, 'Capital-Skill Complementarity and the Skill-Premium.' *American Economic Journal: Macroeconomics*, **5** (2), pp. 72–117, *2013*.

Ritter, Moritz, 'Trade and inequality in a directed search model with firm and worker heterogeneity.' *Canadian Journal of Economics*, **48** (5), pp. 1902–1916, *2015*.

Robinson, Joan, *The Economics of Imperfect Competition*: London: Macmillan, *1933*.

Rosen, Sherwin, 'The Economics of Superstars.' *American Economic Review*, **71** (5), pp. 845–858, *1981*.

—, 'The theory of equalizing differences.' In O. Ashenfelter and R. Layard (eds.) *Handbook of Labor Economics*, **1** of Handbook of Labor Economics: Elsevier, Chapter 12, pp. 641–692, *1987*.

Sampson, Thomas, 'Selection into Trade and Wage Inequality.' *American Economic Journal: Microeconomics*, **6** (3), pp. 157–202, *2014*, DOI: http://dx.doi.org/10.1257/mic.6.3.157.

Sato, Ryuzo, 'The Most General Class of CES Functions.' *Econometrica*, **43**, pp. 999–1003, *1975*.

Shimer, Robert, 'The Assignment of Workers to Jobs in an Economy with Coordination Frictions.' *Journal of Political Economy*, **113** (5), pp. 996–1025, *2005*.

Sobol, Ilya, 'A Distribution of points in a cube and approximate evaluation of integrals.' *USSR Comput. Maths. Math, Phys*, **7**, pp. 86–112, *1967*.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom and Till von Wachter, 'Firming up Inequality*.' *The Quarterly Journal of Economics*, p. qjy025, *2018*, DOI: http://dx.doi.org/10.1093/qje/qjy025.

Stole, Lars A. and Jeffrey Zwiebel, 'Intra-Firm Bargaining under Non-Binding Contracts.' *The Review of Economic Studies*, **63** (3), pp. 375–410, *1996*.

Verhoogen, E., 'Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector.' *Quarterly Journal of Economics*, **123** (2), *2008*.

Yeaple, Stephen, 'A simple model of firm heterogeneity, international trade, and wages.' *Journal of International Economics*, **65** (1), pp. 1–20, *2005*.

A Data Appendix

A.1 Data Overview

Throughout the paper, I use the LIAB, a linked employer-employee data set provided by the German Institute for Employment Research (IAB). It combines information from the IAB Establishment Panel with information on all workers who were employed in one of these firms as of the 30th of June in each year between 1993 and 2014. The firm-level variables used in this paper include:

- Sales, as measured by overall revenues. Values before 2001 are reported in Deutsche Mark and converted to EUROs using an exchange rate of 1.95.
- Intermediate input shares in sales.
- Value-added: Measured as the non-intermediate input share in sales.
- Exports: Export share in sales. Across all years, the overall export share in sales is available. For selected years, this is broken down into sales in countries added in the course of the EU 2004 expansion and sales in countries belonging to the EURO monetary currency union.
- Industry affiliation.

To calculate aggregates, I use the provided sampling weights.

The data on workers is drawn from a sample of administrative social security records in Germany and is representative of all individuals covered by the social security system, roughly 80 percent of the German workforce and contains information on daily earnings, unemployment spells, education (no high school, high school, university), training (completed vocational training, university degree), gender, age, tenure at the firm and occupation at the three-digit level. Throughout, I focus on full-time workers and exclude trainees or voluntary workers from the sample.

An important limitation of the dataset is the top coding of wages. In particular, about 9% of wages of the overall sample in each year are censored at the reporting maximum. Therefore, I impute wages of top earners following the approach in Card *et al.* (2013) and Dustmann *et al.* (2009). I estimate censored regressions estimated for each year and a total of 25 age-training cells, allowing the variance to differ within each cell. For each year, censored wages are imputed as the sum of the predicted wage and a random component, drawn from a normal distribution with mean zero and a separate variance for each year-age-training group, obtained from the standard error of the forecast for uncensored wages. This imputation procedure increases the mean of log wages by about 0.03 and the standard deviation of log wages by about 0.05 in each year.[32]

A.2 Construction of Empirical Moments for Model Estimation

To estimate the model, I construct the following empirical moments for manufacturing.

- Normalized distribution of log sales: Sales of establishments are normalized by average sales of their sector of production. Both averages and the resulting firm sales distribution are weighted using the sampling weights provided by the IAB establishment panel.
- Wage distribution by worker types: Moments related to workers' wages are constructed by constructing wages $W_{s,f}$ at the worker group, and establishment level. The provided establishment sampling weights are used to construct aggregate wage distributions. To control for heterogeneity among that is not captured by the model, $W_{s,f}$ is defined as establishment-worker-group-year fixed effects of a Mincer wage regression of individual log wages on worker observables fixed effects for age groups, occupations, sectors of employment, federal state that the employer is located in and years. Wages $W_{s,f}$ are then used to construct aggregate wage distributions, using the establishment panel weights provided by the survey.

[32] For affected wages, wages on average increase by 0.5 with a standard deviation of 0.4. This corresponds to an average increase of 10% over the reporting maximum.

B Theoretical Appendix

B.1 Properties of Non-Homothetic CES

Following the exposition in Comin *et al.* (2017), consider the following generalization of the production function in the main text:

$$\sum_s \Omega_s^{1/\sigma} \left(\frac{L_i}{g_s(Q)} \right)^{(\sigma-1)/\sigma} = 1. \tag{B.1}$$

The production function in the main text is a special case for $g_s(Q) = Q^{-\frac{\varepsilon_i - \sigma}{\sigma - 1}}$. Thus, $g_s'(Q) < 0$ corresponds to the case where $-\frac{\varepsilon_s - \sigma}{\sigma - 1} < 0$, that is $\frac{\varepsilon_s - \sigma}{\sigma - 1} > 0$.

Lemma 1. *If $\sigma > (<)1$ and $g_s'(Q) > (<)0$, then $Q(\mathbf{L})$ as implicitly defined by (B.1) is strictly increasing and concave.*

Proof. Establishing monotonicity is straightforward. To show concavity, suppose for the sake of a contradiction that there exist Q' and Q'' such that $\mathcal{Q} \equiv \mathcal{Q}(\alpha Q' + (1-\alpha)Q'')$ is strictly smaller than both Q' and Q''. For the case $\sigma \geq 1$ and thus $g_s'(Q) > 0$,

$$\begin{array}{rcll}
1 & = & \sum_s \theta_s^{1/\sigma} \left(\alpha \frac{L_i}{g_s(\mathcal{Q})} + (1-\alpha) \frac{L_i}{g_s(\mathcal{Q})} \right)^{\frac{\sigma-1}{\sigma}} & g_s(\mathcal{Q}) < \min\{g_s(\mathcal{Q}'), g_s(\mathcal{Q}'')\} \\
 & > & \sum_s \theta_s^{1/\sigma} \left(\alpha \frac{L_i}{g_s(Q')} + (1-\alpha) \frac{L_i}{g_s(Q'')} \right)^{\frac{\sigma-1}{\sigma}} & \sigma > 1 \\
 & \geq & \alpha \sum_s \theta_s^{1/\sigma} \left(\frac{L_i}{g_s(Q')} \right)^{(\sigma-1)/\sigma} + (1-\alpha) \sum_s \theta_s^{1/\sigma} \left(\frac{L_i}{g_s(Q'')} \right)^{(\sigma-1)/\sigma}, &
\end{array}$$

where in the first inequality we have used monotonicity of $g_s(.)$ and in the second we have used the assumption that $\sigma > 1$. The last line implies a contradiction to the initial assertion. For the case that $\sigma < 1$,the inequalities are reversed, also implying a contradiction. ■

This implies that $\frac{\varepsilon_s - \sigma}{1-\sigma} > 0$ for all s is a sufficient conditions to ensure that the implicitly defined production function in is strictly increasing and concave in all its arguments.

B.2 Microfoundations for non-homothetic Production Functions

B.2.1 Multinomial Choice with Scale Effects

Random utility models may microfound CES utility if an appropriate unobserved underlying shock to preferences taking the form of a Frechet or logit distribution is assumed (McFadden (1973)). Here, I demonstrate how this argument extends to production in the context of a task-based theory of input choice. In particular, I will demonstrate how such model maps into observational equivalent relative wage bill shares.

Assume that a firm has to complete a unit interval of tasks $j \in [0,1]$ in order to produce a unit of output. For each task s, the firm can hire a unit of time of a worker of type $s \in S$ at wage w_s to complete the specified task. The overall cost of having a task j completed by occupation s is assumed to be log-linear and is specified as follows:

$$\log c\,(j,s,Q) = \sigma \log w_s - \log \theta_s^{1/\sigma} - \frac{\varepsilon_s - 1}{1-\sigma} \log Q + \nu_{j,s}, \tag{B.2}$$

where $\nu_{j,s}$ is an idiosyncratic cost-shock that is assumed to be Frechet distributed according to

$$G(\nu) = e^{-z^{-(\sigma-1)}}. \tag{B.3}$$

Minimizing cost for a given target level of output, the firm then chooses for each task j a worker type s that can complete this task most effectively:

$$s^*(j) = \arg\min_s \log c\,(s,j,Q)$$

The probability that a given worker type is chosen to complete task j therefore depends directly on firm scale Q if $\varepsilon_s \neq \varepsilon_{s'}$. An alternative interpretation is that the nature of the tasks themselves change as firms grow larger, given that organizational needs and production arrangements need to vary accordingly. The following result confirms that the resulting cost function will be equivalent to the cost function implied by the non-homothetic CES.

PROPOSITION 6 *The mixed random cost task assignment model defined by the cost functions (B.2) and Type-II Extreme Value distributed idiosyncratic cost shocks is isomorphic to a non-homothetic CES model in which firms produce according to a non-homothetic production technology with respect to output and elasticity of substitution σ between worker groups.*

Proof. The proof is a straightforward application of the well-known Frechet algebra. ■

B.2.2 Technological Choice

Non-homotheticities in production are a reduced form way of modeling the effects of an underlying technological choice that a firm may face. To illustrate this point, I show that a non-homothetic CES production function can be derived from such a model. To do so, I assume that firms optimally choose technologies from a menu. Firms trade off the fixed cost of adopting the technology against potential savings in variable costs of production. Technologies might thereby vary in factor intensities and use, for example, different proportions of high skilled workers.

For simplicity, I assume that firms only hire two types of workers of either high or low skill. Firms choose a technology t from a menu of technologies $\mathcal{T}$ that is characterized by costs:

$$\mathcal{T} = \left\{ t \in \mathbb{R}_+ : TC(t,Q) = (fh(t) + Qg(t)) \left(\alpha(t) w_l^{1-\sigma} + (1-\alpha(t)) w_h^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \right\},$$

where f denotes a fixed cost of production. Firms thus trade off fixed cost, as governed by $h(.)$ agains variable cost governed by $g(.)$. Further, $\alpha(t)$ governs the relative technological advantage that high and low skill workers have across different technologies. Conditional on a desired level of output q, firms choose $t^*(q)$ so as to minimize total cost:

$$t^*(Q) = \min_t \{TC(t,Q) : TC \in \mathcal{T})\}. \tag{B.4}$$

Note that for $\sigma = 1$ and $h(t) = 1 + (\eta - 1) 1_{t=1}$, and $g(t) = 1_{t=0} + \frac{1}{\gamma} 1_{t=1}$ this reduces to the production function in Bustos (2011b).

Under the condition that this problem has a unique solution, the optimal choice reduces to a total cost function that is equivalent to the one uniquely associated with the non-homothetic production structure. Given an optimal choice t^*, the cost function can be written as:

$$TC(q,\mathbf{w}) = \underbrace{(fh(t^*(Q)) + Qg(t^*(Q)))}_{\equiv A(q)} \underbrace{\left(\alpha(t^*(Q)) w_l^{1-\sigma} + (1-\alpha(t^*(Q))) w_h^{1-\sigma} \right)^{\frac{1}{1-\sigma}}}_{\equiv \left(\alpha(Q) w_l^{1-\sigma} + (1-\alpha(Q)) w_h^{1-\sigma} \right)^{\frac{1}{1-\sigma}}}$$

Lemma 1 *Assume that $h(.)$,$g(.)$, $\alpha(.) \in C^1$. If (i) $h(t)$ and $g(t)$ are log-concave and (ii) $\alpha(t)$ is log-convex if $\sigma > 1$ and log-concave if $\sigma < 1$, then $\forall Q$, there exists a unique technlogy $t^*(Q)$ that solves the cost-minimization problem in equation (B.4).*

Proof. The first order condition with respect to t implies that the optimal technology choice t is characterized by:

$$0 = \frac{a'(t)}{a(t)} + \frac{1}{1-\sigma} \frac{b'(t)}{b(t)},$$

where $a(t) \equiv fh(t) + Qg(t)$ and $b(t) = \alpha(t) \left(w_l^{1-\sigma} - w_h^{1-\sigma} \right) + w_h^{1-\sigma}$. By the assumptions stated in the lemma, the RHS is monotonously decreasing. By the Intermediate Value Theorem, a unique solution to the firm's optimal technology choice exists. ■

This implies that that cost can be written as:

$$C(w_l, w_h, Q) = A(Q) \left(\alpha(Q) w_l^{1-\sigma} + (1-\alpha(Q)) w_h^{1-\sigma} \right)^{\frac{1}{1-\sigma}}, \tag{B.5}$$

where $A(Q) \equiv fh(t^*(Q)) + gq(t^*(Q))$.

To derive the resulting production function, I use the fact that the cost function is homogeneous of degree 1 in prices and that Shephard's lemma implies that the input demand for type $i \in \{H, L\}$ is given by $L_i = \frac{\partial C(w_l, w_h, Q)}{\partial w_i}$. Using these two facts, the respective labor demands can be written as:

$$L_l = A(Q)\left\{1 - \alpha(Q) + \alpha(Q)\left(\frac{w_h}{w_l}\right)^{1-\sigma}\right\}^{\frac{\sigma}{1-\sigma}}, \tag{B.6}$$

and

$$L_h = A(q)\left\{(1 - \alpha(Q))\left(\frac{w_h}{w_l}\right)^{\sigma-1} + \alpha(Q)\right\}^{\frac{\sigma}{1-\sigma}}. \tag{B.7}$$

Combining both equalities gives:

$$1 = A(Q)\left[\alpha(Q)L_l^{\frac{\sigma-1}{\sigma}} + (1 - \alpha(Q))\, L^{\frac{\sigma-1}{\sigma}}\right] \tag{B.8}$$

(B.8) corresponds to the general functional form of the non-homothetic CES derived in Sato (1975) which in turn generalizes the technology used in the main text.

B.3 Derivations & Proofs

B.3.1 Consumption Problem

All workers i of a given skill type have Cobb-Douglas preferences over consumption and jobs that take the following form:

$$U(i, f) = \log C_i + \log A_f + \frac{1}{\beta}\epsilon_{i,f}, \tag{B.9}$$

where $C_i = \left(\sum_{f\in\mathcal{F}} \varphi_f^{\frac{\eta-1}{\eta}} c_{f,i}^{\frac{\eta-1}{\eta}} + \left(c_i^*\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}$ is a CES consumption aggregator over domestic and foreign varieties.

Workers supply one unit of labor inelastically to the firm that maximizes their utility. This utility maximization problem can be solved in two steps. First, workers maximize utility from consumption, subject to the following budget constraint:

$$\sum_{f\in\mathcal{F}} p_f c_f = W_{s,f}.$$

By standard arguments, this yields the following demand schedule for domestic varieties:

$$c_{f,i} = \varphi^{\eta-1} p_f^{-\eta} P^{\eta-1} W_{s,f}.$$

Aggregated over all domestic individuals, this yields the demand schedules for workers in the main text:

$$D_f(p, \varphi) = \varphi^{\eta-1} p^{-\eta} P^{\eta-1} \left(\sum_s \int_{f \in \mathcal{F}} W_{s,f} l_{s,f} df \right).$$

The indirect utility of an individual worker working form firm f, in turn, can be written as:

$$U(i, f) = \log \left(W_{s,f} / P \right) + \log A_f + \frac{1}{\beta_s} \log \epsilon_{i,f}.$$

In the second step, workers choose the job that maximizes their indirect utility. This is described in the next section.

B.3.2 Logit Choice Probabilities

I derive the probability $\lambda_{s,f}$ that workers of skill type s chooses to work for firm f. The idiosyncratic preferences $\epsilon_{i,f}$ over jobs are independently, identically distributed extreme value with cumulative distribution function:

$$F(x) = e^{-e^{-x}}.$$

Following McFadden (1973), the probability that a worker chooses job f is given by:

$$\begin{array}{rcl} P_{s,f,i} & = & \mathbb{P}\left(\beta_s \log W_{s,f} A_f + \epsilon_{i,f} > \beta_s \log W_{s,f'} A_{f'} + \epsilon_{i,f'}, \forall f' \neq f \right) \\ & = & \mathbb{P}\left(\epsilon_{i,f'} < \epsilon_{i,f} + \beta_s \left(\log W_{s,f} A_f - \log W_{s,f'} A_{f'} \right), \forall f' \neq f \right) \end{array}.$$

Taking $\epsilon_{i,f}$ as given, this expression equals the cumulative distribution function for each $\epsilon_{i,f'}$ evaluated at $\epsilon_{i,f} + \beta_s \left(\log W_{s,f} A_f - \log W_{s,f'} A_{f'} \right)$. Since the preference shocks are independent, this cdf can be calculated for each $\epsilon_{i,f}$:

$$P_{s,f,i} | \epsilon_{i,f} = \prod_{f' \neq f} e^{-e^{-\left(\epsilon_{i,f} + \beta_s \left(\log W_{s,f} A_f - \log W_{s,f'} A_{f'} \right) \right)}}.$$

The choice probability integrates this expression over all possible realizations of $\epsilon_{i,f}$:

$$P_{s,f,i} = \int \left(\prod_{f' \neq f} e^{-e^{-\left(\epsilon_{i,f} + \beta_s \left(\log W_{s,f} A_f - \log W_{s,f'} A_{f'} \right) \right)}} \right) d\epsilon_{i,f}.$$

Calculating this integral gives the choice probability for an individual worker:

$$P_{s,f,i} = \frac{e^{\beta_s \log W_{s,f} A_f}}{\sum_{f'} e^{\beta_s \log W_{s,f'} A_{f'}}} = \frac{\left(W_{s,f} A_f \right)^{\beta_s}}{\sum_{f'} \left(W_{s,f'} A_{f'} \right)^{\beta_s}}.$$

This choice probability is the same for all workers i belonging to type s.

B.3.3 Derivation of the cost function

Throughout the derivation, I suppress amenities A.

The Lagrangean solving the firm's cost minimization problem defined in equation (3.8) is given by:

$$\mathcal{L} = \sum_s W_s\left(l_s\right) l_s + \lambda \left(1 - \sum_s \Omega_s^{\frac{1}{\sigma}} l_s^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_S-\sigma}{\sigma}}\right),$$

where it is understood that $W_s\left(l_s\right)$ is given by the inverse of labor supply given by equation (3.6). The uniqueness of the solution follows from the fact that we are minimizing a strictly convex objective over a convex set, given that the production function is globally quasi-concave in labor inputs.

Taking first-order conditions with respect to l_s and solving for the multipliers on the labor supply equations, we obtain:

$$\frac{\beta_s + 1}{\beta_s} W_s = \lambda \frac{\sigma - 1}{\sigma} \Omega_s^{\frac{1}{\sigma}} l_s^{-\frac{1}{\sigma}} Q^{\frac{\varepsilon_S-\sigma}{\sigma}}.$$

where $\beta_s = \frac{d \log \mathcal{L}_s}{d \log W_s}$ is the elasticity of labor supply. This immediately implies that we can write the wage bill share of skill group s as:

$$\omega_s \equiv \frac{W_s l_s}{\sum_{s'} W_{s'} l_{s'}} = \frac{\frac{\beta_s}{\beta_s+1} \Omega_s^{\frac{1}{\sigma}} l_s^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_S-\sigma}{\sigma}}}{\sum_{s'} \frac{\beta_s'}{\beta_{s'}+1} \Omega_{s'}^{\frac{1}{\sigma}} l_{s'}^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_{s'}-\sigma}{\sigma}}}.$$

By the envelope theorem marginal cost are given by:

$$MC(Q) = -\lambda \frac{1}{Q} \sum_s \Omega_s^{\frac{1}{\sigma}} l_s^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_S-\sigma}{\sigma}} \frac{\varepsilon_S - \sigma}{\sigma}.$$

Solve for the multiplier λ and substitute this term into the first-order condition to obtain the expression for wages given in the main text:

$$W_s = \frac{\beta_s}{\beta_s + 1} MC(Q) \frac{\Omega_s^{\frac{1}{\sigma}} l_s^{-\frac{1}{\sigma}} Q^{\frac{\varepsilon_S-\sigma}{\sigma}}}{\frac{1}{Q} \sum_{s'} \Omega_{s'}^{\frac{1}{\sigma}} l_{s'}^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_{s'}-\sigma}{\sigma}} \frac{\varepsilon_{s'}-\sigma}{1-\sigma}}.$$

To derive marginal cost, multiply both sides of this expression by l_s and sum over all s to obtain total labor cost $C(Q)$:

$$C(Q) \equiv \sum_s W_s l_s = QMC(Q) \frac{\sum_s \frac{\beta_s}{\beta_s+1} \Omega_s^{\frac{1}{\sigma}} l_s^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_s-\sigma}{\sigma}}}{\sum_{s'} \Omega_{s'}^{\frac{1}{\sigma}} l_{s'}^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_{s'}-\sigma}{\sigma}} \frac{\varepsilon_{s'}-\sigma}{1-\sigma}}.$$

$$\sum_{s'} \Omega_{s'}^{\frac{1}{\sigma}} l_{s'}^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_{s'}-\sigma}{\sigma}} \frac{\varepsilon_{s'}-\sigma}{1-\sigma} = \frac{Q}{C(Q)} MC(Q) \left(\sum_s \frac{\beta_s}{\beta_s+1} \Omega_s^{\frac{1}{\sigma}} l_s^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_s-\sigma}{\sigma}} \right)$$

Rearrange to solve for marginal cost:

$$MC(Q) = \frac{C(Q)}{Q} \sum_{s'} \frac{\Omega_{s'}^{\frac{1}{\sigma}} l_{s'}^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_{s'}-\sigma}{\sigma}}}{\sum_s \frac{\beta_s}{\beta_s+1} \Omega_s^{\frac{1}{\sigma}} l_s^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_s-\sigma}{\sigma}}} \frac{\varepsilon_{s'}-\sigma}{1-\sigma}.$$

Using the expression for wage bill shares, note that $\frac{\Omega_{s'}^{\frac{1}{\sigma}} l_{s'}^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_{s'}-\sigma}{\sigma}}}{\sum_s \frac{\beta_s}{\beta_s+1} \Omega_s^{\frac{1}{\sigma}} l_s^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_s-\sigma}{\sigma}}} = \omega_s \frac{\beta_s+1}{\beta_s}$. Plug this back into the above expression to derive the expression for marginal cost given in the main text:

$$MC(Q) = \frac{C(Q)}{Q} \sum_s \omega_s \frac{\beta_s+1}{\beta_s} \frac{\varepsilon_s-\sigma}{1-\sigma}.$$

It is immediately evident that total cost are increasing so long as $\frac{\varepsilon_s-\sigma}{1-\sigma} > 0$. Lastly, note that using the first order condition for wages and the constraint posed by the production technology, the multiplier λ equals:

$$\sum_s \left(W_s l_s \frac{\beta_s+1}{\beta_s} \right) = \lambda \frac{\sigma-1}{\sigma}.$$

Therefore, an alternative expression for wages that will be handy in the next proof is given by:

$$W_s = \frac{\beta_s}{\beta_s+1} \left\{ \sum_s W_s l_s \frac{\beta_s+1}{\beta_s} \right\} \Omega_s^{\frac{1}{\sigma}} l_s^{-\frac{1}{\sigma}} Q^{\frac{\varepsilon_s-\sigma}{\sigma}}. \tag{B.10}$$

B.3.4 Convexity of Marginal Cost

PROPOSITION 7 *Suppose* $\frac{\varepsilon_s-\sigma}{1-\sigma} \frac{\beta_s+1}{\beta_s} \geq 1$. *If* $\sigma < 1$, *then (variable) cost are convex. If* $\sigma > 1$,*then cost are convex if there exists a worker type* $s \in \mathcal{S}$ *such that for all* $s' \neq s$, $\varepsilon_s > \varepsilon_{s'}$ *and* $\frac{\varepsilon_s-\sigma}{1-\sigma} \frac{\beta_s+1}{\beta_s} \geq \frac{\varepsilon_{s'}-\sigma}{1-\sigma} \frac{\beta_{s'}+1}{\beta_{s'}}$.[33]

Proof. Define $\tilde{\omega}_s \equiv \frac{\omega_s \frac{\beta_s+1}{\beta_s} \frac{\varepsilon_s-\sigma}{\sigma-1}}{\sum_s \omega_s \frac{\beta_s+1}{\beta_s} \frac{\varepsilon_s-\sigma}{\sigma-1}}$. Defining the wage bill share of worker group s by ω_s,

[33] If $\sigma > 1$, a sufficient condition for locally convex marginal cost is given by: $\forall s$, $\frac{\beta_s+\sigma}{\sigma(1+\beta_s)} > \omega_s$, where ω_s denotes the wage-bill share of worker type s.

the elasticity of cost with respect to output equals: $\frac{\partial \log C(Q)}{\partial \log Q} = \sum_s \omega_s (1+\beta_s) \frac{\partial \log W_s}{\partial \log Q} = \sum_s \omega_s \frac{\beta_s+1}{\beta_s} \frac{\varepsilon_s-\sigma}{\sigma-1}$. The first equality follows from B.10. The super elasticity of total cost with respect to output is given by

$$\frac{\partial^2 \log C(Q)}{\partial \log Q^2} = \sum \tilde{\omega}_s \left(\frac{\partial \log W_s l_s}{\partial \log Q} - \frac{\partial \log C(Q)}{\partial \log Q} \right) = \sum_s \tilde{\omega}_s \left\{ (1+\beta_s) \frac{\partial \log W_s}{\partial \log Q} - \sum \omega_s (1+\beta_s) \frac{\partial \log W_s}{\partial \log Q} \right\}$$

Assumption 1 implies that $\sum_s \omega_s \frac{\beta_s+1}{\beta_s} \frac{\varepsilon_s-\sigma}{1-\sigma} > 0$ and thus rearranging the above expression gives a sufficient condition for a strictly increasing super elasticity is given by:

$$\sum_s \omega_s (1+\beta_s) \frac{\partial \log W_s}{\partial \log Q} \frac{\beta_s+1}{\beta_s} \frac{\varepsilon_s-\sigma}{1-\sigma} - \left(\sum \omega_s (1+\beta_s) \frac{\partial \log W_s}{\partial \log Q} \right)^2 > 0$$

From the two expressions for the cost elasticity, this condition be restated as:

$$\sum_s \omega_s (1+\beta_s) \frac{\partial \log W_s}{\partial \log Q} \frac{\beta_s+1}{\beta_s} \frac{\varepsilon_s-\sigma}{1-\sigma} - \left(\sum_s \omega_s (1+\beta_s) \frac{\partial \log W_s}{\partial \log Q} \right) \left(\sum_s \omega_s \frac{\beta_s+1}{\beta_s} \frac{\varepsilon_s-\sigma}{1-\sigma} \right) > 0 \quad \text{(B.11)}$$

This condition can be understood as a restriction on the covariance between labor supply elasticity adjusted wage growth and increases in cost:

$$\mathrm{Cov}_{\omega_s} \left\{ (1+\beta_s) \frac{\partial \log W_s}{\partial \log Q}, \frac{1+\beta_s}{\beta_s} \frac{\varepsilon_s-\sigma}{1-\sigma} \right\} \geq 0. \quad \text{(B.12)}$$

This condition will be satisfied if ε_s and β_s are sufficiently negatively correlated - that is if workers that are more productive are also on average harder to hire.

Finally, I show that if the parametric conditions in the statement of the proposition are satisfied, marginal costs are increasing. To show this, consider the elasticity of marginal cost with respect to output:

$$\frac{\partial \log MC(Q)}{\partial \log Q} = \frac{\partial \log C(Q)}{\partial \log Q} - 1 + \frac{\partial^2 \log C}{\partial \log Q^2}.$$

If $\sigma < 1$, it is easy to see that the condition in equation (B.12) is satisfied and the last term is thus greater than 0. A sufficient condition for marginal cost to be increasing is then $\frac{\beta_s+1}{\beta_s} \frac{\varepsilon_s-\sigma}{1-\sigma} \geq 1$, as stated in the first part of the proposition.

If $\sigma > 1$, use the previous results to show that the elasticity of marginal cost can be written as:

$$\frac{\partial \log MC(Q)}{\partial \log Q} = -1 + \frac{\sum \omega_s \frac{1+\beta_s}{\beta_s} \frac{\varepsilon_s-\sigma}{1-\sigma} \left\{ (1+\beta_s) \frac{\partial \log W_s}{\partial \log Q} \right\}}{\sum \omega_s \left\{ \frac{\beta_s+1}{\beta_s} \frac{\varepsilon_s-\sigma}{1-\sigma} \right\}}.$$

A sufficient condition for increasing marginal cost is given by:

$$\sum \omega_s \left\{ \frac{1+\beta_s}{\beta_s} \frac{\varepsilon_s - \sigma}{1-\sigma} (1+\beta_s) \frac{\partial \log W_s}{\partial \log Q} \right\} > \sum \omega_s \left\{ \frac{\beta_s + 1}{\beta_s} \frac{\varepsilon_s - \sigma}{1-\sigma} \right\}. \tag{B.13}$$

The elasticity of wages equals:

$$\frac{\partial \log W_s}{\partial \log Q} = \frac{\sigma}{\beta_s + \sigma} \left\{ \underbrace{\sum_k \left(\frac{W_k l_k \frac{1}{\beta_k}}{\sum_l W_l l_l \frac{1}{\beta_l}} \right)}_{\hat{\omega}_s} (\beta_k + 1) \frac{\partial \log W_k}{\partial \log Q} - 1 + \sum_s \omega_s \frac{\beta_s + 1}{\beta_s} \frac{\varepsilon_s - \sigma}{1-\sigma} + \frac{\varepsilon_s}{\sigma} \right\}.$$

Let $a_s \equiv \frac{\partial \log W_s}{\partial \log Q}$. The wage elasticities then define a system of equations which we may write as:

$$Aa = \varepsilon,$$

where the matrix A is given by:

$$A = \begin{bmatrix} \frac{\beta_1+\sigma}{\sigma} - \hat{\omega}_1 (1+\beta_1) & -(\beta_2+1)\tilde{\omega}_2 & \dots & -(\beta_S+1)\tilde{\omega}_S \\ & \frac{\beta_2+\sigma}{\sigma} - \hat{\omega}_2(1+\beta_2) & & \\ & & \dots & \\ -(\beta_1+1)\tilde{\omega}_1 & & & \frac{\beta_S+\sigma}{\sigma} - \hat{\omega}_S(1+\beta_S) \end{bmatrix}.$$

If $\frac{\beta_s+\sigma}{\sigma} - \hat{\omega}_s (1+\beta_s) > 0$ for all worker types, then Gershgorin's circle theorem can be applied to show that A is a M-matrix, implying that $a >> 0$. If $\frac{\partial \log W_s}{\partial \log Q} \geq 0$ for all worker types, then equation (B.13) obviously holds under the parametric restrictions stated in the proposition. For $\sigma < 1$, we therefore have that marginal cost are increasing globally under the conditions stated in the proposition.

For $\sigma > 1$, first observe that marginal cost are positive and further equal $\sum \omega_s (1+\beta_s) \frac{\partial \log W_s}{\partial \log Q}$. Given that marginal cost are positive, there therefore exists at least one s such that $\frac{\partial \log W_s}{\partial \log Q} > 0$. Wlog, let $s = 1$ be that worker group. Then we know that since marginal cost are positive, that $\omega_1 (1+\beta_1) \frac{\partial \log W_1}{\partial \log Q} > -\sum_{k \neq 1} \omega_k (1+\beta_k) \frac{\partial \log W_k}{\partial \log Q}$. Use the relative growth in wages to re-arrange this term as:

$$\omega_s (1+\beta_s) \frac{\partial \log W_1}{\partial \log Q} > -\sum_{k \neq s} \omega_k (1+\beta_k) \left(\frac{\beta_1 + \sigma}{\beta_k + \sigma} \frac{\partial \log W_1}{\partial \log Q} + \frac{\varepsilon_s - \varepsilon_1}{\beta_k + \sigma} \right).$$

The increase in the wage by worker group 1 can be bound as:

$$\frac{\partial \log W_1}{\partial \log Q}\left\{\omega_1\left(1+\beta_1\right)+\sum_{k\neq s}\omega_k\frac{1+\beta_k}{\beta_k+\sigma}\left(\beta_1+\sigma\right)\right\}>\sum_{k\neq s}\omega_k\frac{1+\beta_k}{\beta_k+\sigma}\left(\varepsilon_1-\varepsilon_k\right)$$

Next, turning to equation (B.13), we can apply similar steps to isolate $\frac{\partial \log W_1}{\partial \log Q}$ and rewrite the condition for increasing marginal cost as the following inequality:

$$\frac{\partial \log W_1}{\partial \log Q}\left\{\omega_1\frac{1+\beta_1}{\beta_1}\frac{\varepsilon_1-\sigma}{1-\sigma}\left(1+\beta_1\right)+\sum_{k\neq s}\omega_k\frac{1+\beta_k}{\beta_k}\frac{\varepsilon_k-\sigma}{1-\sigma}\frac{1+\beta_k}{\beta_k+\sigma}\left(\beta_1+\sigma\right)\right\}+\sum_{k\neq s}\omega_s\frac{1+\beta_k}{\beta_k}\frac{\varepsilon_k-\sigma}{1-\sigma}\frac{1+\beta_k}{\beta_k+\sigma}\left(\varepsilon_k-\varepsilon_1\right)$$

$$>\frac{\partial \log W_1}{\partial \log Q}\left\{\omega_1\left(1+\beta_1\right)+\sum_{k\neq s}\omega_k\frac{1+\beta_k}{\beta_k+\sigma}\left(\beta_1+\sigma\right)\right\}+\sum_{k\neq s}\omega_k\frac{1+\beta_k}{\beta_k+\sigma}\left(\varepsilon_k-\varepsilon_1\right)$$

$$\Leftrightarrow\frac{\partial \log W_1}{\partial \log Q}\omega_1\left(1+\beta_1\right)>-\sum_{k\neq s}\omega_k\left(1+\beta_k\right)\left(\frac{\beta_1+\sigma}{\beta_k+\sigma}\frac{\partial \log W_1}{\partial \log Q}+\frac{\varepsilon_s-\varepsilon_1}{\beta_k+\sigma}\right)\frac{\left(\frac{1+\beta_k}{\beta_k}\frac{\varepsilon_k-\sigma}{1-\sigma}-1\right)}{\left\{\frac{1+\beta_1}{\beta_1}\frac{\varepsilon_1-\sigma}{1-\sigma}-1\right\}}$$

$$\Leftrightarrow\frac{\partial \log W_1}{\partial \log Q}\left\{\omega_1\left(1+\beta_1\right)+\sum_{k\neq s}\omega_k\frac{\left(1+\beta_k\right)}{\beta_k+\sigma}\left(\beta_1+\sigma\right)\frac{\left(\frac{1+\beta_k}{\beta_k}\frac{\varepsilon_k-\sigma}{1-\sigma}-1\right)}{\left\{\frac{1+\beta_1}{\beta_1}\frac{\varepsilon_1-\sigma}{1-\sigma}-1\right\}}\right\}>\sum_{k\neq s}\omega_k\frac{\left(1+\beta_k\right)}{\beta_k+\sigma}\left(\varepsilon_1-\varepsilon_s\right)\frac{\left(\frac{1+\beta_k}{\beta_k}\frac{\varepsilon_k-\sigma}{1-\sigma}-1\right)}{\left\{\frac{1+\beta_1}{\beta_1}\frac{\varepsilon_1-\sigma}{1-\sigma}-1\right\}}$$

There exist two possibilities. Either, the term on the right-hand side is negative, in which case the inequality holds as $\frac{1+\beta_k}{\beta_k}\frac{\varepsilon_k-\sigma}{1-\sigma}>1$. If the terms on the RHS are negative, in other words if wages were decreasing for all other worker groups, then given that the proposition implies that $\frac{\left(\frac{1+\beta_k}{\beta_k}\frac{\varepsilon_k-\sigma}{1-\sigma}-1\right)}{\left(\frac{1+\beta_1}{\beta_1}\frac{\varepsilon_1-\sigma}{1-\sigma}-1\right)}<1$ and given that we can bound the wage growth by worker group 1, we have that the inequality must hold and therefore have shown a sufficient condition for increasing marginal cost if $\sigma>1$.

■

B.3.5 Proof of Proposition 3

Proof. To show concavity of the profit function, consider the second derivative of operating profits with respect to output Q :

$$\frac{\partial^2\pi\left(\varphi,A\right)}{\partial Q}=-\left(\frac{\eta-1}{\eta^2}\right)Q^{-1/\eta-1}\left(\left(\varphi Y\right)^{\frac{1}{\eta}}P^{\left(\eta-1\right)/\eta}+\left(\varphi Y^*\right)^{\frac{1}{\eta}}\right)-\frac{\partial MC\left(Q,A\right)}{\partial Q}<0,$$

since $\eta>1$ and marginal cost are convex.

Note also that the profit function is given by:

$$\pi\left(\varphi,A\right)=\varphi\left(\frac{\eta}{\eta-1}MC\left(Q,A\right)\right)^{1-\eta}\left(YP^{\eta-1}+1_X\tau^{1-\eta}Y^*\right)-C\left(Q,A\right)-1_XF_X.$$

Cost and marginal cost are not directly affected by φ, and thus $\pi(\varphi, A)$ is increasing in φ. Marginal cost and cost are decreasing in amenities A, implying that $\pi(\varphi, A)$ is increasing in A.

If there is no international trade, that is if $P^* = Y^* = 0$, the fixed point theorems in Kennan (2001) apply and imply a fixed point exists. Consider now the case where $P^* > 0, \quad Y^* > 0$. Using the CES demand structure, the fixed point can be written as a system of equations in quantities, prices and the aggregate price index.

$$f_1(\mathbf{p}; \mathbf{Q}, P) = P - \left(\int_{f \in \mathcal{F}} \varphi p_f^{1-\eta} df + (P^*)^{1-\eta}\right)^{\frac{1}{1-\eta}}$$

$$f_k(\mathbf{p}; \mathbf{Q}, P) = p_k - \left(Q_{H,k}^{-1/\eta} + (\tau Q_{F,k})^{-1/\eta}\right) \varphi^{1/\eta} \left(YP^{\eta-1} + Y^*\right)^{1/\eta}.$$

$f = \left(f_1, \{f_k\}_{k \in 2,\dots,|\mathcal{F}|+1}\right)$ is strictly monotone in all its arguments. Given that $P^*, Y^* > 0$, $f(\mathbf{0}) < 0$ and $\lim_{\lambda \to \infty} f(\lambda \mathbf{x}) > 0$. The intermediate value theorem thus implies the existence of a fixed point. To prove uniqueness, I show that gross substitution holds. $\frac{\partial f_1^2}{\partial x \partial y} = 0$, for all $x \neq y$. Also, $\frac{\partial^2 f_k}{\partial p_k \partial Q} = \frac{\partial^2 f_k}{\partial p_k \partial P} = 0$. Lastly, $\frac{\partial^2 f_k}{\partial Q \partial P} = \frac{\eta-1}{\eta} P^{\eta-2} \frac{1}{\eta} Q_{i,k}^{-1/\eta-1} > 0$ for $i \in \{H, F\}$. Therefore, gross substitution holds and standard arguments imply that a unique fixed point exists.

■

B.3.6 Proof of Proposition 4

Proof. The relative change in wages in response to a change in trade cost can be expressed as:

$$\frac{d \log W_s(\varphi)}{d \log \tau} - \frac{d \log W_{s'}(\varphi)}{d \log \tau} = \left\{-\frac{1}{\sigma}\left(\frac{d \log l_s(\varphi)}{d \log \tau} - \frac{d \log l_{s'}(\varphi)}{d \log \tau}\right) + \left(\frac{\varepsilon_s - \varepsilon_{s'}}{\sigma}\right)\frac{d \log Q(\varphi)}{d \log \tau}\right\} \tag{B.14}$$

To show the first statement, assume that $\beta_s = \beta$ and $\varepsilon_s = \varepsilon$. In this case, the relative change in wages for two skill groups s and s' can be written as:

$$\frac{d \log W_s(\varphi)}{d \log \tau} - \frac{d \log W_{s'}(\varphi)}{d \log \tau} = -\frac{1}{\sigma + \beta}\left(\frac{d \log \Lambda_s}{d \log \tau} - \frac{d \log \Lambda_{s'}}{d \log \tau}\right).$$

Therefore within firm wage inequality changes proportionally across all firms, independently of any heterogeneous equilibrium effects that a change in market size across firms.

To show the second statement, assume that labor markets are equally competitive, that is $\beta_s = \beta$. Using the expression for labor supply, we can then express the relative passthrough of a trade shock as:

$$\frac{d\log W_s(\varphi)}{d\log\tau} - \frac{d\log W_{s'}(\varphi)}{d\log\tau} = \frac{1}{\sigma+\beta}\left(\underbrace{-\left(\frac{d\log\Lambda_s}{d\log\tau} - \frac{d\log\Lambda_{s'}}{d\log\tau}\right)}_{\text{Common to all firms}} + \underbrace{(\varepsilon_s - \varepsilon_{s'})\frac{d\log Q(\varphi)}{d\log\tau}}_{\text{Output Complementarity}}\right) \tag{B.15}$$

For firms that see an increase in market size, $\frac{d\log Q}{d\log\tau} > 0$, employees with higher technological scale complementary will therefore see larger relative wage gains.

To show the third statement, consider now the case where $\varepsilon_s = \varepsilon_{s'}$. Heterogeneous changes in firm level wage premia across firms now reflect differences in labor supply elasticities:

$$\frac{d\log W_s(\varphi)}{d\log\tau} - \frac{d\log W_{s'}(\varphi)}{d\log\tau} = -\frac{1}{\sigma}\left(\beta_s\frac{d\log W_s}{d\log Q} - \beta_{s'}\frac{d\log W_{s'}}{d\log Q}\right)\frac{d\log Q(\varphi)}{d\log\tau} - \frac{1}{\sigma}\left(\frac{d\log\Lambda_s}{d\log\tau} - \frac{d\log\Lambda_{s'}}{d\log\tau}\right) \tag{B.16}$$

As is evident in this expression, lower labor supply elasticity β_s in labor markets will put larger upward pressure on wages at firms that grow in scale in response to a change in trade cost.

■

B.3.7 Proof of Proposition 5

Proof. I first characterize the Walrasian equilibrium allocation under competitive labor markets. In a competitive equilibrium, firms act as price takers and wages equal the marginal revenue product of labor, that is

$$W_{s,f} = MRPL_{s,f}. \tag{B.17}$$

In the following, I drop the dependency of cost on amenities from the algebraic expressions. Following the steps from the previous derivation of the cost function, marginal cost in this case are given by:

$$MC(Q) = \frac{C}{Q}\sum_s \omega_s\left(\frac{\varepsilon_s - \sigma}{1-\sigma}\right), \tag{B.18}$$

where ω_s denotes the market share of worker group s in the competitive allocation. Showing that the proposed tax policy achieves the same worker allocation with competitive labor markets is equivalent to showing that labor demand and cost correspond to equation (B.17) and equation (B.18) respectively.

Under the proposed tax reform, firms pay a fraction $1-\tau_s$ of the wage bill while the government subsidizes the remainder. The first-order condition for a firm's optimal employment of worker type s then becomes:

$$(1-\tau_s)\, W_{s,f} = \frac{\beta_s}{\beta_s + 1} MRPL_{s,f}.$$

If $\tau_s = \frac{1}{1+\beta_s}$, then labor demand is equal to equation (B.17).

I now show that marginal cost are equivalent to those of a firm that sets wages equal to the marginal revenue product of labor. Using the first-order condition, the gross wage-bill share is given by $\omega_s = \frac{\Omega_s^{\frac{1}{\sigma}} l_s^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_S-\sigma}{\sigma}}}{\sum_{s'} \Omega_{s'}^{\frac{1}{\sigma}} l_{s'}^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_{S'}-\sigma}{\sigma}}}$. Marginal cost are by the envelope theorem given by:

$$MC(Q) = -\lambda \frac{1}{Q} \sum_s \Omega_s^{\frac{1}{\sigma}} l_s^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_S-\sigma}{\sigma}} \frac{\varepsilon_S - \sigma}{\sigma}. \tag{B.19}$$

Wages can be expressed as: $W_s = MC(Q) \frac{\Omega_s^{\frac{1}{\sigma}} l_s^{-\frac{1}{\sigma}} Q^{\frac{\varepsilon_S-\sigma}{\sigma}}}{\frac{1}{Q}\sum_{s'} \Omega_{s'}^{\frac{1}{\sigma}} l_{s'}^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_{S'}-\sigma}{\sigma}} \frac{\varepsilon_{S'}-\sigma}{1-\sigma}}$ and total cost are given by $C(Q) \equiv \sum_s (1-\tau_s) W_s l_s = QMC(Q) \frac{\sum_s (1-\tau_s) \Omega_s^{\frac{1}{\sigma}} l_s^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_S-\sigma}{\sigma}}}{\sum_{s'} \Omega_{s'}^{\frac{1}{\sigma}} l_{s'}^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_{S'}-\sigma}{\sigma}} \frac{\varepsilon_{S'}-\sigma}{1-\sigma}}$. Using the definition of ω_s from above, this can be rearranged to:

$$\frac{\sum_s (1-\tau_s) W_s l_s}{Q} \sum \omega_s \frac{\varepsilon_{s'} - \sigma}{1-\sigma} = MC(Q) \left(1 - \sum_s \omega_s \tau_s\right). \tag{B.20}$$

Noting that $(1 - \sum_s \omega_s \tau_s) = \sum_s \omega_s (1-\tau_s) = \frac{\sum_s W_s l_s (1-\tau_s)}{\sum_s W_s l_s}$, marginal cost are therefore given by $MC(Q) = \frac{\sum_s W_s l_s}{Q} \sum_s \omega_s \left(\frac{\varepsilon_s - \sigma}{1-\sigma}\right)$.

Consequently, marginal cost are equivalent to the competitive case that is characterized by equation (B.18). Finally, the labor allocation is not distorted, given that non-progressive income taxes enter log-additively into workers' optimal choice of employment. As a result, the allocation of workers to firms and output in the equilibrium with labor market frictions equals that of the equilibrium with competitive labor markets.

■

B.4 Theoretical Extensions

B.4.1 Upward Sloping Labor Supply Curves in a Model of Directed Search

While the model presented in the main text microfounds upward-sloping labor supply curves by means of idiosyncratic employment preference shocks to individual workers, another microfoundation is given by standard search models, as in Burdett & Mortensen (1998).

Assume as in the main text that labor markets are segregated by worker type s and that the mass of workers is given by L_s. Workers have linear indirect utility in income. Unemployed workers search for employment, receiving wage offers at Poisson rate λ_s. Workers can choose to accept a wage offer, in which case they work for a given firm until the match is terminated, which happens at exogenous Poisson rate χ_s.

Such a setting gives rise to upward sloping labor supply curves. The following proposition shows that the labor supply elasticity, albeit not constant, is closely related to the elasticity in the main text if the underlying source of heterogeneity across firms is drawn from a Pareto distribution.

Proposition 8 *Assume that the underlying firm productivity distribution is Pareto with shape parameter p. Denoting labor supply by $\mathcal{L}_s\left(W_{s,f}\right)$,we have that the elasticity of labor supply to an individual firm f with respect to wages is given by*

$$\frac{d\log\mathcal{L}_s\left(W_{s,f}\right)}{d\log W_{s,f}} = 2p\frac{\lambda_s/\chi_s\left(1-F_s\left(W_{s,f}\right)\right)}{1+\lambda_s/\chi_s\left(1-F_s\left(W_{s,f}\right)\right)},$$

where $F(.)$ is the equilibrium distribution of wages.

Proof. In this set-up, the steady state mass of workers $\mathcal{L}_s$ that accept jobs at a firm posting wages $W_{s,f}$ can be shown to equal:

$$\mathcal{L}_s\left(W_{s,f}\right) = \frac{\lambda_s/\chi_s}{L_s\left[1+\lambda_s/\chi_s\left(1-F_s\left(W_{s,f}\right)\right)\right]^2}, \tag{B.21}$$

where $F_s\left(W_{s,\varphi}\right)$ is the share of firms that offers wages less than $W_{s,f}$ in equilibrium. The labor supply function in equation (B.21) thus displays similar behavior to the expression for labor supply presented in the main text. In particular, firms that offer higher wages attract more workers and upward shifts in the wages offered by all other firms decrease the labor supply to firm f. Further, the distribution of wages is a sufficient statistic for labor supply to the firm.

Ignoring issues of differentiability, the elasticity of a firm's labor supply takes the following form:

$$\frac{d\log\mathcal{L}_s}{d\log W_{s,f}} = \frac{2W_{s,f}\left(\lambda_s/\chi_s\frac{\partial F_s\left(W_{s,f}\right)}{\partial W}\right)}{\left(1+\lambda_s/\chi_s\left(1-F_s\left(W_{s,f}\right)\right)\right)}.$$

If firms are only heterogeneous in productivity, then the wage offer distribution inherits the shape of the underlying productivity distribution, as wages are monotonous in productivity. If productivity is drawn from the Pareto distribution, that is $F_s\left(W_{s,f}\right) = 1-\left(\frac{a}{\varphi}\right)^p$, then in equilibrium, the above expression simplifies to:

$$\frac{d\log\mathcal{L}_s}{d\log W_{s,\varphi}} = 2\frac{\lambda_s/\chi_s p\left(\frac{a}{\varphi}\right)^p}{1+\lambda_s/\chi_s\left(\frac{a}{\varphi}\right)^p} = 2p\frac{\lambda_s/\chi_s p\left(1-F_s\left(W_{s,f}\right)\right)}{1+\lambda_s/\chi_s\left(1-F_s\left(W_{s,f}\right)\right)}.$$

■

B.4.2 Heterogeneity within Worker Types and Individual Worker Fixed Effects in Structural Wages

The baseline model assumes that workers of a given type s are homogeneous inputs in production. As a consequence, wages in the baseline model are pinned down at the level of worker types and firms. In this section, I show that the framework readily extends to allow for variation in wages at the level of individual workers and firms, while still preserving the key implications displayed in the main text.

To this end, I allow for exogenous differences in labor efficiency units - or ability - across workers i, denoted e_i. Firms compensate workers for labor efficiency units.

Firms production technologies are as described in Section 3, however labor inputs are now aggregates of efficiency units. Denoting by $\mathcal{I}_{s,f}$ the set of workers of type s employed at firm φ, labor inputs are now given by:

$$l_{s,f} = \int_{i\in\mathcal{I}_{s,f}} e_i di. \tag{B.22}$$

Firms' wages reflect efficiency wages $w_{i,s,f} = e_i W_{s,f}$ that closely resemble the wages in the main text, adjusted for efficiency units.

The total number of workers of worker type s is given by L_s, while the (type specific) distribution of efficiency units is denoted by $F_s(e)$. Workers still receive idiosyncratic preference shocks and choose their optimal employer, subject to being eligible to their rejection cutoffs. As a result, the total number of efficiency units supplied to a firm posting a wage $w_{s,\varphi}$ and non-wage amenities $a_{s,\varphi}$ is now given by:

$$l_{s,f} = L_s a_{s,f} W_{s,f}^{\beta_s} \int_{\underline{e}_s}^{\infty} e^{\beta_s}\Delta(e)_s^{-1} dF_s(e). \tag{B.23}$$

and $\Delta_s(e) = \int_{\varphi'} a_{s,\varphi}\left(eW_{s,f}\right)^{\beta_s}$ is a measure of aggregate demand for worker type s that is similar to the one given in the main text.

To pin down wages, we can solve for the cost minimization problem of the firm, as in the main text. In particular, the problem can now be written:

$$TC(Q) = \min_{w_{s,\varphi}, \underline{e}_{s,\varphi}} \sum_s W_{s,f}^{1+\beta_s} L_s a_{s,\varphi} \tilde{\Delta}^{-1}.$$

subject to $1 = \sum_s \Omega_s^{1/\sigma} l_s^{\frac{\sigma-1}{\sigma}} Q^{\varepsilon_s/\sigma - 1}$ and $\tilde{\Delta}_s = \int_{\underline{e}_s}^{\infty} e^{\beta_s} \Delta_s^{-1}(e) dF_s(e)$.

Following similar steps as in Appendix B, it is simple proof the following proposition, which generalizes the results in Proposition 2 to the current setting:

PROPOSITION 9 *1. Structural wages are given by the following expression:*

$$\log \tilde{w}_{i,s,\varphi} \equiv \log e_i w_{s,\varphi} = \chi_s + \eta_{\beta_s} \psi_i + \eta_{\varepsilon_s} \lambda_{\varphi,1} + \eta_{\beta_s} \lambda_{\varphi,1} + \nu_{\varphi,s}. \tag{B.24}$$

2. *If $\beta_s = \beta$ and $\varepsilon_s = \varepsilon$, wages are log-additive in a worker fixed effect, worker type fixed effect and firm fixed effect:*

$$\log \tilde{w}_{i,s,\varphi} = \chi_s + \psi_i + \lambda_\varphi + \nu_{\varphi,s}. \tag{B.25}$$

Proof. The first order condition pinning down wages $w_{s,\varphi}$ is given by:

$$W_{s,f} = \frac{\beta_s}{\beta_s + 1} \frac{\sigma - 1}{\sigma} \Omega_s^{\frac{1}{\sigma}} l_s^{-\frac{1}{\sigma}} Q^{\frac{\varepsilon_s - \sigma}{\sigma}} \lambda.$$

The earnings of employee i with efficiency unit e_i are then given by

$$w_{i,s,f} = \frac{\beta_s}{\beta_s + 1} e_i \frac{\sigma - 1}{\sigma} l_s^{-\frac{1}{\sigma}} Q^{\frac{\varepsilon_s - \sigma}{\sigma}} \lambda.$$

Solving for l_s and rearranging yields the following expression for efficiency wages:

$$\log \tilde{w}_{i,s,f} = \frac{\sigma}{\sigma + \beta_s} \log \left(\Omega_s \frac{\beta_s}{\beta_s + 1} \left(L_s A_{s,f} \tilde{\Delta}^{-1} \right)^{-\frac{1}{\sigma}} \right) + \frac{\sigma}{\sigma + \beta_s} \log e_i + \frac{\varepsilon_s - \sigma}{\sigma + \beta_s} \log Q_\varphi + \frac{\sigma}{\sigma + \beta_s} \log \lambda_f.$$

∎

As evident from the first part of the proposition, wages are not log-additive in worker and firm types. Independently of a worker's inherent ability, the effect of working for a better employer depends on her type s, as in the baseline version of the model. The proposition shows that the model, under the same conditions as stated in Proposition 2, rationalizes a reduced form model for wages that is log-additive in individual worker and firm type fixed effects.

B.4.3 Multilateral Wage Bargaining and Convex Vacancy Posting Cost

The model in the main text employs a simple model of monopsonistic competition in labor markets to generate wage differences for similar workers across firms. Here, I show that the key implications of the model for wage outcomes are robust to alternative assumptions on the structure of labor markets and are not dependent on the type of competition that is assumed in product markets. In particular, I show that a model of multilateral wage bargaining between firms and workers expressions for wages that inherit a log-additive structure with multiplicative firm and worker type effects.

Assume that labor markets are segregated by worker type $s \in \mathcal{S}$. Workers are assumed to be randomly matched with firms. Upon matching, firms multilaterally bargain with all worker types over wages. Let β_s denote the bargaining weight of worker group s,while the outside option of a worker of type s is equal to $\overline{W}_s$. The cost of posting V vacancies equals V^{γ_s} per unit of time, the (equilibrium) probability of being matched with workers - taken as given by firms - by q_s and the job desctruction rate by χ_s.

As shown in Stole & Zwiebel (1996), wages are pinned down by a set of differential equations. Supressing firm subscripts, wages $\{W_s\}_{s\in\mathcal{S}}$ solve:

$$W_s\left(\mathbf{L}\right) \;\; = \left(1-\beta_s\right)\overline{W}_s + \beta_s\left(MRPL_s - \textstyle\sum_{s'} l_s \frac{\partial W_s}{\partial l_s}\right) \quad , \quad \forall s \in \mathcal{S}. \tag{B.26}$$

where $\mathbf{L} \equiv [L_s]_{s\in\mathcal{S}}$ and $MRPL_s$ is the marginal revenue product of workers of type s. The solution to this set of differential equations can be dervived from minimally adjusting the results in Cahuc *et al.* (2008).

The following proposition shows that this alternative model gives rise to near equivalent structural expressions for wages. It also highlights that labor supply elasticities are closely related to inverse measures of the convexity of iso-elastic vacancy cost, as shown for simple cases in Manning (2011).

Proposition 10 *Assume that outside options are equal to zero for all worker types, that is $\forall s$, $\overline{W}_s = 0$. A model with multilateral bargaining, random search and convex vacancy posting cost then implies the following structural expression for wages:*

$$\begin{aligned} \log W_{s,f} &= \alpha_s + \tfrac{\varepsilon_s}{\sigma}\left(1-\tfrac{\sigma}{\tilde{\gamma}_s+\sigma}\right)\log Q_f + \left(1-\tfrac{\sigma}{\tilde{\gamma}_s+\sigma}\right)\log\left(P_f\int_0^1 \tfrac{1}{\beta_s} z^{\frac{1-\beta_s}{\beta_s}-\frac{1}{\sigma}} G_s\left(\mathbf{L}A_i\left(z\right)\right)dz\right), \\ &= \alpha_s + \eta_{\varepsilon_s}\psi_{1,f} + \eta_{\gamma_s}\psi_{2,f} \end{aligned} \tag{B.27}$$

where $\tilde{\gamma}_s \equiv \frac{1}{\gamma_s} - 1$, $\alpha_s = \frac{\sigma}{\tilde{\gamma}_s+\sigma}\log\frac{1-\beta_s}{\iota_s}$ and $\iota_s \equiv \frac{\chi_s}{\gamma_s q_s}$. $A_i(z)$ is a $S\times S$ diagonal matrix with entries $a_{j,i} = z^{\frac{\beta_j}{1-\beta_j}\frac{1-\beta_i}{\beta_i}}$ and P_φ is the equilibrium price that firm φ charges in equilibrium. G_s denotes the function $G_s \equiv \frac{\frac{\sigma-1}{\sigma}\Omega_s^{\frac{1}{\sigma}}}{\sum\Omega_s^{\frac{1}{\sigma}} l_s^{\frac{\sigma-1}{\sigma}} Q^{\frac{\varepsilon_s-\sigma}{\sigma}\frac{\varepsilon_s-\sigma}{1-\sigma}}}$.

Proof. I first proof the following auxillary lemma.

LEMMA 2 *Under non-homothetic technologies given in equation (3.1), wages in the multilateral bargaining model are given by:*

$$W_s = (1-\beta_s)\bar{W}_g + PQ^{\frac{\varepsilon_s}{\sigma}} l_s^{-\frac{1}{\sigma}} \int_0^1 z^{\frac{1-\beta_s}{\beta_s}-\frac{1}{\sigma}} G_s(\mathbf{L}A_i(z))\, dz, \tag{B.28}$$

,

where $A_i(z)$ is a $S \times S$ diagonal matrix with entries $a_{j,i} = z^{\frac{\beta_j}{1-\beta_j}\frac{1-\beta_i}{\beta_i}}$ and P is the price that firm f charges.

Proof. The result is a direct consequence of the result in given in appendix B.4. in Cahuc *et al.* (2008), generalized to allow for non-competitive input markets. Denote by F the implicit production function defined by the non-homothetic CES. Imposing that the $p'th$ derivative of $l_s^{(p)} \times MRPl_s^{(p)} = P \times l_s^{(p)} \frac{\partial F}{\partial l_s}$ in continuous at zero and that $\lim_{N\to 0} Nw(N) = 0$, replacing $\frac{\partial F}{\partial L_s}$ in Cahuc *et al.* (2008) by $MRPL_s$ and following the steps of the proof implies that wages in equilibrium to satisfy

$$W_s = (1-\beta_s)\bar{W}_s + P\int z^{\frac{1-\beta_s}{\beta_s}} \frac{\partial \log F}{\partial \log L_s}(\mathbf{L}A_i(z))\, dz \tag{B.29}$$

Taking derivatives and pulling out the term involving output Q, and redefining the remainding term to be equal to G_s gives the result stated in the lemma.

■

In steady state, firm-level employment of each type s is equal to the ratio of match and separation probabilities multiplied by the number of vacancies that a firm posts: $l_s = V_s \frac{q_s}{\chi_s}$. Under convex vacancy posting cost, the optimal number of vacancies solves the following first-order condition:

$$MRPL_s = W_s + \sum_{s'} l_s \frac{\partial W_s}{\partial l_s} + \iota_s l_s^{\frac{1}{\gamma_s}-1},$$

where $\iota_s \equiv \frac{\chi_s}{\gamma_s q_s}$. Following the same steps as in in Cahuc *et al.* (2008), labor demand can be shown to solve:

$$PQ^{\frac{\varepsilon_s}{\sigma}} l_s^{-\frac{1}{\sigma}} \int_0^1 \frac{1}{\beta_s} z^{\frac{1-\beta_s}{\beta_s}-\frac{1}{\sigma}} G_s(\mathbf{L}A_i(z))\, dz = W_s + \iota_s l_s^{\frac{1}{\gamma_s}-1}. \tag{B.30}$$

Assuming that outside options are normalized to zero, that is $\overline{W}_s \equiv 0$, combining equation (B.28) and equation (B.30), implies that labor demand equals:

$$l_s = \left(\frac{(1-\beta_s)}{\iota_s} PQ^{\frac{\varepsilon_s}{\sigma}} \int_0^1 \frac{1}{\beta_s} z^{\frac{1-\beta_s}{\beta_s}-\frac{1}{\sigma}} G_s(\mathbf{L}A_i(z))\, dz\right)^{\frac{\sigma}{\gamma_s+\sigma}},$$

where $\tilde{\gamma}_s \equiv \frac{1}{\gamma_s} - 1$. Plugging this expression into the expression for wages, we obtain the result.

■

C Table and Figure Appendix

Table C.1 The Effect of Firm Sales on Worker Wages by Skill Group

		(1)	(2)	(3)	(4)	(5)	(6)
		Sales	Rev/L	VA	Sales	Rev/L	VA
Skill Group	1	0.016***	0.007	0.012	0.002	−0.008**	−0.009*
		(3.57)	(0.88)	(1.37)	(0.27)	(-2.45)	(-1.79)
	2	0.021***	0.014*	0.017*	0.005	−0.004	−0.007
		(4.73)	(1.72)	(1.94)	(0.78)	(−1.07)	(−1.25)
	3	0.024***	0.019**	0.020***	0.008	0.002	−0.003
		(5.37)	(2.24)	(2.27)	(1.32)	(0.46)	(-0.63)
	4	0.026***	0.022***	0.023***	0.012*	0.006*	0.001
		(5.97)	(2.66)	(2.55)	(1.76)	(1.73)	(-0.1)
	5	0.028***	0.025***	0.025***	0.015**	0.011***	0.002
		(6.48)	(3.06)	(2.81)	(2.24)	(3.11)	(0.5)
	6	0.030***	0.029***	0.027***	0.018***	0.017***	0.006
		(6.91)	(3.48)	(3.06)	(2.72)	(4.61)	(1.18)
	7	0.033***	0.032***	0.029***	0.022***	0.022***	0.009*
		(7.31)	(3.92)	(3.32)	(3.21)	(6.12)	(1.88)
	8	0.036***	0.037***	0.033***	0.026***	0.029***	0.014***
		(8.24)	(4.71)	(3.77)	(3.71)	(7.68)	(2.75)
	9	0.041***	0.045***	0.038***	0.032***	0.038***	0.020***
		(9.87)	(5.51)	(4.24)	(4.51)	(9.92)	(4.13)
	10	0.051***	0.060***	0.049***	0.042***	0.053***	0.03***
		(10.45)	(7.58)	(5.26)	(5.5)	(12.02)	(8.19)
Within R^2		0.32	0.33	0.33	0.31	0.31	0.31
Firm Controls		✓	✓	✓			
Worker Controls		✓	✓	✓	✓	✓	✓
Industry FE		✓	✓	✓			
Occupation FE		✓	✓	✓			
Year FE		✓	✓	✓			
Education FE		✓	✓	✓			
Worker FE					✓	✓	✓
Firm FE					✓	✓	✓
Observations		5061226	5061226	5061226	6483778	6483778	4902074

Notes: This table presents the estimated relation between firm sales on individual worker wages across skill groups, as in equation (2.1). Skill groups are defined in Section 2, with 1 denoting the lowest and 10 denoting the highest skill group. Firm controls include the revenue share of intermediate inputs and the average skill intensity of all full-time employees. Worker controls include age, gender, and nationality. Regressions (1) to (3) include fixed effects for a worker's occupation, sector of occupation, the federal state of residence, and years. Standard errors are clustered at the establishment level, and t-statistics are in parentheses. *** indicates significance at the 1%, ** at the 5% and * at the 10% level.

Table C.2 The Effect of Firm Size on Worker Wages Relative to the Mean Wage of Their Employer by Skill Group

		(1)	(2)	(3)	(4)	(5)	(6)
Log Sales×		Sales	Rev/L	VA	Sales	Rev/L	VA
Skill Group	1	-0.014^{***}	-0.020^{***}	-0.015^{***}	-0.007^{*}	-0.012^{***}	-0.006
		(-13.36)	(-11.25)	(-12.92)	(-1.79)	(-14.8)	(-1.43)
	2	-0.011^{***}	-0.014^{***}	-0.011^{***}	-0.007^{*}	-0.012^{***}	-0.005
		(-10.45)	(-8.48)	(-10.28)	(-1.79)	(-12.2)	(-1.41)
	3	-0.008^{***}	-0.009^{***}	-0.005^{***}	-0.007^{*}	-0.012^{***}	-0.005
		(–7.69)	(–6.04)	(–7.54)	(-1.80)	(-12.2)	(-1.40)
	4	-0.005^{**}	-0.006^{***}	-0.006^{**}	-0.006^{*}	-0.011^{***}	-0.005
		(-5.44)	(-3.88)	(-5.31)	(-1.78)	(-11.4)	(-1.39)
	5	-0.003^{***}	-0.003^{*}	-0.003^{***}	-0.006^{*}	-0.011^{***}	-0.005
		(-3.09)	(-1.79)	(-2.99)	(-1.71)	(-10.7)	(-1.35)
	6	-0.001	0.001	-0.001	-0.006	-0.010^{***}	-0.005
		(-0.83)	(0.24)	(0.69)	(-1.59)	(-9.77)	(-1.28)
	7	0.002	0.004^{**}	0.002	-0.005	-0.009^{***}	-0.004
		(1.38)	(2.30)	(1.58)	(-1.39)	(-8.29)	(-1.11)
	8	0.005^{**}	0.009^{***}	0.005^{**}	-0.004	-0.07^{***}	-0.003
		(4.51)	(5.48)	(4.92)	(-1.08)	(-6.68)	(-0.72)
	9	0.009^{***}	0.002^{***}	0.009^{***}	-0.002	-0.003^{***}	-0.00
		(8.12)	(9.18)	(8.77)	(-0.44)	(-2.75)	(-0.01)
	10	0.018^{***}	0.029^{***}	0.018^{***}	0.004	0.006^{***}	0.006^{*}
		(13.70)	(15.72)	(14.77)	(1.13)	(2.7)	(1.7)
Within R^2		0.22	0.23	0.22	0.06	0.06	0.05
Firm Controls		✓	✓	✓			
Worker Controls		✓	✓	✓	✓	✓	✓
Industry FE		✓	✓	✓			
Occupation FE		✓	✓	✓			
Year FE		✓	✓	✓			
Education FE		✓	✓	✓			
Worker FE					✓	✓	✓
Firm FE					✓	✓	✓
Observations		5061226	5061226	5061226	6483788	6483788	4902074

Notes: Coefficient estimates of estimating equation (2.2). Firm control: Revenue share of intermediate inputs. Worker controls include age, gender and nationality. Regressions (1) to (3) include fixed effects for a worker's occupation, sector of occupation, the federal state of residence, and years. Standard errors are clustered at the establishment level, and t-statistics are in parentheses. *** indicates significance at the 1%, ** at the 5% and * at the 10% level.

Table C.3 Workforce Composition and Firm Performance

	Skill Intensity
Log Revenue	0.33***
	(24.01)
Firm Controls	✓
Fixed Effects	✓
Within R^2	0.11
Observations	69004

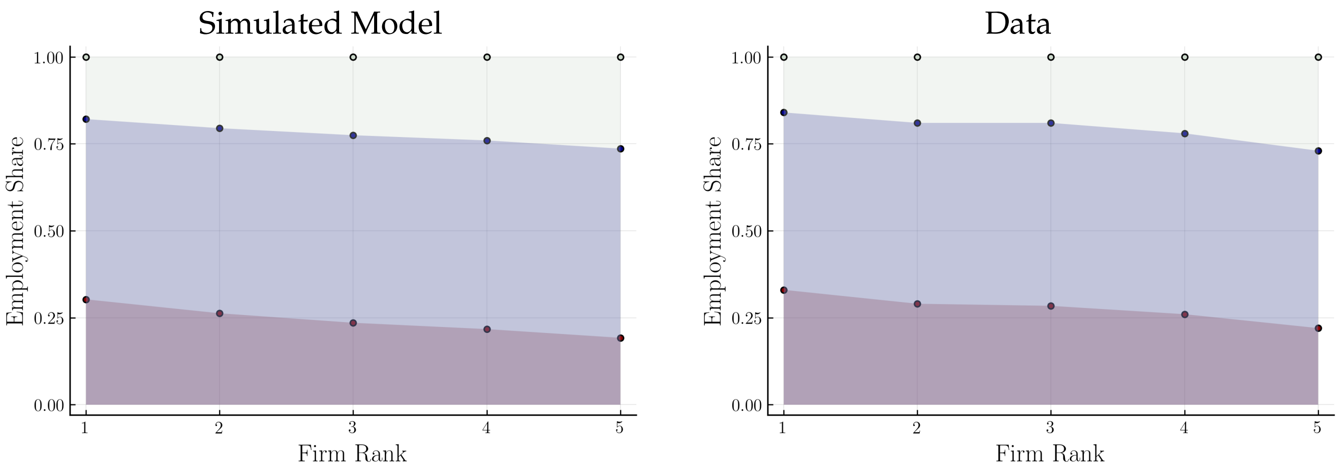
Notes: This table presents the regression results of projecting a firm's skill-intensity, measured by the average wage rank of its employees, on log revenues, firm controls, and fixed effects. Standard errors are displayed in parentheses and clustered at the level of employers f. Fixed effects are included for sectors, federal states, and years. Firm controls include the share of intermediate inputs in total revenues, the total number of workers, average employee age, the share of female employees, and the share of german employees.

Table C.4 Moments for Production Function Estimation

Moment	Skill Group s	Region	
		East	West
$\frac{\text{Cov}\left[d\log\left(l_{s,f,t}/l_{L,f,t}\right), d\log\xi_{f,t}\right]}{\text{Cov}\left[d\log\left(W_{s,f,t}l_{s,f,t}/W_{M,f,t}l_{M,f,t}\right), d\log\xi_{f,t}\right]}$	Low	1.22	0.88***
		(0.85)	(5.96)
	High	0.59**	0.95***
		(1.98)	(5.97)
$\frac{\text{Cov}\left[d\log Q_{f,t}, d\log\xi_{f,t}\right]}{\text{Cov}\left[d\log\left(W_{s,f,t}l_{s,f,t}/W_{M,f,t}l_{M,f,t}\right), d\log\xi_{f,t}\right]}$	Low	3.76	−5.07***
		(0.53)	(-6.6)
	High	1.33	5.65***
		(0.98)	(6.80)
First Stage F		24.3	
Group×Year×Lag Fixed Effect		✓	
Observations		123,082	

Notes: Estimated moments used for the estimation of the production function parameters. t-statistics are displayed in paentheses. *** indicates significance at the 1%, ** at the 5% and * at the 10% level.

Figure C.1 Sorting: Employment Shares of Medium Skilled Workers in Blue, of Low Skilled Workers in Red



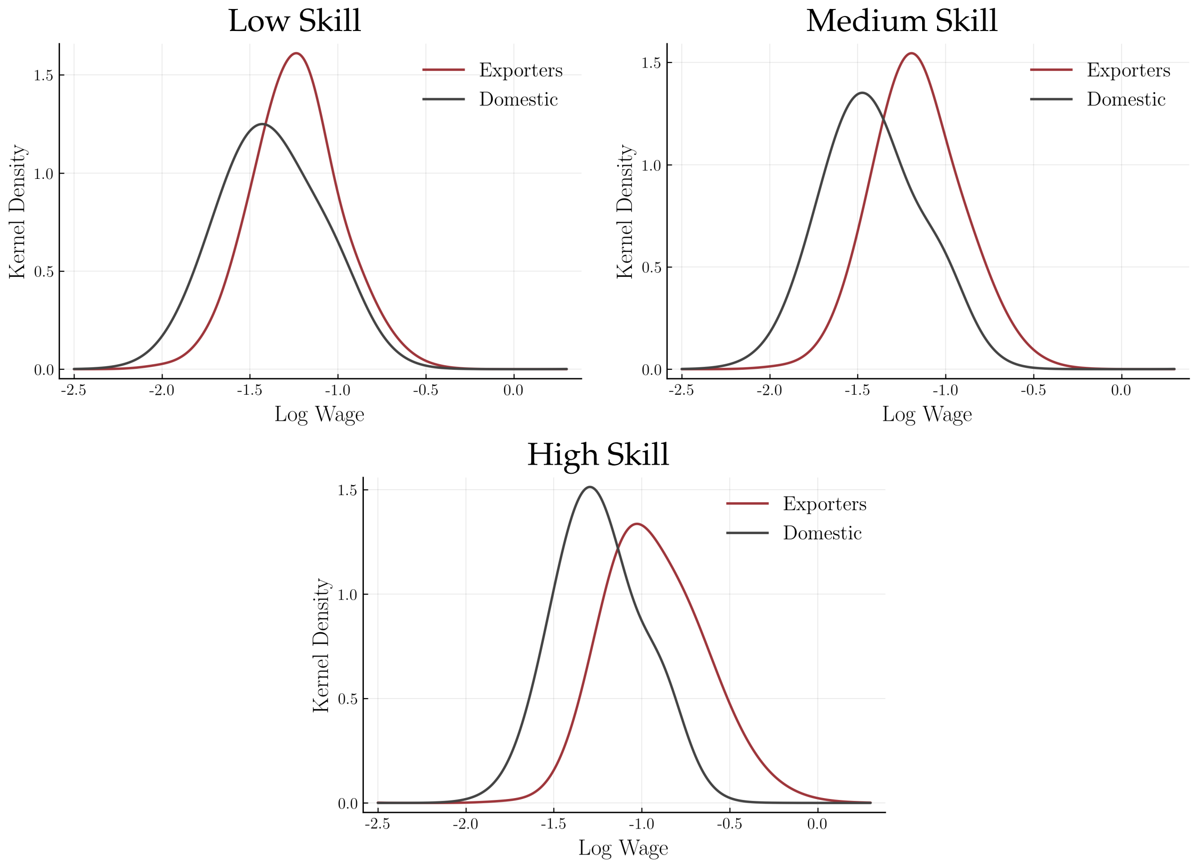
Notes: This figure plots model simulated employment distributions along the firm sales distribution on the left and the empirical distribution on the right. Firm ranks correspond to quintiles of the firm sales distribution.

Table C.5 Over-Identification Check: Returns to Skill and Firm Size

		Wage Relative to Mean Firm Wage	
		Model	Data
Log Sales × Skill group	L	-0.04	-0.02
	M	-0.003	-0.001
	H	0.05	0.02
		$sd\left(\log W_f\right)$	
		Model	Data
Log Sales		0.01	0.01

Notes: The first panel displays the coefficient estimate β_s of estimating: $\log W_{f,s} - \log \overline{W}_f = \alpha + \beta_s \log \text{Sales}_{f,t} + \epsilon_{f,t}$. H, M, and L denote high, medium, and low skill worker groups. The estimation results reported from the data point are based on individual worker observations and control for worker heterogeneity not captured by the model: Age, gender, tenure, sector, occupation, education. The second panel displays the coefficient estimate β of estimating: $sd\left(\log W_{f,s}|f\right) = \alpha + \beta \log \text{Sales}_f + \gamma \log L_{f,t} + \epsilon_f$. The estimation results reported from the data control for firm heterogeneity that is not captured by the model: Sector and intermediate input share.

Figure C.2 Simulated Wage Distributions Across Exporting and Domestic Firms



Notes: This figure displays simulated wage distributions by skill type and exporter status of a worker's employers.

Table C.6 Detailed Counterfactual Effects of Trade on Wages

	Change		
	Model	Data	% Explained
Low Skill			
sd(log W)	0.003	0.017	17
90-10 ratio	0.019	0.03	62
Medium Skill			
sd(log W)	0.004	0.02	21
90-10 ratio	0.011	0.13	9
50-10 ratio	0.03	0.05	61
High Skill			
sd(log W)	0.008	0.05	16
90-10 ratio	0.03	0.12	26
50-10 ratio	0.03	0.05	61
Mean Wage Differences			
Between High and Medium Skill	0.01	0.029	32
Between Medium and Low Skill	-0.008	-0.031	24

*Notes:*This table displays a detailed breakdown of the counterfactual effects of trade liberalization on earnings inequality in Germany between 1993-2002 and 2003-2014.

D Estimation Appendix

D.1 Production function estimation

D.1.1 Estimator Properties

The estimation approach leverages a classical minimum distance estimator. This is natural, given that the estimation procedure first estimates regional pass-throughs of firm-level demand shocks into relative employment and wages and then uses these moments to a system of moment equations.

Denote by $\hat{\beta}^{X}_{Y,\hat{\xi}}$ the coefficient estimate of a two-stage IV regression of an outcome X on a dependent variable Y,instrumented by changes in $\hat{\xi}$ as a first stage. The set of moment condition in equation (4.5) can then be written as:

$$\mathcal{G}\left(\Theta,\hat{\beta}\right)=0. \tag{D.1}$$

For a weighting matrix W and set of estimates $\hat{\beta}$, the minimum distance estimator for the parameters of interest $\hat{\Theta}$ then solves the following problem:

$$\hat{\Theta}=\arg\min_{\Theta}\mathcal{G}\left(\Theta,\hat{\beta}\right)'W\mathcal{G}\left(\Theta,\hat{\beta}\right). \tag{D.2}$$

The estimate will thus solve the following first-order condition:

$$\left(\frac{\partial\mathcal{G}\left(\hat{\Theta},\hat{\beta}\right)}{\partial\Theta}\right)'W\mathcal{G}\left(\hat{\Theta},\hat{\beta}\right)=0. \tag{D.3}$$

As we assume that $\hat{\beta}$ is a consistent estimator of β, the continuous mapping theorem implies that $\hat{\Theta}$ is a consistent estimator of the true underlying set of parameters, that is $\hat{\Theta}\xrightarrow{p}\Theta$. To construct standard errors, note that the estimates inherit any imprecision stemming from the set of two-stage IV estimates $\hat{\beta}$. To derive the variance-covariance matrix, define a mapping H that maps the estimates of reduced form moments into the structural parameters of interest by way of the first-order condition (D.3):

$$H\left[\hat{\beta},\Theta\left(\hat{\beta}\right);W\right]\equiv\left(\frac{\partial\mathcal{G}\left(\hat{\Theta},\hat{\beta}\right)}{\partial\Theta}\right)'W\mathcal{G}\left(\hat{\Theta},\hat{\beta}\right)=0. \tag{D.4}$$

Denoting by P the total number of observations used to compute the moments β, the asymptotic distribution of $\hat{\Theta}$ can be derived by applying first the implicit function theorem and then the Delta method. The asymptotic distribution of $\hat{\Theta}$ is then given by:

$$\sqrt{P}\left(\hat{\Sigma}-\Sigma\right)\xrightarrow{d}\mathcal{N}\left(0,\tilde{H}'V_{\hat{\beta}}\tilde{H}\right), \tag{D.5}$$

where $\tilde{H}\equiv\left(H_2\left[\hat{\beta},\hat{\Theta};W\right]\right)^{-1}H_1\left[\hat{\beta},\hat{\Theta};\Omega\right]$. H_2 and H_1 denote multivariate derivatives of $\frac{\partial\mathcal{G}(\Theta,\beta)}{\partial\Theta}'W\mathcal{G}(\Theta,\beta)$ with respect to Θ and β respectively. $V_{\hat{\beta}}$ denotes the variance-covariance matrix of the estimates of β. The standard errors of the parameters depend directly on the variance-covariance structure obtained from the estimation of β.

D.1.2 Identification

Endogenous elasticity of labor supply to the firm The identification argument laid out in Section 4.1.2 allows the elasticity of the elasticity of labor supply to be firm-specific. However, it assumes that the elasticity to labor supply is exogenous to the actions of the firm. However, in many models of monopsony, the elasticity of the labor supply to the firm depends on the wage that it offers (e.g. Burdett & Mortensen (1998), Manning (2011), Manning & Petrongolo (2017), Berger *et al.* (2021)). Here I lay out sufficient conditions under which the estimation approach outlined in the main text can be augmented to accommodate endogenous heterogeneity in labor supply to the firm.

To that end, consider an (equilibrium) labor supply function $\mathcal{L}(W,\Omega)$ to an individual firm (omitting subscripts for skill types) and assume that $\mathcal{L}$ is continuously differentiably. W denotes the wage, and Ω denotes general equilibrium objects, such as wages paid by all other firms and the stock of total workers. Denote $\beta(W,\Omega)\equiv\frac{\partial\log\mathcal{L}}{\partial\log W}$ the elasticity of labor supply to individual firms. Wages are be given by equation (3.9) with labor supply elasticities being given by $\beta(W,\Omega)$.

The estimation approach outlined in the main text requires partial identification of the parameters of the production function from a shock that is exogenous to unobserved sources of heterogeneity in labor demand. If changes in the mark-down on wages $M(\beta)\equiv\frac{\beta(W)}{\beta(W)+1}$ in response to the shock are estimable, partial identification can be attained. Intuitively, changes in the mark-down term are estimable, if $\beta(W)$ admits "sufficient statistics" in the sense of Chetty (2009). $\beta(W,\Omega)$ is said to admit sufficient s tatistics, if there exist statistics $\mathbf{t}=(t_1,...,t_N)$ such that

$$d\log\beta=\alpha_w d\log W+\sum_{n=1}^{N}\alpha_N d\log t_n.$$

Then changes in the mark-down on wages can be written:

$$dm\equiv d\log M(\beta)=\underbrace{\left(\alpha_w d\log W+\sum_{n=1}^{N}\alpha_N d\log t_n\right)}_{\equiv dz}\underbrace{\left(-\frac{\beta}{\beta+1}\right)}_{\equiv\tilde{M}}.$$

The covariance between changes in the mark-down and an exogenous shock ξ can be written:

$$Cov\left(dm, d\xi\right) = E\left(dz\right)Cov\left(\tilde{M}, d\xi\right) + E\left(\tilde{M}\right)Cov\left(dz, d\xi\right) + \nu,$$

where $\nu = E\left[\left(\tilde{M} - E\left(\tilde{M}\right)\right)\left(dz - E\left(dz\right)\right)\left(\xi - E\left(\xi\right)\right)\right]$.

If *changes* in ξ are independent from initial *levels* in mark-downs M, then $Cov\left(\tilde{M}, d\xi\right) = 0$ and $\nu = 0$. Then the above expression simplifies to:

$$Cov\left(dm, d\xi\right) = E\left(M\right)Cov\left(dz, d\xi\right).$$

Partial identification requires that the following term is estimable, up to structural constants:

$$E(\tilde{M}_s)Cov\left(dz_s, d\xi\right) - E(\tilde{M}_{s'})Cov\left(dz_{s'}, d\xi\right) + \left(\varepsilon_s - \varepsilon_{s'}\right)dq.$$

If the statistics $\mathbf{t} = (t_1, ..., t_N)$ can be observed, then $Cov\left(dz, d\xi\right)$ is computable, up to the constants $\alpha_w, (\alpha_n)$. If $\beta > 1$, the last term in turn can be expanded:

$$E\left(\tilde{M}\right) = \frac{b}{1+b} + \frac{b}{\left(1+b\right)^2}E\left(dz\right) + \mathcal{O}\left(\frac{b^2}{\left(1+b\right)^3}\right),$$

for some constant b. Thus partial identification to first-order is ensured.

Sufficient statistics are admitted in many models. An extension of the present framework where firms internalize their effect on aggregate labor demand admits the sufficient statistic $t_{s,f} = \frac{W_{s,f}l_{s,f}}{\sum_{f'} W_{s,f'}l_{s,f'}}$, as in Berger *et al.* (2021).

Variation across regions A condition for full identification is that the system of moment conditions is not collinear. The corresponding condition given in the main text reads:

$$\frac{\mathrm{Cov}_r\left(d\log\left(\frac{l_{s,f,t}}{l_{s',f,t}}\right), d\log\hat{\xi}_{f,t}\right)}{\mathrm{Cov}_r\left(d\log\left(Q_{f,t}\right), d\log\hat{\xi}_{f,t}\right)} \neq \frac{\mathrm{Cov}_{r'}\left(d\log\left(\frac{l_{s'',f,t}}{l_{f,s',t}}\right), d\log\hat{\xi}_{f,t}\right)}{\mathrm{Cov}_{r'}\left(d\log\left(Q_{f,t}\right), d\log\hat{\xi}_{f,t}\right)}.$$

Here, I use the structural model to illuminate the economic mechanisms.

Using the structural expression for labor supplies, $\mathrm{Cov}_r\left(d\log\left(\frac{l_{s,f,t}}{l_{s',f,t}}\right), d\log\hat{\xi}_{f,t}\right)$ can be written as:

$$\begin{aligned}
\mathrm{Cov}_r\left(d\log\left(\frac{l_{s,f,t}}{l_{s',f,t}}\right), d\log\hat{\xi}_{f,t}\right) \quad = \quad & \beta_s\mathrm{Cov}_r\left(d\log\left(W_{s,f,t}\right), d\log\hat{\xi}_{f,t}\right) \\
- \quad & \beta_{s'}\mathrm{Cov}_r\left(d\log\left(W_{s',f,t}\right), d\log\hat{\xi}_{f,t}\right) \\
+ \quad & \mathrm{Cov}_r\left(d\log\left(\Lambda_s/\Lambda_{s'}\right), d\log\hat{\xi}_{f,t}\right) \\
+ \quad & \mathrm{Cov}_r\left(d\log\left(A_{s,f}/\Lambda_{s',f}\right), d\log\hat{\xi}_{f,t}\right)
\end{aligned}$$

The effect operating through aggregate labor demand can be written as:

$$\text{Cov}_r\left(d\log\left(\Lambda_s\right), d\log\hat{\xi}_{f,t}\right) = \sum_{f\in\mathcal{F}} \frac{W_{s,f}l_{s,f}}{\sum_{s,f'} W_{s,f'}l_{s,f'}}\beta_s \text{Cov}_r\left(d\log\left(\frac{W_{s,f}}{A_{s,f}}\right), d\log\hat{\xi}_{f,t}\right).$$

These terms will differ across regions so long as employment shares and shock exposure, as measured by the covariance term on the RHS, are not perfectly correlated across regions. As the empirical implementation uses a foreign demand shock, this condition will be satisfied if, for example, the market wage bill shares of exporting and domestic firms differ across both regions. Given the sizable structural differences between former East and West German federal states, this is likely to be the case in the data.

$\text{Cov}_r\left(d\log\left(Q_{f,t}\right), d\log\hat{\xi}_{f,t}\right)$ can be written as:

$$\begin{aligned}\text{Cov}_r\left(d\log\left(Q_{f,t}\right), d\log\hat{\xi}_{f,t}\right) = \; & \text{Cov}_r\left(\frac{\tau Q^X_{f,t}}{Q_{f,t}}\left(-\eta\frac{d\log MC_{f,t}}{d\log\hat{\xi}_{f,t}} - \eta d\log\tau + d\log Y^*\right), d\log\hat{\xi}_{f,t}\right) \\ & + \text{Cov}_r\left(\frac{Q^H}{Q_{f,t}}\left(-\eta\frac{d\log MC_{f,t}}{d\log\hat{\xi}_{f,t}} + d\log P^{\eta-1}Y\right), d\log\hat{\xi}_{f,t}\right)\end{aligned}.$$

If the distribution of export-shares $\frac{\tau Q^X_{f,t}}{Q_{f,t}}$ differs across regions, then $\text{Cov}_r\left(d\log\left(Q_{f,t}\right), d\log\hat{\xi}_{f,t}\right)$ will differ across regions.

To summarize, the set of partially identified parameters across regions will inform the technological parameters if the regions differ characteristics of exporting firms, export shares of firms and/or the distribution of workers across firms.

D.2 Model Solution Algorithm

I first discretize the joint distribution of demand shifters φ and amenities A by calculating nodes and their associated weights through a quadrature routine (qnwlogn) that is supplies as part of both the "QuantEcon.jl" and "CompEcon.jl" packages in Julia.

The solution algorithm of the model is detailed below.

1. Guess the aggregate number of firms.
2. Guess aggregate income D = Labor Income + Profits - Fixed Cost
3. Loop: Solve for equilibrium in product and labor markets, given the guess for D.
    (a) Guess labor market tightness for all worker groups $s : \Delta_s = \left(\int W^{\beta_s}_{s,f} df\right)^{-1}$
        i. Loop: Solve for price index P
            A. Guess P.

B. Given guesses for Δ_s, D and P, solve for profit-maximizing wages, labor demands and market entry decisions of firms.

C. Use the marginal costs obtained in step B. to calculate an updated guess for the aggregate price index. Update guess for P until convergence.

ii. Use wages from step ii. to update the guess on aggregate labor demand. Update guess until convergence.

4. Use wages, profits, and market entry decisions obtained in step 2. to update guess of income D. Update guess with step-size halfway between the old and new guess and return to step 1. Repeat until convergence.

5. Given the implied expected profits, calculate the mass of entering firms. Update the initial guess by taking a convex combination of initial and implied mass of firms and return to step 1 until convergence.

D.3 Calibration Algorithm

To calibrate the model, I define the following loss function is given by:

$$L(\Theta) = (\mathcal{M}(\Theta) - \overline{m})' \, W \, (\mathcal{M}(\Theta) - \overline{m}),$$

where $\mathcal{M}(\Theta)$ is a vector that contains the simulated moments for parameter values Θ, $\overline{m}$ denotes the vector of empirical target moments and W is a weighting matrix. I weight moments by the square of their inverse target value.

To find the set of parameters that best fits the target moments $\mathcal{M}(\Theta)$, I implement the following algorithm.

1. Initialize guess for the parameter space. $a_0 = \left[\underline{\Theta}_0, \overline{\Theta}_0\right]$

2. Evaluate the objective function at points given by a Sobol sequence with 500 elements through the parameter space.

3. Keep the guesses falling into the 30th percentile of Loss function evaluations and update the guess of the parameter space through keeping the respective highest and lowest values.

4. Return to step 2 and increase the number of elements in the Sobol sequence.

5. Repeat this procedure until convergence.

A Sobol sequence (Sobol (1967)) generates quasi-random numbers through $A \subset \mathbb{R}^n$ that fill the space of possibilities more evenly than pseudo-random numbers and therefore allows for faster convergence of the algorithm. To implement the algorithm, I use the "Sobol.jl" package in Julia.