Supplemental Appendix

Division of Labor in the Global Economy

— Sascha O. Becker, Hartmut Egger, Michael Koch, and Marc-Andreas Muendler —

September 10, 2025

This Supplemental Appendix has three parts: Section S1, presents additional empirical evidence, Section S2 comprises further derivation details, and Section S3 discusses the implementation of our estimation model.

S1 Further evidence

S1.1 Summary statistics

As described in Section 2.3, we combine the BIBB-BAuA labor force survey information with the LIAB linked plant–worker records. To include task information from BIBB-BAuA alongside the LIAB linked plant–worker data, we use the within occupation variance of log daily wage by plant, job experience, squared job experience, indicators for (i) gender, (ii) 7 schooling and vocational training indicators, (iii) 16 regions, (iv) 34 sectors, (v) 7 plant-size categories, and (vi) 335 occupations over the years 1999, 2006, 2012, and 2018. We predict using a probit estimation the probability that a worker reports performing a given task in the BIBB-BAuA sample and, using the same regressors, the probability that a worker in the LIAB linked plant–worker sample performs the task. Table S.1 shows the raw data from LIAB as well as the imputed task information.

Table S.1 also reports summary statistics on revenues and other relevant plant attributes from the combined LIAB and BIBB-BAuA data. For our final dataset, we eliminate marginal workers with a daily wage below a minimum threshold (see Stüber et al. 2023) as well as workers with imputed daily wages higher than EUR 2,000 in 1998 prices.[1] Moreover, excluding plants for which we lack relevant information as well as plants with employment of less than two full-time workers (for which we cannot compute meaningful measures of wage dispersion) our sample covers 24,993 plant-year observations, with 7,572 of these observations referring to exporters.

[1] By eliminating workers with unreliable wage information we lose only 56 plant-year observations. However, we increase the lowest plant-level daily wage average from an unrealistically low level of EUR 4.40 to (a still low level of) EUR 10.87, while we reduce the highest plant-level daily wage average from several billion euros to EUR 255,09.

Table S.1: Descriptive Statistics for Combined Data

	Obs.	Mean	Median	StDev.	Min.	Max.
log Revenues	24,993	13.694	13.469	1.410	6.626	24.538
log Export revenues	7,572	17.211	17.105	2.130	11.819	28.677
Export indicator	24,993	0.174	0.000	0.379	0.000	1.000
Employment (full-time)	24,993	17.994	6.000	112.621	2.000	53,809
log Daily wage	24,993	4.111	4.127	0.376	2.386	5.542
CV Residual daily wage	24,993	0.202	0.175	0.141	0.000	1.368
CV Daily wage	24,993	0.198	0.167	0.149	0.000	1.126
Count 3-digit occupations n	24,993	4.276	3.000	5.124	1.000	153
Average number of tasks b	24,993	6.331	6.593	1.538	0.596	10.846
Number of distinct tasks $\tilde{\beta}$	24,993	8.625	8.816	2.082	0.911	14.769
Normalized number of tasks $b/\tilde{\beta}$	24,993	0.748	0.751	0.144	0.083	1.000

Sources: LIAB and BIBB-BAuA 1999, 2006, 2012 and 2018. Sample restricted to plants with more than two full-time workers, excluding marginal workers with a daily wage below a minimum threshold (see Stüber et al. 2023) and excluding workers with daily wages higher above EUR 2,000 in 1998 prices.
Notes: Descriptive statistics based on annual plant observations, using inverse probability weights to make plant sample representative of Germany economy, as suggested by the Research Data Centre at the IAB. *CV* is coefficient of variation of daily (raw or residual) wages within a plant-occupation. The daily wage residual is obtained from a Mincer regression (in logs), including demographic, education and tenure information as well as time, sector and region fixed effects and plant revenues.

S1.2 Workplace operations

Using the BIBB-BAuA labor force survey data for the four waves 1999, 2006, 2012 and 2018, Table S.2 shows the frequency of workplace operations (tasks) for the overall sample period as well as the individual observation years. We inversely weight the frequency of worker observations by their sampling frequency to achieve representativeness.

A comparison across columns of Table S.2 shows a shift towards multitasking between 1999 and 2006 that is reflected by an increase in the performance of all workplace operations. Since 2006 the total number of tasks conducted by German workers appears to be fairly stable, whereas there seems to be a decreasing importance of activities related to “Manufacture, Produce Goods” and an increasing importance of activities related to "Apply Legal Knowledge”. This may speak for a general (heavily criticized) increase in the bureaucracy at German workplaces after the millennium.[2]

[2]Focusing on subsamples of workers, we observe total frequencies across the four sample periods of 7.230 for workers earning above the median daily wage; 6.997 for workers aged 45 and older; 7.427 for those holding a college-qualifying secondary education diploma (Abitur or equivalent); and 7.573 for supervisors and managers. To ensure representativeness, we inversely weight the frequency of worker observations by their sampling probability. These patterns suggest that German workers engage in multitasking to a similar extent across skill levels, age groups, hierarchical layers, and the entire wage distribution.

Table S.2: Frequency of Workplace Operations

Workplace Operations (Tasks)		Individual years			
		1999	2006	2012	2018
1. Manufacture, Produce Goods	0.177	0.162	0.208	0.184	0.157
2. Repair, Maintain	0.354	0.308	0.407	0.364	0.338
3. Entertain, Accommodate, Prepare Foods	0.208	0.226	0.195	0.213	0.200
4. Transport, Store, Dispatch	0.447	0.353	0.500	0.495	0.436
5. Measure, Inspect, Control Quality	0.626	0.463	0.672	0.664	0.689
6. Gather Information, Develop, Research, Construct	0.796	0.517	0.851	0.864	0.921
7. Purchase, Procure, Sell	0.461	0.421	0.485	0.479	0.459
8. Program a Computer	0.110	0.052	0.140	0.111	0.132
9. Apply Legal Knowledge	0.591	0.194	0.680	0.689	0.759
10. Consult and Inform	0.866	0.750	0.887	0.892	0.921
11. Train, Teach, Instruct, Educate	0.546	0.375	0.568	0.577	0.642
12. Nurse, Look After, Cure	0.262	0.280	0.256	0.260	0.255
13. Advertise, Promote, Conduct Marketing and PR	0.417	0.282	0.462	0.442	0.470
14. Organize, Plan, Prepare Others' Work	0.690	0.595	0.673	0.693	0.784
15. Control Machinery and Technical Processes	0.346	0.300	0.391	0.362	0.333
Total Number of Tasks	6.896	5.279	7.375	7.288	7.494

Source: BIBB-BAuA 1999, 2006, 2012 and 2018 (inverse sampling weights).
Note: Frequencies of performing a workplace operation (task) at the worker level.

S1.3 The link between plant size, task assignment, and wage dispersion

In this supplement, we revisit Facts 1 to 3 of Section 2 in the main text with proper econometric methods, using plants instead of workers as units of observation for our analysis. To eliminate potential biases from omitted variables, we control for time, commuting zone, and NACE 2-digit industry fixed effects, and to address endogeneity concerns in the relationship between plant size, the normalized number of tasks, and residual wage dispersion, we implement in addition to OLS also IV regressions, in which we instrument revenues by time-varying industry-level exports to China from the United Nations Commodity Trade Statistics Database (Comtrade) and the trade in services database (TSD) at the World Bank. To ensure exogeneity of our instrument, we follow Autor et al. (2013) and Dauth et al. (2014) in using shipments of third countries—Australia, Canada, Japan, Norway, New Zealand, Sweden, Singapore and the United Kingdom—instead of Germany. In the regression analysis, we exclude marginal workers with a daily wage below a minimum threshold (see Stüber et al. 2023) as well as workers with imputed daily wages above EUR 2,000 in 1998 prices to eliminate extreme outliers with implausible wage information.[3]

[3]The reported results are qualitatively unchanged, if we consider log revenues instead of log employment to measure firm size by an output measure more closely related to the profitability of the plant.

Table S.3: Occupation Count, Normalized Task Number, and Residual Wage Dispersion

	Dependent variable (in logs):							
	Number of occupations		CV residual daily wages		Normalized task number		CV residual daily wages	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Second stage								
log Employment	0.597***	0.597***	0.144***	0.308***	-0.130***	-0.099**		
	(0.015)	(0.104)	(0.016)	(0.082)	(0.005)	(0.039)		
log Normalized task nr.							-0.473***	-3.130*
							(0.098)	(1.710)
First stage								
log Industry exports		0.149***		0.149***		0.149**		-0.015*
		(0.035)		(0.035)		(0.035)		(0.008)
Adj. R^2	0.740		0.137		0.616		0.117	
AR F-stat. (p-val.)		0.008		0.000		0.065		0.000
F-stat 1^{st} (p-val.)		0.000		0.000		0.000		0.057
Observations	20,004	20,004	20,004	20,004	20,004	20,004	20,004	20,004

Sources: LIAB and BIBB-BAuA 1999, 2006, 2012 and 2018, all sectors. Plants with at least two full-time workers, excluding marginal workers with a daily wage below a minimum threshold and workers with imputed daily wagers higher than EUR 2,000 in 1998 prices.
Notes: Specifications include time, NACE 2-digit industry, and commuting zone fixed effects. Standard errors clustered at the industry and the commuting zone level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

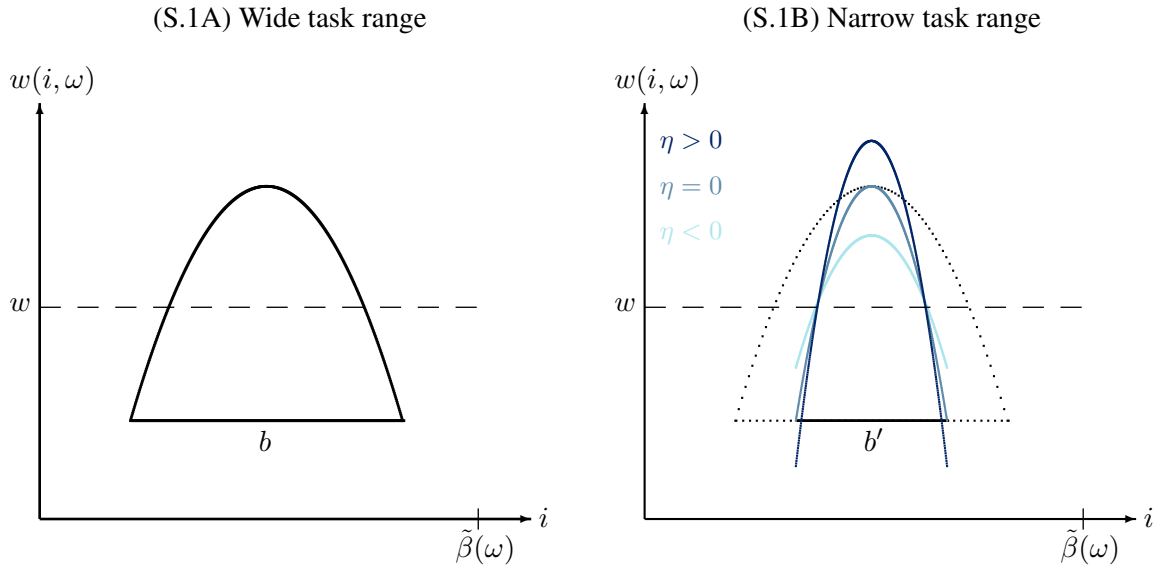
Table S.3 summarizes the results from our empirical analysis, with Columns 1 to 4 looking into the relationship between plant size, occupation counts and plant-occupation internal wage dispersion proposed by Fact 1: larger plants have a higher occupation count as well as higher plant-occupation internal wage dispersion. For instance, the OLS results reported in Columns 1 and 3 indicate that a ten percent increase in employment is associated with a six percent increase in the count of occupations and more than a one percent increase in residual wage dispersion within plant-occupations. In Columns 2 and 4, we report results from corresponding IV regressions. These parameter estimates provide further support for Fact 1 and the reported p-values of the relevant F-statistics indicate that the chosen instrument has explanatory power.

In Columns 5 and 6 we report evidence for a negative link between plant size and the normalized number of tasks conducted by workers, which confirms Fact 2: larger plants are internally more specialized. The estimated positive coefficient of Column 5 suggests that a ten percent increase in plant-level employment is associated with a more than one percent decrease in the normalized number of tasks conducted by workers. Column 6 shows that choosing an IV estimator does not change this result. Finally, Columns 7 and 8 confirm the negative link between the number of normalized tasks and the plant-occupation internal wage dispersion proposed by Fact 3. However, choosing an IV estimator strongly increases the (now less precisely) estimated coefficient, which may speak for an omitted variable bias of OLS that exists, for instance, if a confounding factor that may be rooted in uncontrolled technology differences affects both wage dispersion and the normalized number of tasks. We interpret the overall evidence from Table S.3 as suggestive of a direct reorganization channel in the plant's internal labor market, by which product-market expansions in the wake of globalization trigger a more specialized division of labor and lead to more residual wage dispersion within plant-occupations.

S1.4 The role of η for wage variability and employer success

Given η's important role in shaping wage dispersion (and worker efficiency differences), we illustrate its impact by showing how narrower task ranges affect within-occupation wage variability under different values of η. Panel A of Figure S.1 shows the individual wages following from eq. (9) within a plant-occupation that covers a broad task range $b(\omega) = b$. Now suppose the plant optimally adopts a narrower task range $b(\omega) = b' < b$ as depicted in Panel B of Figure S.1. The wage schedule will still vary around the unchanged economy-wide wage w, but it depends on the plant's sensitivity of worker performance to task mismatch η whether the worker efficiency dispersion, and hence the wage variability

Figure S.1: Within Plant-Occupation Wage Schedule and the Task Range

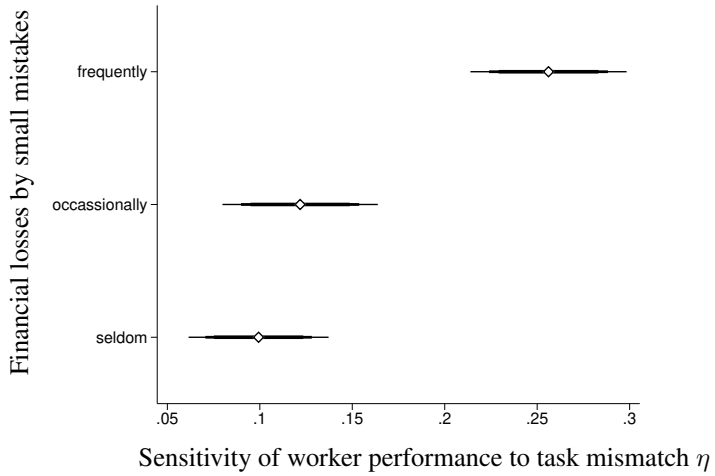


Notes: The graph displays individual wages $w(i,\omega)$ over task interval $[0, b(\omega)]$ as a function of workers' core ability i, according to eq. (9).

around the economy-wide mean, stays constant, rises, or falls at the plant. For a positive sensitivity parameter $\eta > 0$, a narrower task range $b' < b$ magnifies the worker efficiency dispersion and thus induces more variation in workers' wages—with the opposite being true if $\eta < 0$. In practice, workers with badly matched abilities near the boundary of a narrow task range might exhibit a more than proportionally diminished efficiency, if their mistakes on the job can result in heavier losses to the employer than in wider task ranges (as captured by $\eta > 0$). A priori, it is equally conceivable that badly matched workers in narrow task ranges suffer only a less than proportional reduction in efficiency, compared to their efficiency in wide task ranges, if their mistakes matter little to the employer, because narrower task ranges may have a lesser impact on overall production (as captured by $\eta < 0$).

To assess empirically whether and to what extent the sensitivity of worker performance is important for the employer, we can rewrite eq. (10) to obtain a theory-consistent measure of the unobservable sensitivity of worker performance η as a function of the normalized task range $b(\omega)/\tilde{\beta}(\omega)$ and the coefficient of variation of wages $cv(\omega)$, which are both observable in our combined plant-worker dataset. We can then relate the resulting value of η at the plant level to a question about whether workers' small mistakes in their occupation cause the employer financial losses ("Financial losses by small mistake," see Becker and Muendler 2015) from the BIBB-BAuA surveys. Our proposed efficiency-

Figure S.2: Sensitivity of Performance and Financial Losses from Small Mistakes



Source: BIBB-BAuA 1999, 2006 and 2012 with imputed information on the normalized task range $b(\omega)/\tilde{\beta}(\omega)$ and the coefficient of variation $cv(\omega)$ from LIAB.
Notes: Prediction of sensitivity of worker performance by categorical variable on financial losses by small mistakes of workers, controlling for plant size categories as well as state and year fixed effects. Results are differences to omitted category of small mistakes "never" leading to financial losses. Thick, medium, and thin lines represent the 99, 95, and 90 percent confidence intervals.

wage mechanism associates a higher sensitivity of worker performance with the tenet that an employer's surplus ("financial losses") is more responsive to (bad) match quality—thus the stronger wage variability in plants with higher levels of η depicted by Panel B of Figure S.1.

Answers to the question "Financial losses by small mistake" in the BIBB-BAuA survey come in four categories: "never", "seldom", "occasionally", and "frequently or almost always". We run a worker-level regression of our theory-consistent measure for the sensitivity of worker performance to task mismatch, η, on the three worker-reported categories of loss frequencies after mistakes, relative to the omitted category "never". In this regression, we control for plant-size categories as well as state and year fixed effects and report the results in a coefficient plot displayed in Figure S.2.[4] The evidence reported there supports the conclusion that the more likely a worker's mistake causes losses for the plant, the higher is the theory-consistent measure for the sensitivity of worker performance to mismatch, confirming that η indeed captures an important facet of the production process.

[4]BIBB-BAuA does not provide information on "Financial losses by small mistake" for the survey year 2018.

S2 Theoretical background material

S2.1 Extension to Stole-Zwiebel bargaining

A plant ω's revenues are

$$r(\omega) = A^{\frac{1}{\sigma}} \left\{ \tilde{\varphi}(\omega)\tilde{\beta}(\omega)[n(\omega)+1] \exp\left[\frac{1}{n(\omega)+1} \sum_{j=1}^{n(\omega)+1} \ln\left(\int_0^{b(\omega)} \ell_j(i,\omega)\, \lambda(i,\omega)\, \mathrm{d}i \right) \right] \right\}^{1-\frac{1}{\sigma}}$$

where $n(\omega)+1$ is the plant's occupation count, $b(\omega)$ is its task range per occupation, $\tilde{\beta}(\omega)$ is its full task range required for production, $\ell_j(i,\omega)$ is employment of workers of type (core ability) i in the task interval of job j, $\lambda(i,\omega)$ is the labor efficiency of type-i workers in a task interval with range $b(\omega)$, $\tilde{\varphi}(\omega)$ is plant-specific elemental productivity, and A is a constant that captures demand shifters. We assume that hiring is subject to search frictions and wage setting is the result of individual bargaining of the employer with a continuum of workers as derived by Stole and Zwiebel (1996). We can distinguish $n(\omega)+1$ groups of workers by their occupation j and characterize the bargaining outcome at the employer with two equations of the following form:[5]

$$\psi(\omega) = \frac{1}{\ell(\omega)} \int_0^{\ell(\omega)} r\left[k\, \mathbf{s}(\omega)\right] \mathrm{d}k, \qquad \frac{\partial \psi(\omega)}{\partial \ell_j(i\omega)} = w_j(i,\omega), \tag{S.1}$$

where $\psi(\omega)$ is the plant's operating profit, k denotes a proportional increase in employment symmetrically over all the plant's occupations $n(\omega)+1$, $r[\cdot]$ are the plant's revenues as a function of its occupational employment-shares vector $\mathbf{s}(\omega)$, $\ell_j(\omega) \equiv \int_0^{b(\omega)} \ell_j(i,\omega)\mathrm{d}i$ is employment in a task interval with range b, $\ell(\omega) \equiv \sum_{j=1}^{n(\omega)+1} \ell_j(\omega)$ is the plant's total employment, $w_j(i,\omega)$ is type-i worker's wage in an occupation j with task range $b(\omega)$, and each occupation j's employment share at the plant $s_j(\omega) \equiv \ell_j(\omega)/\ell(\omega)$ enters the occupational employment-share vector

$$\mathbf{s}(\omega) \equiv \left(s_1(\omega), ..., s_{n(\omega)+1}(\omega)\right)^T.$$

The first expression in eq. (S.1) links the result of the employer-worker bargaining outcome to the Aumann-Shapley value (Aumann and Shapley 1974).[6] Intuitively, the

[5] Existence and uniqueness of this solution follow from Theorem 9 in Stole and Zwiebel (1996).

[6] Brugemann et al. (2015) point to a conceptual problem with Stole and Zwiebel bargaining because, unlike the argument in the original paper, the order in which workers bargain with the employer does matter for the payoff they receive. As a result, the outcome of the Stole and Zwiebel game differs from the equilibrium prescribed by Aumann-Shapley values. As a remedy, Brugemann et al. (2015) propose

first expression in eq. (S.1) assures that the employer's entire revenues are fully exhausted through bargaining. By the second expression in eq. (S.1), the employer and every worker split the surplus equally so that revenues are divided by the mass of all workers and the employer but, since the employer is non-atomic, it does not affect the mass $\ell(\omega)$ and revenues are divided by $\ell(\omega)$. The plant's operating profit is therefore $\psi(\omega)$.

Employers allocate workers symmetrically over the task range of jobs, so $\ell_j(i,\omega) = \ell_j(0,\omega) = \ell_j(b(\omega),\omega)$ for all $i \in (0, b(\omega))$. Therefore, we obtain $\psi(\omega) = [\sigma/(2\sigma - 1)]r(\omega)$, where $r(\omega)$ follows from above. Substitution into the second expression of eq. (S.1) yields

$$w_j(i,\omega) = \frac{\sigma - 1}{2\sigma - 1} \frac{r(\omega)\lambda_j(i,\omega)}{\lambda_j(\omega)\ell_j(\omega)} \frac{1}{n(\omega) + 1}, \tag{S.2}$$

where occupation-level labor efficiency is

$$\lambda_j(\omega) \equiv \frac{1}{\ell_j(\omega)} \int_0^{b(\omega)} \ell_j(i,\omega)\lambda(i,\omega)\,\mathrm{d}i.$$

Combining $\psi(\omega) = [\sigma/(2\sigma - 1)]r(\omega)$ and eq. (S.2) establishes

$$\frac{w_j(i,\omega)}{\lambda_j(i,\omega)}\lambda_j(\omega)\ell_j(\omega) = \frac{\sigma - 1}{\sigma} \frac{\psi(\omega)}{n(\omega) + 1}. \tag{S.3}$$

Every worker in occupation j therefore receives the same wage per efficiency unit of labor, $w_j(i,\omega)/\lambda_j(i,\omega) \equiv w_j^e(\omega)$, and this condition is sufficient to guarantee a symmetric allocation of workers over their task range, if worker types are uniformly distributed over the employer's full task range $\tilde{\beta}(\omega)$ and an employer gets a random draw of the workers.

With the bargaining solution at hand, we can turn to hiring. We assume that hiring takes place prior to the wage negotiation and involves the costs of advertising jobs for employers. Risk-neutral workers apply for those jobs that promise the highest expected return given the imperfect signal they receive regarding their suitability for executing the tasks required in an occupation, according to a posted vacancy. We assume that the signal the workers receive through a vacancy posting only informs them about whether their core ability i falls within the respective task range, but does not provide further details regarding their core ability's exact position within the task interval. Vacancy posting costs are given by $w^s f_b$, where w^s is a service fee equal to the return on labor used for

to replace the Stole and Zwiebel game by a Rolodex game, by which workers are randomly picked to bargain from a Rolodex shuffle, so as to anchor the bargaining outcome of Stole and Zwiebel (1996) in non-cooperative game theory. The outcome of the Rolodex game remains the same as the one posited in Stole and Zwiebel (1996), so we acknowledge the correction but refer to Stole and Zwiebel (1996) when discussing the solution concept.

providing services. Following Helpman et al. (2010), we propose that vacancy posting costs are positively related to labor market tightness, and decrease in the unemployment rate u. The ex ante probability of workers to be matched with an employer is $(1-u)$. Vacancy posting costs are specified to equal $w^s f_b = w^s B(1-u)^\varepsilon$, where $B > 1$ is a constant parameter and $\varepsilon > 0$ is the elasticity of vacancy posting costs with respect to the employment rate. The hiring problem of the employer can therefore be stated as follows:

$$\max_{\ell_j(\omega)} \psi(\omega) - \sum_{j=1}^{n(\omega)+1} w^s B(1-u)^\varepsilon \ell_j(\omega) - w^s \tilde{\zeta}(\omega)\lambda(\omega)^\gamma - w^s \tilde{\zeta}(\omega) f_0. \tag{S.4}$$

The first-order condition of this optimization problem is equivalent to

$$[n(\omega)+1]\ell_j(\omega) = \frac{\sigma-1}{\sigma}\frac{\psi(\omega)}{w^s B(1-u)^\varepsilon} = \ell(\omega), \tag{S.5}$$

so that employers hire the same number of workers for all of their (symmetric) jobs.

Combining the results yields $r(\omega) = A\left[mc(\omega)\right]^{1-\sigma}$, $mc(\omega) \equiv w/\big(\tilde{\varphi}(\omega)\{\eta+\pi[\nu n(\omega)+1]\}\big)$, $\lambda_j(\omega) = \lambda(\omega) = \eta+\pi\tilde{\beta}/b(\omega) = \{\eta+\pi[\nu n(\omega)+1]\}$—from eqs. (2), (5), and (8)—and

$$\lambda(\omega)w^e(\omega) = w^s B(1-u)^\varepsilon = \frac{\sigma-1}{2\sigma-1}\frac{r(\omega)}{\ell(\omega)} \equiv w. \tag{S.6}$$

Moreover, combing $\psi(\omega) = [\sigma/(2\sigma-1)]r(\omega)$ and eqs. (S.4), (S.5), we compute $\psi(\omega) = r(\omega)/(2\sigma-1) - w^s\tilde{\zeta}(\omega)\left\{\eta+\pi[\nu n(\omega)+1]\right\}^\gamma - w^s\tilde{\zeta}(\omega)f_0$. The optimal count of occupations is then determined by maximizing $\psi(\omega)$ with respect to $n(\omega)$, which yields $r(\omega)(\sigma-1)/[\gamma(2\sigma-1)] = w^s\tilde{\zeta}(\omega)\left\{\eta+\pi[\nu n(\omega)+1]\right\}^\gamma$. The zero-cutoff profit condition then establishes $r(\omega)/\tilde{\zeta}(\omega) = [(2\sigma-1)\xi f_0/(\sigma-1)]w^s$ and can be solved for $f_0(\sigma-1)/(\gamma-\sigma+1) = \{\eta+\pi[\nu n(\omega)+1]\}^\gamma$. The rest of the analysis follows as in the main text in Section 3.

However, the derivations of equilibrium in the closed (Section 4) and open economy (Section 5) differ because, under Stole-Zwiebel bargaining, there is unemployment in equilibrium. Risk-neutral workers must be indifferent between applying for jobs in the production sector (with an ex-ante expected wage w) or providing service inputs at a pay w^s (which is associated with self-employment). The unemployment rate of production workers is then given by the requirement that $w^s = (1-u)w$, establishing $B(1-u)^{1+\varepsilon} = 1$ from eq. (S.6). This equal-pay condition implies for the employment rate $1-u = B^{-1/(1-\varepsilon)} < 1$, which is a constant in our model because labor is used for production as well as services provision. Finally, we need to check that the wages paid to production

workers are (weakly) higher than their expected income outside the job $(1-u)w$. The wage of the least productive worker at employer ω is given by

$$w(0,\omega) = \frac{w\lambda(0,\omega)}{\lambda(\omega)} = w\frac{\eta + 2[\nu n(\omega)+1]}{\eta + \pi[\nu n(\omega)+1]} \equiv \underline{w}(n(\omega)). \tag{S.7}$$

Note that $\underline{w}'(n(\omega)) < 0$ and that $\lim_{n(\omega)\to\infty} \underline{w}(n(\omega)) = 2w/\pi$. It follows that $\underline{w}(n(\omega)) > (1-u)w$ is satisfied for all employers if $B < (\pi/2)^{1+\varepsilon}$. In this case, no workers who is matched to a production job will quit ex post. Therefore, we can maintain the parameter constraint $B > (\pi/2)^{1+\varepsilon}$ throughout our extended analysis.

S2.2 Variance of log revenues in the canonical model

We consider the canonical model with deterministic fixed costs equal to $\tilde{\delta}(\omega) = 1$. Moreover, we consider the case of selection into exporting by postulating $f_x^0 > f_0(t-1) > 0$. In this case, the fraction of exporting plants is given by $\chi = \Phi(-\ln\hat{\theta}_x^*/\sigma_u)/\Phi(-\ln\hat{\theta}_1^*/\sigma_u)$. Moreover, making use of $\mu_Y^d \equiv \ln[w\sigma\xi f_0/(\sigma-1)] - \ln\hat{\theta}_1^*$, $\mu_Y^e = \mu_Y^d + \ln t$, $u = \xi\ln\tilde{\theta}(\omega)$, and $v = \ln\hat{\zeta}(\omega)$ as well as eqs. (15) and (16) from the main text to substitute for $r^d(\omega)/\tilde{\zeta}(\omega)$, we can compute the average log revenues of non-exporters as follows:

$$\mathbb{E}\left[y\middle| u \geq \ln\hat{\theta}_1^*, u < \ln\hat{\theta}_x^*\right] = \mu_Y^d - (\sigma_u + \rho_{uv}\sigma_v)\frac{\phi\left(\ln\hat{\theta}_x^*/\sigma_u\right) - \phi\left(\ln\hat{\theta}_1^*/\sigma_u\right)}{\Phi\left(\ln\hat{\theta}_x^*/\sigma_u\right) - \Phi\left(\ln\hat{\theta}_1^*/\sigma_u\right)}.$$

Similarly, making use of $\mu_Y^e = \mu_Y^d + \ln t$, $u = \xi\ln\tilde{\theta}(\omega)$, and $v = \ln\hat{\zeta}(\omega)$ as well as eqs. (15), (16), and (24) to substitute for $r^e(\omega)/\tilde{\zeta}(\omega)$, we can compute the average log revenues of exporters as follows:

$$\mathbb{E}\left[y\middle| u \geq \ln\hat{\theta}_x^*\right] = \mu_Y^e + (\sigma_u + \rho_{uv}\sigma_v)\frac{\phi\left(\ln\hat{\theta}_x^*/\sigma_u\right)}{\Phi\left(-\ln\hat{\theta}_x^*/\sigma_u\right)}.$$

In a next step, we determine the variance of log revenues. Making use of $\tilde{\sigma}_v \equiv \sigma_v\sqrt{1-\rho_{uv}^2}$, we compute the second uncentered moment of log revenues of non-exporters:

$$\mathbb{E}\left[y^2\middle| u \geq \ln\hat{\theta}_1^*, u < \ln\hat{\theta}_X^*\right] = \left(\mu_Y^d\right)^2 - 2\mu_Y^d(\sigma_u + \rho_{uv}\sigma_v)\frac{\phi\left(\ln\hat{\theta}_x^*/\sigma_u\right) - \phi\left(\ln\hat{\theta}_1^*/\sigma_u\right)}{\Phi\left(\ln\hat{\theta}_x^*/\sigma_u\right) - \Phi\left(\ln\hat{\theta}_1^*/\sigma_u\right)}$$

$$+\tilde{\sigma}_v^2+(\sigma_u+\rho_{uv}\sigma_v)^2\left[1-\frac{\frac{\ln\hat{\theta}_x^*}{\sigma_u}\phi\left(\frac{\ln\hat{\theta}_x^*}{\sigma_u}\right)-\frac{\ln\hat{\theta}_1^*}{\sigma_u}\phi\left(\frac{\ln\hat{\theta}_1^*}{\sigma_u}\right)}{\Phi\left(\ln\hat{\theta}_x^*/\sigma_u\right)-\Phi\left(\ln\hat{\theta}_1^*/\sigma_u\right)}\right]$$

The variance of log revenues of non-exporters is then computed by $Var\left(y\middle|u\geq\ln\hat{\theta}_1^*,u<\ln\hat{\theta}_x^*\right)=\mathbb{E}\left(y^2\middle|u\geq\ln\hat{\theta}_1^*,u<\ln\hat{\theta}_x^*\right)-\left[\mathbb{E}\left(y\middle|u\geq\ln\hat{\theta}_1^*,u<\ln\hat{\theta}_x^*\right)\right]^2\equiv\mathbb{V}^d$ and follows as

$$\mathbb{V}^d=\tilde{\sigma}_v^2+(\sigma_u+\rho_{uv}\sigma_v)^2\left[1-\frac{\frac{\ln\hat{\theta}_x^*}{\sigma_u}\phi\left(\frac{\ln\hat{\theta}_x^*}{\sigma_u}\right)-\frac{\ln\hat{\theta}_1^*}{\sigma_u}\phi\left(\frac{\ln\hat{\theta}_1^*}{\sigma_u}\right)}{\Phi\left(\frac{\ln\hat{\theta}_x^*}{\sigma_u}\right)-\Phi\left(\frac{\ln\hat{\theta}_1^*}{\sigma_u}\right)}\right]$$
$$-(\sigma_u+\rho_{uv}\sigma_v)^2\left[1-\frac{\phi\left(\frac{\ln\hat{\theta}_x^*}{\sigma_u}\right)-\phi\left(\frac{\ln\hat{\theta}_1^*}{\sigma_u}\right)}{\Phi\left(\frac{\ln\hat{\theta}_x^*}{\sigma_u}\right)-\Phi\left(\frac{\ln\hat{\theta}_1^*}{\sigma_u}\right)}\right]^2.$$

For exporters, the second uncentered moment of log revenues is given by

$$\mathbb{E}\left[y^2\middle|u\geq\ln\hat{\theta}_1^*,u<\ln\hat{\theta}_x^*\right]=\left(\mu_Y^e\right)^2-2\mu_Y^e(\sigma_u+\rho_{uv}\sigma_v)\frac{\phi\left(\ln\hat{\theta}_x^*/\sigma_u\right)}{\Phi\left(-\ln\hat{\theta}_x^*/\sigma_u\right)}$$
$$+\tilde{\sigma}_v^2+(\sigma_u+\rho_{uv}\sigma_v)^2\left[1+\frac{\ln\hat{\theta}_x^*}{\sigma_u}\frac{\phi\left(\ln\hat{\theta}_x^*/\sigma_u\right)}{\Phi\left(-\ln\hat{\theta}_x^*/\sigma_u\right)}\right].$$

The corresponding variance is $Var\left(y\middle|u\geq\ln\hat{\theta}_x^*\right)=\mathbb{E}\left[y^2\middle|u\geq\ln\hat{\theta}_x^*\right]-\left\{\mathbb{E}\left[y\middle|u\geq\ln\hat{\theta}_x^*\right]\right\}^2\equiv\mathbb{V}^e$:

$$\mathbb{V}^e=\tilde{\sigma}_v^2+(\sigma_u+\rho_{uv}\sigma_v)^2\left[1+\frac{\ln\hat{\theta}_x^*}{\sigma_u}\frac{\phi\left(\ln\hat{\theta}_x^*/\sigma_u\right)}{\Phi\left(-\ln\hat{\theta}_x^*/\sigma_u\right)}\right]$$
$$-(\sigma_u+\rho_{uv}\sigma_v)^2\left[1+\frac{\phi\left(\ln\hat{\theta}_x^*/\sigma_u\right)}{\Phi\left(-\ln\hat{\theta}_x^*/\sigma_u\right)}\right]^2.$$

Rearranging terms, we can show that $\mathbb{V}^e>(<)\,\mathbb{V}^d$ iff $T(k_x,k_1)\equiv[1-\Phi(k_x)][\Phi(k_x)-\Phi(k_1)]^{-1}\left[H\left(k_x\right)-H\left(k_1\right)\right]^2>(<)\,H'\left(k_x\right)-H'\left(k_1\right)+2\left[H\left(k_x\right)-H\left(k_1\right)\right]$, where $H(\cdot)$ is the hazard rate of the standard normal distribution and $k_x\equiv\ln\hat{\theta}_x^*/\sigma_u$, $k_1\equiv\ln\hat{\theta}_1^*/\sigma_u$. We next show that (i) $\chi>1/2$ if $k_x<0$ and that (ii) $T(k_x,k_1)\leq 2\left[H\left(k_x\right)-H\left(k_1\right)\right]$ if $k_x>0$, with (ii) establishing $\mathbb{V}^e<\mathbb{V}^d$. This implies that $\chi\leq 1/2$ is inconsistent with

$\mathbb{V}^e < \mathbb{V}^d$ in the canonical model with $\tilde{\delta}(\omega) = 1$.

To establish part (i), we can note that $\chi \leq 1/2$ is equivalent to $\chi/(1-\chi) \leq 1$ and thus $\Phi(k_x) \geq [1+\Phi(k_1)]/2$, so that $\chi \leq 1/2$ is impossible if $\Phi(k_x) < 1/2$ and thus $k_x < 0$. To establish part (ii), we can note that $T(k_x, k_1) \leq 2\left[H\left(k_x\right) - H\left(k_1\right)\right]$ is equivalent to $\tilde{T}(k_x; k_1) \equiv 2\left[\Phi(k_x) - \Phi(k_1)\right] - \phi(k_x) + [1-\Phi(k_x)]H(k_1) \geq 0$. To confirm this inequality, we first determine a critical level k_x, denoted by $\tilde{k}_x$ that is implicitly determined as a function of k_1 by $\chi = \left[1-\Phi\left(\tilde{k}_x\right)\right]/\left[1-\Phi\left(k_1\right)\right] = 1/2$. We compute $\lim_{k_1 \to -\infty} \tilde{k}_x = 1/2$, $\lim_{k_1 \to \infty} \tilde{k}_x = \infty$. Moreover, applying the implicit function theorem establishes $d\tilde{k}_x/dk_1 = \phi(k_1)/\left[2\phi\left(\tilde{k}_x\right)\right] > 0$. We can therefore conclude that $\chi \leq 1/2$ implies $k_x \geq \tilde{k}_x$ and thus[7] $\tilde{T}(k_x, k_1) \geq \tilde{T}(\tilde{k}_x, k_1) = 1 - \Phi(k_1) - \phi\left(\tilde{k}_x\right) + \phi(k_1)/2 \equiv \tilde{T}_0(k_1)$, with $\tilde{T}_0'(k_1) = [\phi(k_1)/2](\tilde{k}_x - k_1 - 2)$. We can alternatively write $\tilde{T}_0(k_1) = [1-\Phi(k_1)][1+H(k_1)/2] - \phi\left(\tilde{k}_x\right)$, which is positive for all $k_1 \leq 0$. Moreover, it must be true that $\tilde{k}_x < k_1 + 2$, implying $\tilde{T}_0'(k_1) < 0$ for all $k_1 > 0$. To see this, we rewrite condition $\chi = 1/2$ as $\Gamma(\tilde{k}_x, k_1) = 1 + \Phi(k_1) - \Phi\left(\tilde{k}_x\right) = 0$, with $\partial\Gamma(\cdot)/\partial\tilde{k}_x < 0$. Setting $\tilde{k}_x = k_1 + 2$, we obtain $\Gamma\left(k_1+2, k_1\right) = 1 + \Phi(x) - 2\Phi(x+2) \equiv \tilde{\Gamma}(k_1)$, with $\Gamma(0) = 3/2 - 2\Phi(2) < 0$, $\lim_{k_1 \to \infty} \tilde{\Gamma}(k_1) = 0$, and $\tilde{\Gamma}'(k_1) > 0$, establishing $\Gamma\left(k_1+2, k_1\right) < 0$ for all finite $k_1 > 0$ This implies $k_x < k_1 + 2$ and thus $\tilde{T}_0'(k_1) < 0$ if $k_1 > 0$, which—recollecting $\tilde{T}_0(0) > 0$ and $\lim_{k_1 \to \infty} \tilde{T}_0(k_1) = 0$—is sufficient for $\tilde{T}(k_x, k_1) > 0$ to hold if $\chi \leq 1/2$. This shows part (ii) and completes the proof.

S2.3 Derivation details for the free-entry condition in eq. (28)

We first determine average profits (per entrant into the lottery), beginning with non-exporters. Making use of eqs. (15) and (16), we compute $[(\sigma-1)/(\sigma\xi)]\bar{r}\left(\exp[u]\middle|\exp[v]\right) = wf_0 \exp[u]/\hat{\theta}_1^*$, where $\bar{r}(\omega)$ is the revenue of plant ω conditional on its draw of $\tilde{\zeta}(\omega)$ and $u = \xi \ln \tilde{\theta}(\omega)$, $v = \ln \hat{\delta}(\omega)$ have been considered. We can then express the operating profits of non-exporters as $wf_0(\hat{\theta}/\hat{\theta}_1^* - 1)$, with operating profits of the marginal non-exporter being equal to zero. We introduce the auxiliary variables

$$g_{ij} \equiv \frac{1}{2\pi\sigma_i\sigma_j\sqrt{1-\rho_{ij}^2}} \exp\left\{-\frac{1}{2(1-\rho_{ij}^2)}\left[\left(\frac{i}{\sigma_i}\right)^2 - 2\rho_{ij}\frac{ij}{\sigma_i\sigma_j} + \left(\frac{j}{\sigma_j}\right)^2\right]\right\}$$

[7] Note that $\partial\tilde{T}(\cdot)/\partial k_x = \phi(k_x)[2 + k_x - H(k_1)]$. Then making use of $2 + k_x - H(k_1) > 2 + k_x - H(k_x)$, $H\prime(\cdot) \in (0,1)$ and $2 - H(0) = 2(1 - 1/\sqrt{2\pi}) > 0$, it follows that $\partial\tilde{T}(\cdot)/\partial k_x > 0$.

and $\bar{g}_i \equiv \frac{1}{\sqrt{2\pi}\bar{\sigma}_i}\exp\left[-\frac{1}{2}\left(\frac{i-\bar{\mu}_i}{\bar{\sigma}_i}\right)^2\right]$. Then, aggregating over all non-exporters delivers $\bar{\psi}_0 \equiv \bar{\psi}_0^1 - \bar{\psi}_0^2$, with

$$\bar{\psi}_0^1 \equiv \frac{w f_0}{\hat{\theta}_1^*}\int_{\ln\delta_x}^{\infty}\int_{\ln\hat{\theta}_1^*}^{\ln\hat{\theta}_x^*(\exp[z])}\exp\left[u\right]g_{uz}\int_{-\infty}^{\infty}\exp[v]\bar{g}_v\,\mathrm{d}v\,\mathrm{d}u\,\mathrm{d}z,$$
$$\bar{\psi}_0^2 \equiv w f_0\int_{\ln\delta_x}^{\infty}\int_{\ln\hat{\theta}_1^*}^{\ln\hat{\theta}_x^*(\exp[z])}g_{uz}\int_{-\infty}^{\infty}\exp[v]\bar{g}_v\,\mathrm{d}v\,\mathrm{d}u\,\mathrm{d}z,$$

and

$$\bar{\mu}_v = u\frac{\sigma_v}{\sigma_u}\frac{\rho_{uv}-\rho_{uz}\rho_{vz}}{1-\rho_{uz}^2} + z\frac{\sigma_v}{\sigma_z}\frac{\rho_{vz}-\rho_{uv}\rho_{uz}}{1-\rho_{uz}^2}, \quad \bar{\sigma}_v \equiv \sqrt{\frac{1-\rho_{uv}^2-\rho_{uz}^2-\rho_{vz}^2+2\rho_{uv}\rho_{uz}\rho_{vz}}{1-\rho_{uz}^2}}.$$

Solving the integrals, we compute $\bar{\psi}_0 = w f_0\int_{\ln\delta_x}^{\infty}\left\{\left[\tilde{F}(z,\hat{\theta}_1^*,1)/\hat{\theta}_1^* - \tilde{F}(z,\hat{\theta}_1^*,0)\right] - \left[\tilde{F}(z,\hat{\theta}_x^*(\exp[z]),1)/\hat{\theta}_x^*(\exp[z]) - \tilde{F}(z,\hat{\theta}_x^*(\exp[z]),0)\right]\right\}\mathrm{d}z$

We next consider exporters with a draw of $\tilde{\delta}(\omega) \leq \delta_x$. Following our reasoning above, we can express the profits of exporters as $w(f_o + f_x\exp[z])(\hat{\theta}/\hat{\theta}_0^* - 1)$, where profits of the marginal exporter, drawing $z = \ln\tilde{\delta}(\omega) = -\infty$ and an augmented productivity level of $u = \ln\hat{\theta}_0^*$, are equal to zero. We can then express the average profit (per entrant into the lottery) of plants with a fixed cost draw $\tilde{\delta}(\omega) \leq \delta_x$ as $\bar{\psi}_1 \equiv \bar{\psi}_1^1 - \bar{\psi}_1^2$, with

$$\bar{\psi}_1^1 \equiv w\int_{-\infty}^{\ln\delta_x}\int_{\ln\hat{\theta}^*(\exp[z])}^{\infty}\frac{\exp\left[u\right]}{\hat{\theta}^*(\exp[z])}(f_0 + f_x\exp[z])g_{uz}\,\exp[v]\bar{g}_v\,\mathrm{d}v\,\mathrm{d}u\,\mathrm{d}z,$$
$$\bar{\psi}_1^2 \equiv w\int_{-\infty}^{\ln\delta_x}\int_{\ln\hat{\theta}^*(\exp[z])}^{\infty}(f_0 + f_x\exp[z])g_{uz}\int_{-\infty}^{\infty}\exp[v]\bar{g}_v\,\mathrm{d}v\,\mathrm{d}u\,\mathrm{d}z.$$

Solving the integrals gives $\bar{\psi}_1 = w\int_{-\infty}^{\ln\delta_x}(f_0+f_x\exp[z])\left[\tilde{F}(z,\hat{\theta}^*(\exp[z]),1)/\hat{\theta}^*(\exp[z]) - \tilde{F}(z,\hat{\theta}^*(\exp[z]),0)\right]\mathrm{d}z$.

In a final step, we consider exporters with a draw of $\tilde{\delta}(\omega) \leq \delta_x$, which happen to be plants with $u > \ln\hat{\theta}_x^*(\exp[z])$. We can express the average profits of such exporters (per entrant into the lottery) as $\bar{\psi}_2 \equiv \bar{\psi}_2^1 - \bar{\psi}_2^2$, with

$$\bar{\psi}_2^1 \equiv w f_0\frac{t}{\hat{\theta}_1^*}\int_{\ln\delta_x}^{\infty}\int_{\ln\hat{\theta}_x^*(\exp[z])}^{\infty}\exp[u]g_{uz}\int_{-\infty}^{\infty}\exp[v]g_v\,\mathrm{d}v\,\mathrm{d}u\,\mathrm{d}z,$$
$$\bar{\psi}_2^2 \equiv w\int_{\ln\delta_x}^{\infty}\int_{\ln\hat{\theta}_x^*(\exp[z])}^{\infty}(f_0 + f_x^0\exp[z])g_{uz}\int_{-\infty}^{\infty}\exp[v]g_v\,\mathrm{d}v\,\mathrm{d}u\,\mathrm{d}z,$$

Solving the integrals gives $\bar{\psi}_2 = w\int_{\ln\delta_x}^{\infty}\left[t f_0 F(z,\ln\hat{\theta}_x^*(\exp[z]),1)/\hat{\theta}_1^* - (f_0 + f_x^0\exp[z])\right.$

$\times F(z, \ln \hat{\theta}_x^*(\exp[z]), 0)\big]\, \mathrm{d}z$. We can now add up the three parts of average profits, use $\hat{\theta}_x^* = \hat{\theta}_1^* f_x \exp[z]/[(t-1)f_0]$ and set the resulting expression equal to wf_e to compute the free entry condition in eq. (28).

We complete our analysis here with a discussion of the properties of the free entry condition. For this purpose, we first differentiate the left-hand side of eq. (28) with respect to $\hat{\theta}_1^*$, acknowledging $\hat{\theta}^* = \hat{\theta}_1^*(1 + f_x \exp[z])/t$ and $\hat{\theta}_x^* = \hat{\theta}_1^* f_x \exp[z]/[(t-1)f_0]$. This gives $\partial LHS_{eq.\ (28)}/\partial \hat{\theta}_1^* = -wf_0 \left(\frac{1}{\hat{\theta}_1^*}\right)^2 \int_{\ln \delta_x}^{\infty} \left[\tilde{F}(z, \hat{\theta}_1^*, 1) + (t-1)\tilde{F}(z, \hat{\theta}_x^*(\exp[z]), 1)\right] \mathrm{d}z + \int_{-\infty}^{\ln \delta_x} t\tilde{F}(z, \hat{\theta}^*(\exp[z]), 1)\, \mathrm{d}z < 0$, with $LHS_{eq.\ (28)}$ falling from infinity if $\hat{\theta}_1^* = 0$ to zero if $\hat{\theta}_1^* \to \infty$. Moreover, computing the partial derivative with respect to f_x, we obtain $\partial LHS_{eq.\ (28)}/\partial f_x = -w \int_{\ln \delta_x}^{\infty} \tilde{F}(z, \hat{\theta}_x^*(\exp[z]), 0)\, \mathrm{d}z + \int_{-\infty}^{\ln \delta_x} \tilde{F}(z, \hat{\theta}^*(\exp[z]), 0)\, \mathrm{d}z < 0$. Applying the implicit function theorem, we compute $d\hat{\theta}_1^*/df_x < 0$, while for an interior solution implicitly determined by eq. (28), it must be true that $\hat{\varepsilon} \equiv -d\hat{\theta}_1^*/df_x \times f_x/\hat{\theta}_1^* \in (0, 1)$.

We finally acknowledge $\hat{\theta}_1^* = \hat{\theta}_a^*$—leaving the formal proof for this intuitive result to the interested reader—and determine $\hat{\theta}_1^*$ for the liming case of $f_x \to 0$. We first compute $\lim_{f_x \to 0} \int_{\ln \delta_x}^{\infty} \left\{\tilde{F}(z, \hat{\theta}_1^*, 1)/\hat{\theta}_1^* - \tilde{F}(z, \hat{\theta}_1^*, 0)\right\} \mathrm{d}z = \int_{\infty}^{\infty} \left\{\tilde{F}(z, \hat{\theta}_1^*, 1)/\hat{\theta}_1^* - \tilde{F}(z, \hat{\theta}_1^*, 0)\right\} \mathrm{d}z = 0$. In a second step, we note that $\hat{\theta}_x^*(\exp[z]) = \hat{\theta}_1^* f_x \exp[z]/[f_0(t-1)]$ implies $\lim_{f_x \to 0} \hat{\theta}_x^*(\exp[z]) = 0$ and thus $\lim_{f_x \to 0} \tilde{F}(z, \hat{\theta}_x^*(\exp[z]), 1) = \tilde{F}(z, 0, 1)$ for any finite $\hat{\theta}_1^* > 0$. We then have $\lim_{f_x \to 0} \int_{\ln \delta_x}^{\infty} (f_x \exp[z]/f_0) \left\{\tilde{F}(z, \hat{\theta}_x^*(\exp[z]), 1)/\hat{\theta}_x^*(\exp[z]) - \tilde{F}(z, \hat{\theta}_x^*(\exp[z]), 0)\right\} \mathrm{d}z = \int_{\infty}^{\infty} [(t-1)/\hat{\theta}_1^*] \tilde{F}\,(z, 0, 1)\ \mathrm{d}z = 0$, where the second equality presumes $\hat{\theta}_1^* > 0$. In a third step, we acknowledge $\hat{\theta}^*(\exp[z]) = \hat{\theta}_1^*(1 + f_x \exp[z]/f_0)$, so that $\lim_{f_x \to 0} \hat{\theta}^*(\exp[z]) = \hat{\theta}_1^*$. We can then write

$$\int_{-\infty}^{\ln \delta_x} \left(1 + \frac{f_x \exp[z]}{f_0}\right) \left[\frac{1}{\hat{\theta}^*(\exp[z])} \tilde{F}\left(z, \hat{\theta}_x^*(\exp[z]), 1\right) - \tilde{F}\left(z, \hat{\theta}^*(\exp[z]), 0\right)\right] \mathrm{d}z$$
$$= \int_{-\infty}^{\infty} \left[\frac{1}{\hat{\theta}_1^*} \tilde{F}\left(z, \hat{\theta}_1^*, 1\right) - \tilde{F}\left(z, \hat{\theta}_1^*, 0\right)\right] dz = \frac{1}{\hat{\theta}_1^*} F\left(\hat{\theta}_1^*, 1\right) - F\left(\hat{\theta}_1^*, 0\right) = 0$$

Adding up the three integrals, we observe that $\lim_{f_x \to 0} LHS_{eq.\ (28)} = F(\hat{\theta}_1^*, 1)//\hat{\theta}_1^* - F(\hat{\theta}_1^*, 0)$, which establishes $\hat{\theta}_1^* = \hat{\theta}_a^*$ if $f_x \to 0$ from contrasting eqs. (19) and (28). This completes the proof.

S2.4 Derivation details for Λ and CV in eq. (33)

We organize the proof in four steps.

S2.4.1 Variable labor input L_v

We first compute total variable labor input in non-exporting and exporting plants, L_v^d and L_v^e, respectively. Starting with non-exporters, we compute

$$L_v^d = N\xi f_0(\hat{\theta}_1^*)^{-1} \int_{\ln \delta_x}^{\infty} \int_{\ln \hat{\theta}_1^*}^{\ln \hat{\theta}_x^*(\exp[z])} \exp[u]\, g_{uz} \int_{-\infty}^{\infty} \exp[v] \bar{g}_v \,\mathrm{d}v \,\mathrm{d}u \,\mathrm{d}z,$$

where ξf_0 is employment of the marginal non-exporter, N is the mass of plants entering the lottery, and $g_{uz}, \bar{g}_v$ are given in Supplement S2.3. This can be solved for $L_v^d = N\xi f_0(\hat{\theta}_1^*)^{-1} \int_{\ln \delta_x}^{\infty} \left\{ \tilde{F}(z, \ln \theta_1^*, 1) - \tilde{F}(z, \ln \theta_x^*(\exp[z]), 1) \right\} dz$.

Similarly, total variable labor input in exporting plants can be computed according to

$$L_v^e = N\xi f_0(\hat{\theta}_0^*)^{-1} \left\{ \int_{-\infty}^{\ln \delta_x} \int_{\ln \hat{\theta}^*(\exp[z]}^{\infty} \exp[u]\, g_{uv} \int_{-\infty}^{\infty} \exp[v] \bar{g}_v \,\mathrm{d}v \,\mathrm{d}u \,\mathrm{d}z \right.$$
$$\left. + \int_{\ln \delta_x}^{\infty} \int_{\ln \hat{\theta}_x^*(\exp[z])}^{\infty} \exp[u]\, g_{uz} \int_{-\infty}^{\infty} \exp[v] \bar{g}_v \,\mathrm{d}v \,\mathrm{d}u \,\mathrm{d}z \right\},$$

where ξf_0 is also employment of the marginal exporter. Solving the integrals, we compute $L_v^e = N\xi f_0(\hat{\theta}_1^*)^{-1} \int_{\ln \delta_x}^{\infty} t \left\{ \tilde{F}(z, \ln \theta_x^*(\exp[z]), 1) - \tilde{F}(z, \ln \theta^*(\exp[z]), 1) \right\} \mathrm{d}z$, where $\hat{\theta}_1^* = t\hat{\theta}_0^*$ has been used. Total variable labor input of all plants is therefore given by $L_v = N\xi f_0(\hat{\theta}_1^*)^{-1} \int_{\ln \delta_x}^{\infty} \left\{ \tilde{F}(z, \ln \theta_1^*, 1) + (t-1)\tilde{F}(z, \ln \theta_x^*(\exp[z]), 1) - t\tilde{F}(z, \ln \theta^*(\exp[z]), 1) \right\}] \,\mathrm{d}z$.

S2.4.2 Economy-wide worker efficiency Λ

We first compute average worker efficiency of non-exporters, Λ^d. Averaging $\lambda^d(\hat{\theta}) = \left[(\xi f_0 \hat{\theta}) / (\gamma \hat{\theta}_1^*) \right]^{1/\gamma}$ over non-exporters using their employment levels as weights, gives

$$\Lambda^d = \frac{N\xi f_0}{\hat{\theta}_1^* L_v^d} \left(\frac{\gamma \hat{\theta}_1^*}{\xi f_0} \right)^{-\frac{1}{\gamma}} \int_{\ln \delta_x}^{\infty} \left\{ \tilde{F}(z, \hat{\theta}_1^*, 1 + 1/\gamma) - \tilde{F}(z, \hat{\theta}_x^*(\exp[z]), 1 + 1/\gamma) \right\} \mathrm{d}z.$$

We next determine average worker efficiency among exporters, Λ^e. Averaging $\lambda^e(\hat{\theta}) = \left[(\xi f_0 \hat{\theta}) / (\gamma \hat{\theta}_0^*) \right]^{1/\gamma}$ over non-exporters using their employment levels as weights, gives

$$\Lambda^e = \frac{N\xi f_0}{\hat{\theta}_1^* L_v^e} \left(\frac{\gamma \hat{\theta}_1^*}{\xi f_0} \right)^{-\frac{1}{\gamma}} \int_{\ln \delta_x}^{\infty} t^{1+\frac{1}{\gamma}} \left\{ \tilde{F}(z, \hat{\theta}_x^*(\exp[z]), 1 + 1/\gamma) \right.$$
$$\left. - \tilde{F}(z, \hat{\theta}^*(\exp[z]), 1 + 1/\gamma) \right\} \mathrm{d}z,$$

We can finally compute the economy-wide average of worker efficiency according to $\Lambda = (L_v^d/L_v)\Lambda^d + (L_v^e/L_v)\Lambda^e$, which making use of auxiliary function $G(\hat{\theta}_1^*, f_x/f_0, 1+1/\gamma)$ from the main text can be solved for Λ in eq. (33).

S2.4.3 Economy-wide wage dispersion CV

We first compute the average wage dispersion of non-exporters, CV^d. Using $\mathcal{X} \equiv \sqrt{4-\pi(\pi-2)}/\pi$ and averaging $cv^d(\hat{\theta}) = \mathcal{X}\left\{1-\eta\left[\gamma\hat{\theta}_1^*/(\xi f_0\hat{\theta})\right]^{1/\gamma}\right\}$ over non-exporters using their employment levels as weights, we compute

$$CV^d = \frac{N\xi f_0\mathcal{X}}{\hat{\theta}_1^* L_v^d}\left\{\int_{\ln\delta_x}^{\infty}\left[\tilde{F}(z,\hat{\theta}_1^*,1)-\tilde{F}(z,\hat{\theta}_x^*(\exp[z]),1)\right]\mathrm{d}z - \eta\left(\frac{\gamma\hat{\theta}_1^*}{\xi f_0}\right)^{\frac{1}{\gamma}}\int_{\ln\delta_x}^{\infty}\left[\tilde{F}(z,\hat{\theta}_1^*,1-1/\gamma)-\tilde{F}(z,\hat{\theta}_x^*(\exp[z]),1-1/\gamma)\right]\mathrm{d}z\right\}.$$

We next determine average wage dispersion among exporters, CV^e. Averaging $cv^e(\hat{\theta}) = \mathcal{X}\left\{1-\eta\left[\gamma\hat{\theta}_0^*/(\xi f_0\hat{\theta})\right]^{1/\gamma}\right\}$ over non-exporters using their employment levels as weights, we compute

$$CV^e = \frac{N\xi f_0\mathcal{X}}{\hat{\theta}_1^* L_v^e}\left\{\int_{\ln\delta_x}^{\infty} t\left[\tilde{F}(z,\hat{\theta}_x^*(\exp[z]),1)-\tilde{F}(z,\hat{\theta}^*(\exp[z]),1)\right]\mathrm{d}z - \eta\left(\frac{\gamma\hat{\theta}_1^*}{\xi f_0}\right)^{\frac{1}{\gamma}}\int_{\ln\delta_x}^{\infty} t^{1+\frac{1}{\gamma}}\left[\tilde{F}(z,\hat{\theta}_x^*(\exp[z]),1-1/\gamma)-\tilde{F}(z,\hat{\theta}^*(\exp[z]),1-1/\gamma)\right]\mathrm{d}z\right\}.$$

We can finally compute the economy-wide average of plant-occupation internal wage dispersion by $CV = (L_v^d/L_v)CV^d + (L_v^e/L_v)CV^e$, which making use of auxiliary function $G(\hat{\theta}_1^*, f_x/f_0, 1-1/\gamma)$ from the main text can be solved for CV in eq. (33).

S2.4.4 Properties of Λ and CV

As before, we leave the formal proof of the intuitive results $\lim_{f_x\to\infty}\Lambda = \Lambda_a$, $\lim_{f_x\to\infty} CV = CV_a$ to the interested reader and focus on $f_x \to 0$. In this case, we have $\lim_{f_x\to 0}\hat{\theta}^*(\exp[z]) = \hat{\theta}_0^* = \hat{\theta}_a^*$ and $\lim_{f_x\to 0}\int_{\ln\delta_x}^{\infty}\tilde{F}(z,\hat{\theta}^*,\iota)dz = \int_{-\infty}^{\infty}\tilde{F}(z,\hat{\theta}_a^*,\iota)dz = F(\hat{\theta}_a^*,\iota)$. Moreover acknowledging $\hat{\theta}_x^*(\exp[z]) = \hat{\theta}_1^* f_x \exp[z]/[(t-1)f_0]$, we compute $\lim_{f_x\to 0}\hat{\theta}_x^*(\exp[z]) = 0$, while $\lim_{f_x\to 0}\delta_x = \infty$. This is sufficient for $\lim_{f_x\to 0}\int_{\ln\delta_x}^{\infty}\tilde{F}(z,\hat{\theta}_x^*(\exp[z]),\iota)dz = \int_{\infty}^{\infty}\tilde{F}(z,0,\iota)dz = 0$. We can therefore conclude that $\lim_{f_x\to 0} G(\hat{\theta}_1^*, f_x/f_0, \iota) = F(\hat{\theta}_a^*,\iota)$, which establishes $\lim_{f_x\to 0}\Lambda = \Lambda_a$ and $\lim_{f_x\to 0} CV = CV_a$.

We now determine the derivative of $G(\hat{\theta}_1^*, f_x/f_0, \iota)$ with respect to f_x. Making use of auxiliary functions $A(\iota) \equiv \exp\left[\{(\iota\sigma_u)^2 + 2\iota\rho_{uv}\sigma_u\sigma_v + \sigma_v^2\}/2\right]$ and the auxiliary functions $B_0(z,\iota) \equiv z - \rho_{vz}\sigma_v\sigma_z - \iota\rho_{uz}\sigma_u\sigma_z$, and $B_1(z,x,\iota) \equiv \ln x - z\rho_{uz}\sigma_u/\sigma_z - \iota\sigma_u^2(1-\rho_{uz}^2) - \sigma_u\sigma_v(\rho_{uv} - \rho_{uz}\rho_{vz})$, we compute using $\hat{\varepsilon} = -d\hat{\theta}_1^*/df_x \times f_x/\hat{\theta}_1^* \in (0,1)$

$$\frac{d\int_{\ln\delta_x}^{\infty}\tilde{F}(z,\hat{\theta}_1^*,\iota)dz}{df_x} = \frac{A(\iota)}{f_x}\frac{1}{\sqrt{2\pi}\sigma_z}\exp\left[-\frac{1}{2}\left(\frac{B_0(\ln\delta_x,\iota)}{\sigma_z}\right)^2\right]\Phi\left(-\frac{B_1(z,\hat{\theta}_1^*,\iota}{\sigma_u}\right) + \frac{A(\iota)}{f_x}\int_{\ln\delta_x}^{\infty}\frac{\exp\left[-\frac{1}{2}\left(B_0(z,\iota)/\sigma_u\right)^2\right]}{2\pi\sigma_u\sigma_z\sqrt{1-\rho_{uz}^2}}\exp\left[-\frac{1}{2}\left(\frac{B_1(z,\hat{\theta}_1^*,\iota)}{\sigma_u\sqrt{1-\rho_{uz}^2}}\right)^2\right]\hat{\varepsilon}dz,$$

$$\frac{(t^\iota-1)d\int_{\ln\delta_x}^{\infty}\tilde{F}(z,\hat{\theta}_x^*,\iota)dz}{df_x} = \frac{(t^\iota-1)A(\iota)}{f_x}\frac{\exp\left[-\frac{1}{2}\left(B_0(\ln\delta_x,\iota)/\sigma_z\right)^2\right]}{\sqrt{2\pi}\sigma_z}\Phi\left(-\frac{B_1(z,\hat{\theta}_1^*,\iota)}{\sigma_u}\right) - \frac{(t^\iota-1)A(\iota)}{f_x}\int_{\ln\delta_x}^{\infty}\frac{\exp\left[-\frac{1}{2}\left(B_0(z,\iota)/\sigma_z\right)^2\right]}{2\pi\sigma_u\sigma_z\sqrt{1-\rho_{uz}^2}}\exp\left[-\frac{1}{2}\left(\frac{B_1(z,\hat{\theta}_x^*(\exp[z]),\iota)}{\sigma_u}\right)^2\right](1-\hat{\varepsilon})\,dz,$$

and

$$\frac{d\int_{-\infty}^{\ln\delta_x}t^\iota\tilde{F}(z,\hat{\theta}^*,\iota)dz}{df_x} = -\frac{t^\iota A(\iota)}{f_x}\frac{1}{\sqrt{2\pi}\sigma_z}\exp\left[-\frac{1}{2}\left(\frac{B_0(\ln\delta_x,\iota)}{\sigma_z}\right)^2\right]\Phi\left(-\frac{B_1(z,\hat{\theta}_1^*,\iota)}{\sigma_u}\right) - \frac{t^\iota A(\iota)}{f_x}\int_{-\infty}^{\ln\delta_x}\frac{\exp\left[-\frac{1}{2}\left(B_0(z,\iota)/\sigma_z\right)^2\right]}{2\pi\sigma_u\sigma_z\sqrt{1-\rho_{uz}^2}}\exp\left[-\frac{1}{2}\left(\frac{B_1(z,\hat{\theta}^*(\exp[z]),\iota)}{\sigma_u}\right)^2\right] \times\left[\frac{f_x\exp(z)}{f_0+f_x\exp(z)} - \hat{\varepsilon}\right]dz.$$

Adding up the three derivatives gives $dG(\hat{\theta}_1^*, f_x/f_0, \iota)/df_x$.

We now take the limit of $f_x \to 0$. This gives

$$\lim_{f_x\to 0}\frac{dG(\cdot)}{df_x} = -\frac{t^\iota A(\iota)}{f_0}\int_{-\infty}^{\infty}\frac{\exp\left[-\frac{1}{2}(B_0(z,\iota)/\sigma_z)^2\right]}{2\pi\sigma_u\sigma_z\sqrt{1-\rho_{uz}^2}} \times\exp\left[-\frac{1}{2}\left(\frac{B_1(z,\ln\hat{\theta}_0^*,\iota)}{\sigma_u}\right)^2\right][\exp[z]-\hat{\varepsilon}_a]\,dz,$$

with $\hat{\varepsilon}_a \equiv \lim_{f_x\to 0}\varepsilon = F(\hat{\theta}_a^*,0)/[(1/\hat{\theta}_a^*)F(\hat{\theta}_a^*,1)] \in (0,1)$. Substituting for $B_0(z,\iota)$ and

$B_1(z, \ln \hat{\theta}_0^*, \iota)$, we compute

$$\lim_{f_x \to 0} \frac{dG(\cdot)}{df_x} = -\frac{t^\iota A(\iota)}{f_0} \frac{1}{\sqrt{2\pi}\sigma_u} \exp\left[-\frac{1}{2}\left(\frac{\ln \hat{\theta}_0^* - \iota\sigma_u^2 - \rho_{uv}\sigma_u\sigma_v}{\sigma_u}\right)^2\right](\bar{a} - \hat{\varepsilon}_a),$$

with $\bar{a} \equiv \exp[\ln \hat{\theta}_0^* \rho_{uz}\sigma_z/\sigma_u + \sigma_v\sigma_z(\rho_{vz} - \rho_{uv}\rho_{uz}) + \sigma_z^2(1-\rho_{uz}^2)/2] > 0$.

With this insight at hand, we can then determine

$$\lim_{f_x \to 0} \frac{d\Lambda}{df_x} = \frac{\Lambda^a}{f_0}\left\{(g-1)\varepsilon_a - \frac{1}{\sigma_u} H\left(\frac{\ln \hat{\theta}_a^* - g\sigma_u^2 - \rho_{uv}\sigma_u\sigma_v}{\sigma_u}\right)(\bar{a} - \varepsilon_a) \right. \\ \left. + \frac{1}{\sigma_u} H\left(\frac{\ln \hat{\theta}_a^* - \sigma_u^2 - \rho_{uv}\sigma_u\sigma_v}{\sigma_u}\right)(\bar{a} - \varepsilon_a)\right\},$$

where $g \equiv 1 + 1/\gamma > 1$. Differentiating $\bar{\Gamma}(g) \equiv (g-1)\varepsilon_a - \frac{1}{\sigma_u} H\left(\frac{\ln \hat{\theta}_a^* - g\sigma_u^2 - \rho_{uv}\sigma_u\sigma_v}{\sigma_u}\right)(\bar{a} - \varepsilon_a)$ and noting that $H'(\cdot) \in (0,1)$ gives $\bar{\Gamma}'(g) = \varepsilon_a + H'\left(\frac{\ln \hat{\theta}_a^* - g\sigma_u^2 - \rho_{uv}\sigma_u\sigma_v}{\sigma_u}\right)(\bar{a} - \varepsilon_a) > 0$. Making use of $\bar{\Gamma}(1) = -\frac{1}{\sigma_u} H\left(\frac{\ln \hat{\theta}_a^* - \sigma_u^2 - \rho_{uv}\sigma_u\sigma_v}{\sigma_u}\right)(\bar{a} - \varepsilon_a)$, we therefore conclude $\lim_{f_x \to 0} d\Lambda/df_x > 0$. We can follow a similar line of reasoning to establish

$$\lim_{f_x \to 0} \frac{dCV}{df_x} = -\frac{\mathcal{X}\eta}{f_0}\left(\frac{\gamma t \hat{\theta}_a^*}{\xi f_0}\right)^{\frac{1}{\gamma}} \left\{(g-1)\varepsilon_a - \frac{1}{\sigma_u} H\left(\frac{\ln \hat{\theta}_a^* - g\sigma_u^2 - \rho_{uv}\sigma_u\sigma_v}{\sigma_u}\right)(\bar{a} - \varepsilon_a) \right. \\ \left. + \frac{1}{\sigma_u} H\left(\frac{\ln \hat{\theta}_a^* - \sigma_u^2 - \rho_{uv}\sigma_u\sigma_v}{\sigma_u}\right)(\bar{a} - \varepsilon_a)\right\},$$

with $g \equiv 1 - 1/\gamma < 1$. From $\bar{\Gamma}(1) = -\frac{1}{\sigma_u} H\left(\frac{\ln \hat{\theta}_a^* - \sigma_u^2 - \rho_{uv}\sigma_u\sigma_v}{\sigma_u}\right)(\bar{a} - \varepsilon_a)$, we now conclude that $\lim_{f_x \to 0} dCV/df_x > (<)\, 0$ if $\eta > (<)\, 0$. This completes the proof.

S2.5 The share of exporters in the model variant with plant-specific $\tilde{\delta}(\omega)$

We first compute the unconditional share of plants entering the lottery and choosing to export:

$$\text{sh}_x = \int_{-\infty}^{\ln \delta_x} \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left[-\frac{1}{2}\left(\frac{z}{\sigma_z}\right)^2\right] \Phi\left(-\frac{\ln \hat{\theta}^*(\exp[z]) - z\rho_{uz}\sigma_u/\sigma_z}{\sigma_u\sqrt{1-\rho_{uz}^2}}\right) dz \\ + \int_{\ln \delta_x}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left[-\frac{1}{2}\left(\frac{z}{\sigma_z}\right)^2\right] \Phi\left(-\frac{\ln \hat{\theta}_x^*(\exp[z]) - z\rho_{uz}\sigma_u/\sigma_z}{\sigma_u\sqrt{1-\rho_{uz}^2}}\right) dz,$$

which, making use of the definition of $\hat{F}(z,x)$ from the main text, corresponds to the numerator in eq. (35).

In a second step, we determine the share of entrants choosing to produce and compute

$$\text{sh} = \int_{-\infty}^{\ln \delta_x} \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left[-\frac{1}{2}\left(\frac{z}{\sigma_z}\right)^2\right] \Phi\left(-\frac{\ln \hat{\theta}^*(\exp[z]) - z\rho_{uz}\sigma_u/\sigma_z}{\sigma_u\sqrt{1-\rho_{uz}^2}}\right) dz + \int_{\ln \delta_x}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left[-\frac{1}{2}\left(\frac{z}{\sigma_z}\right)^2\right] \Phi\left(-\frac{\ln \hat{\theta}_1^* - z\rho_{uz}\sigma_u/\sigma_z}{\sigma_u\sqrt{1-\rho_{uz}^2}}\right) dz.$$

Making use of the definition of $\hat{F}(z,x)$ from the main text establishes the denominator in eq. (35). This completes the proof.

S3 Implementation of Estimation Model

For convenience, we estimate the transformed equation system

$$y = \begin{cases} \mu_Y^e + u + v & \text{if } I = 1 \\ \mu_Y^d + u + v & \text{if } I = 0 \end{cases}, \tag{S.8a}$$

$$o = \begin{cases} \mu_O^e - (1/\gamma)u & \text{if } I = 1 \\ \mu_O^d - (1/\gamma)u & \text{if } I = 0 \end{cases}, \tag{S.8b}$$

$$I = \begin{cases} 1 & \text{if } \mu_X + e \geq 0 \\ 0 & \text{if } \mu_X + e < 0 \end{cases}, \tag{S.8c}$$

$$y, o, I = \text{missing} \qquad \text{if } u < \ln \hat{\theta}_0^*, \tag{S.8d}$$

instead of (34), where $e = u - z$ is the composite of two stochastic variables that are jointly log-normal distributed. The joint normal distribution of the unobserved plant characteristics (disturbances) can then be stated as

$$(u, v, e)^T \sim \mathcal{N}_{\mathcal{T}}(\mathbf{0}, \tilde{\mathbf{\Sigma}}) \qquad \text{with} \quad \tilde{\mathbf{\Sigma}} = \begin{pmatrix} \sigma_u^2 & \rho_{uv}\sigma_u\sigma_v & \rho_{ue}\sigma_u \\ \rho_{uv}\sigma_u\sigma_v & \sigma_v^2 & \rho_{ve}\sigma_v \\ \rho_{ue}\sigma_u & \rho_{ve}\sigma_v & \sigma_e^2 \end{pmatrix},$$

with $\sigma_e \equiv \sigma_u\sqrt{1 - 2\rho_{uz}\sigma_z/\sigma_u + \sigma_z^2/\sigma_u^2}$, $\rho_{ue} \equiv \sigma_u/\sigma_e - \rho_{uz}\sigma_z/\sigma_e$, and $\rho_{ve} \equiv \rho_{uv}\sigma_u/\sigma_e - \rho_{vz}\sigma_z/\sigma_e$.

With the ML estimation of equation system (S.8) we aim to determine 14 parameters

of our theoretical model. These include the means $\mu_Y^e, \mu_Y^d, \mu_O^e, \mu_O^d, \mu_x$, the second moments of the three stochastic variables $\sigma_u, \sigma_v, \sigma_e$, ρ_{uv}, ρ_{ue}, ρ_{ve}, the two fundamental model parameters σ, γ, and the truncation point $\hat{\theta}_0^*$. As pointed out by Maddala (1986), one of the variance parameters remains undetermined by our estimator, and hence we set $\sigma_e = 1$. We can further reduce the set of parameters to be estimated by making use of the two functional relationships imposed by our model, namely *(i)* $\mu_Y^e - \mu_Y^d = -\gamma(\mu_O^e - \mu_O^d) = \ln t$ and *(ii)* $r^e(\omega)/r_T^e(\omega) = t^{(1-\sigma)/\xi} \equiv 1 - expsh$, where $expsh$ is the average share of foreign to total sales of exporters at the plant level which is observable in our data. We can now combine the two functional relationships to pin down σ by other model parameters:

$$\sigma = 1 + \gamma - \frac{\ln(1 - expsh)}{\mu_O^e - \mu_O^d}.$$

In addition, we know from the main text that μ_Y^e and μ_O^e are linked in our model by

$$\mu_Y^e = \frac{\gamma}{2}\ln\left[4 - \pi(\pi - 2)\right] + \ln\gamma + \ln\left[1 + \frac{1}{\gamma - \ln(1 - expsh)/(\mu_O^e - \mu_O^d)}\right] - \gamma\mu_O^e,$$

where σ has been substituted from above. There exists a similar function relationship in our model linking μ_Y^d and μ_O^d according to

$$\mu_Y^d = \frac{\gamma}{2}\ln\left[4 - \pi(\pi - 2)\right] + \ln\gamma + \ln\left[1 + \frac{1}{\gamma - \ln(1 - expsh)/(\mu_O^e - \mu_O^d)}\right] - \gamma\mu_O^d.$$

Finally, we can note from Section 5 that $\ln\hat{\theta}_1^* - \ln\hat{\theta}_0^* = \ln t$, with $\ln t = -\gamma(\mu_O^e - \mu_O^d)$ from above. Noting further that $\gamma(\mu_O^d - o) = u$ follows from Eq. (34b′), we can conclude that $u \geq \ln\hat{\theta}_1^* = \gamma(\mu_O^d - maxn)$, where $maxn$ is the observed maximum of composite $\ln[cv(\omega)b(\omega)/\tilde{\beta}(\omega)]$ among non-exporters. Putting together, this establishes

$$\ln\hat{\theta}_0^* = \gamma(\mu_O^e - maxn).$$

Using these functional relationships from our model and employing the average export share and the maximum $\ln[cv(\omega)b(\omega)/\tilde{\beta}(\omega)]$ from our data, reduces the number of parameters that have to be estimated to nine.

S3.1 Conditional likelihood functions

To estimate system (S.8), we have to derive the likelihood function. Starting point is the density for the stochastic parameter triple $\{u, v, e\}$, which, under our assumption of a

trivariate normal distribution, can be expressed as $f_{u,v,e} = g_{uv}\,\bar{g}_e / P(u \geq \ln\hat{\theta}_a^*)$, where g_{uv}, $\bar{g}_e$ are defined in Supplement S2.3, with

$$\bar{\mu}_e \equiv u\frac{\sigma_e}{\sigma_u}\frac{\rho_{ue}-\rho_{uv}\rho_{ev}}{1-\rho_{uv}^2} + v\frac{\sigma_e}{\sigma_v}\frac{\rho_{ev}-\rho_{ue}\rho_{uv}}{1-\rho_{uv}^2}, \quad \bar{\sigma}_e \equiv \sigma_e\sqrt{\frac{1-\rho_{ue}^2-\rho_{uv}^2-\rho_{ev}^2+2\rho_{ue}\rho_{uv}\rho_{ev}}{1-\rho_{uv}^2}},$$

and where $P(u \geq \ln\hat{\theta}_a^*) = \Phi\left(-\ln\hat{\theta}_0^*/\sigma_u\right)$ is the ex ante probability of drawing a sufficiently high $u \geq \ln\hat{\theta}_0^*$ to start production.

The (marginal) densities of tuple $\{u, v\}$ conditional on exporting ($I_i = 1$) and non-exporting ($I_i = 0$) can then be computed according to $f_{u,v}^e = \int_{-\mu_X}^{\infty} f_{u,v,e}de$ and $f_{u,v}^d = \int_{-\infty}^{-\mu_X} f_{u,v,e}de$. Solving the integrals gives

$$f_{u,v}^e = g_{uv}\frac{\Phi\left((\mu_X+\mu_e)/\bar{\sigma}_e\right)}{\Phi\left(-\ln\hat{\theta}_0^*/\sigma_u\right)}, \qquad f_{u,v}^d = g_{uv}\frac{\Phi\left(-(\mu_X+\mu_e)/\bar{\sigma}_e\right)}{\Phi\left(-\ln\hat{\theta}_0^*/\sigma_u\right)}. \tag{S.9}$$

It is a notable feature of our model that for observed realizations u and v the conditional likelihoods in eq. (S.9) do not permit separate identification of σ_e from μ_X. This is, why we set $\sigma_e = 1$, acknowledging that the model parameters are "estimable only up to a scale factor" (see Maddala 1986, p. 1635).

Lemma 1. *Denote the observed data with the vectors* $(\mathbf{y}, \mathbf{o}, \mathbf{I})$ *whose characteristic elements for plant* i *are* (y_i, o_i, I_i)*, denote the maximum observable* o_i *among non-exporters* (o_i^d)*, with* $maxn \equiv \max\{o_i^d\}$*, let* N *be the number of observations, and set* $\sigma_e = 1$*. We replace the truncation point* $\ln\hat{\theta}_0^*/\sigma_u$ *by* $\gamma(maxn - \mu_O^e)$*. Making use of structural relationships from our model to eliminate* σ*,* μ_Y^e*,* μ_Y^d*, the conditional likelihood function for system* (S.8) *is denoted*

$$\mathcal{L}\big(\cdot\,\big|\mathbf{y}, \mathbf{o}, \mathbf{I}\big) = \mathcal{L}\big(\gamma, \mu_O^e, \mu_O^d, \mu_X, \sigma_u, \sigma_v, \rho_{ue}, \rho_{uv}, \rho_{ve}\big|\mathbf{y}, \mathbf{o}, \mathbf{I}, maxn, expsh\big),$$

where $expsh$ *is the observed average ratio of export sales to total sales among exporting plants. Expressing* μ_Y^d *as function of parameters* γ, μ_O^d, μ_O^e *and observed* $expsh$ *and using auxiliary functions*

$$x_{i1} = -\frac{o_i - \mu_O^d - I_i(\mu_O^e - \mu_O^d)}{\sigma_u/\gamma}, \qquad x_{i2} = \frac{\left[y_i - \mu_Y^d + \gamma(o_i - \mu_O^d)\right]}{\sigma_v}$$

and

$$\hat{\mu}_i = x_{i1}\frac{\rho_{ue}-\rho_{uv}\rho_{ve}}{1-\rho_{uv}^2} + x_{i2}\frac{\rho_{ve}-\rho_{ue}\rho_{uv}}{1-\rho_{uv}^2}, \qquad \hat{x}_i = \sqrt{\frac{x_{i1}^2 - 2\rho_{uv}x_{i1}x_{i2} + x_{i2}^2}{1-\rho_{uv}^2}},$$

we can express the conditional likelihood function as

$$\mathcal{L}\big(\cdot\,\big|\mathbf{y},\mathbf{o},\mathbf{I},maxn,expsh\big) = \prod_{i=1}^{N}\left\{\frac{1}{\sqrt{2\pi}\sigma_u\sigma_v\sqrt{1-\rho_{uv}^2}}\frac{\phi\left(\hat{x}_i\right)}{\Phi\left(-\frac{maxn-\mu_O^e}{\sigma_u/\gamma}\right)}\Phi\left(\frac{\mu_X+\hat{\mu}_i}{\sqrt{\frac{1-\rho_{ue}^2-\rho_{uv}^2-\rho_{ev}^2+2\rho_{ue}\rho_{uv}\rho_{ve}}{1-\rho_{uv}^2}}}\right)\right\}^{I_i} \times\left\{\frac{1}{\sqrt{2\pi}\sigma_u\sigma_v\sqrt{1-\rho_{uv}^2}}\frac{\phi\left(\hat{x}_i\right)}{\Phi\left(-\frac{maxn-\mu_O^e}{\sigma_u/\gamma}\right)}\Phi\left(-\frac{\mu_X+\hat{\mu}_i}{\sqrt{\frac{1-\rho_{ue}^2-\rho_{uv}^2-\rho_{ev}^2+2\rho_{ue}\rho_{uv}\rho_{ve}}{1-\rho_{uv}^2}}}\right)\right\}^{1-I_i}.$$

Proof. The conditional likelihood function follows from eq. (S.9) after substituting $u = x_{i1}\sigma_u$ and $v = x_{i2}\sigma_v$, setting $\sigma_e = 1$, and accounting for the definitions of $\bar{\mu}_e$ and $\bar{\sigma}_e$. □

References

AUMANN, R. J. AND L. S. SHAPLEY (1974): *Values of non-atomic games*, Princeton, New Jersey: Princeton University Press.

AUTOR, D. H., D. DORN, AND G. H. HANSON (2013): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103, 2121–68.

BECKER, S. O. AND M.-A. MUENDLER (2015): "Trade and Tasks: An Exploration Over Three Decades in Germany," *Economic Policy*, 30, 589–641.

BRUGEMANN, B., P. GAUTIER, AND G. MENZIO (2015): "Intra Firm Bargaining and Shapley Values," *NBER Working Paper*, 21508.

DAUTH, W., S. FINDEISEN, AND J. SUEDEKUM (2014): "The Rise of the East and the Far East: German Labor Markets and Trade Integration," *Journal of the European Economic Association*, 12, 1643–75.

HELPMAN, E., O. ITSKHOKI, AND S. J. REDDING (2010): "Inequality and Unemployment in a Global Economy," *Econometrica*, 78, 1239–1283.

MADDALA, G. S. (1986): "Disequilibrium, Self-selection, and Switching Models," in *Handbook of Econometrics*, ed. by Z. Griliches and M. D. Intriligator, Amsterdam, Oxford and Tokyo: North-Holland, vol. 3 of *Handbooks in Economics*, chap. 28, 1633–88.

STOLE, L. A. AND J. ZWIEBEL (1996): "Intra-firm Bargaining under Non-binding Contracts," *Review of Economic Studies*, 63, 375–410.

STÜBER, H., W. DAUTH, AND J. EPPELSHEIMER (2023): "A Guide to Preparing the Sample of Integrated Labour Market Biographies (SIAB, version 7519 v1) for Scientific Analysis," *Journal for Labour Market Research*, 57, 1–11.