Monopsonistic Labor Markets and International Trade

Priyaranjan Jha, Antonio Rodriguez-Lopez

CES**ifo**

Monopsonistic Labor Markets and International Trade

Abstract

This paper introduces a framework to study the impact of trade liberalization on wage inequality and welfare in the presence of monopsonistic labor markets. The interaction of firm heterogeneity in productivity with idiosyncratic preferences of workers for working at different firms generates between-firm wage inequality for workers with identical skills. The degree of monopsony power is captured by the elasticity of firm-level labor supply, with a lower elasticity implying more wage-setting power by the firm. With more productive firms paying higher wages, monopsony power dampens the impact of firm heterogeneity on the allocation of market shares and allows lower productivity firms to survive. In a closed economy this increases inequality, but in an open economy high levels of monopsony power inhibit exporting, which may reduce inequality by compressing wages on the right side of the distribution. Nevertheless, inequality in the open economy is always higher than in autarky. Monopsony power reduces social welfare (for empirically plausible values of the labor supply elasticity) and the gains from trade.

Priyaranjan Jha
Department of Economics
University of California, Irvine
USA - Irvine, CA 92697
pranjan@uci.edu

Antonio Rodriguez-Lopez
Department of Economics
University of California, Irvine
USA - Irvine, CA 92697
jantonio@uci.edu

July 2019

1 Introduction

While traditional economics research has paid considerable attention to the non-competitive behavior in the product market—so much so that imperfect competition models have become pervasive—it is only recently that economists have started paying increased attention to the non-competitive practices in the labor market and their impact on wages. A recent paper by Naidu, Posner, and Weyl (2018) documents extensive usage of anti-competitive practices in the labor market (such as non-compete covenants and no-poaching agreements) that reduce the mobility of workers between jobs and increase the monopsony power of employers.[1] Whatever the source, the key marker of an employer's monopsony power in the labor market is an upward-sloping labor supply at the firm level, with the wage elasticity, θ, indicating the degree of the employer's wage-setting power. For example, a perfectly inelastic supply ($\theta = 0$) indicates absolute employer's wage-setting power, while a perfectly elastic supply ($\theta = \infty$) corresponds to the conventional case of a perfectly competitive labor market where the employer takes the wage as given. As discussed below, a growing empirical literature obtains estimates for θ that are quite far from the perfectly-competitive case, with values as low as 0.1 and rarely larger than 5.

Despite the resurgence of academic interest in firms' monopsony power over workers, the vast literature studying the labor market implications of trade has ignored this prevalent aspect of the market. This paper fills this gap by studying the implications of trade on the allocation of economic activity, wage inequality, and welfare when the labor market is monopsonistically competitive. We incorporate a static monopsonistic competition model of the labor market in a standard Melitz model with heterogeneous firms. Therefore, in our model there is monopolistic competition in the product market and monopsonistic competition in the labor market. Our modeling of the labor market is similar to Card, Cardoso, Heining, and Kline (2018), who depart from the competitive labor market by assuming that workers have idiosyncratic preferences for jobs across different employers; this could happen either because a worker likes the amenities offered by a particular firm or prefers the location of the firm. As a consequence, each firm ends up facing an upward-sloping labor supply curve with a constant wage elasticity, θ. By capturing monopsony power in a single parameter, our CELS (constant elasticity of labor supply) model provides straightforward intuition for the effects of firms' wage-setting power within a Melitz structure.

Facing an upward-sloping labor supply curve, if a firm wants to attract more workers, it will

---

[1]For example, Krueger and Posner (2018) find that 24.5 percent of U.S. workers said that they are, or have been, bound by non-compete agreements. Relatedly, Starr, Prescott, and Bishara (2019) find that 12.5 percent of U.S. workers without college education and earning less than $40,000 a year were bound by non-compete clauses in 2014.

have to offer higher wages. Hence, firm heterogeneity in productivity translates into heterogeneity in wages, with more productive firms not only hiring more workers but paying higher wages too. In our model workers have identical skills and are equally productive, and thus, we obtain wage inequality for workers with the same skill in a tractable manner.

In a closed-economy version of the model, we show that an increase in monopsony power lowers the cutoff productivity below which firms exit. Intuitively, monopsony power acts as a check on high-productivity firms because it makes hiring more expensive for them as they have to be further up the upward-sloping labor-supply schedule. Compared to a setup with a competitive labor market, monopsony power allows lower productivity firms to survive, dampening the impact of productivity on the allocation of market shares and resulting in a greater mass of firms in the industry. This also has implications for the "superstar firms" phenomenon studied by Autor, Dorn, Katz, Patterson, and Van Reenen (2017). They argue that an increase in product market competition allows the high productivity "superstar firms" to capture a larger share of the market. We show that an increase in monopsony power in the labor market dampens the superstar-firm effect. Alternatively, an increase in labor-market competition enhances the superstar-firm effect.

By allowing the survival of lower productivity firms, which pay lower wages, an increase in monopsony power raises inequality in the closed economy; *i.e.*, inequality increases due to the expansion of the lower tail of the wage distribution. Formally, using a Pareto distribution for firm productivity, we show that wages also have a Pareto distribution and an increase in monopsony power leads to an increase in the Gini coefficient of the distribution of wages.

Turning to the welfare implications in the closed economy, we first look at the impact of monopsony power on the average wage, which is a measure of welfare in a representative-agent economy and is also a measure of the economy's productive efficiency. We find that the relationship between monopsony power and the average wage is non-monotonic. On the one hand, an increase in monopsony power has the potential to increase aggregate output by increasing the number of varieties, but on the other hand it causes a misallocation of resources across heterogeneous firms. The misallocation arises from high-productivity firms paying higher wages and hence hiring less workers compared to the competitive-labor-market case. Despite the theoretical ambiguity, our numerical exercises reveal that an increase in monopsony power is associated with a decrease in the average wage for plausible values of θ. We also explore the welfare implications of monopsony power using an inequality-averse social welfare function. Since an increase in monopsony power increases inequality and likely reduces the average wage, any reasonable social welfare function will

show a reduction in welfare with an increase in monopsony power, which is what we obtain.[2]

In the open-economy case, and similar to Melitz (2003), only high-productivity firms export due to the existence of fixed and iceberg trading costs. A key difference from the standard Melitz model, however, is that the act of exporting has a feedback effect on wages because exporting firms need to hire more workers and therefore are further up the labor-supply schedule. Hence, when a firm is deciding whether to export, it not only has to consider trading costs and the profits in the export market, but also the impact of exporting on wages, which will affect its profits from domestic sales—an exporting firm pays a higher wage, reducing its profits from domestic sales.

Similar to the closed-economy case, the greater the monopsony power the lower the cutoff productivity level below which firms exit. More interestingly, the impact of monopsony power on the cutoff productivity level above which firms export is ambiguous. In fact, we obtain a U-shaped relationship between the labor-supply elasticity θ (inversely related to monopsony power) and the exporting cutoff productivity. That is, as monopsony power increases, the exporting cutoff decreases first but beyond a point it starts increasing. This has important implications for inequality. In contrast to the unambiguous increase in inequality in the closed economy, inequality does not necessarily increase with monopsony power in the open-economy case. For low trade-cost values, an increase in monopsony power increases inequality first and then reduces it. On the one hand, since an increase in monopsony power lowers the survival productivity cutoff, it drives inequality up by stretching the lower tail of the wage distribution. On the other hand, the changes in the exporting cutoff could potentially have an offsetting effect. As monopsony power increases from a low level, the exporting cutoff declines and this reinforces the inequality increasing effect of a decline in the survival cutoff. Beyond a point the exporting cutoff starts increasing, which works to reduce inequality: with exporters paying higher wages than non-trading firms, a reduction in the mass of exporters drives inequality down.

We also study the inequality implications of a reduction in the variable cost of trading for different degrees of monopsony power. For moderate and high degrees of monopsony power we find that trade liberalization increases inequality, but for low degrees of monopsony power trade liberalization first increases inequality but beyond a point further declines in trade costs reduce inequality. Again, the intuition comes from the impact of trade liberalization on the two cutoff productivity levels. An increase in the survival cutoff induced by trade liberalization reduces inequality by compressing the distribution of wages. However, trade liberalization also reduces the exporting cutoff. As argued in

---

[2]When we study the open-economy case, we find some situations where there could be a trade-off between equality and efficiency. Therefore, an inequality-averse social welfare function is useful in that setting because the equality-efficiency trade-off made by a social welfare function will depend on the degree of inequality aversion.

the preceding paragraph, this would tend to increase inequality because exporting firms pay higher wages than purely domestic firms. When the degree of monopsony power is high, the latter effect dominates resulting in an increase in inequality with trade liberalization. This is also the case with a low degree of monopsony power when the trade cost is high. However, as trade cost keeps falling with a low degree of monopsony power, the inequality reducing effect of an increase in the survival cutoff eventually starts dominating, resulting in lower inequality. Despite the non-monotonicity in the relationship between trade liberalization and inequality, we always find that inequality in the open economy is higher than inequality in the closed economy.

Regarding welfare in the open economy, we find that for all levels of monopsony power, trade liberalization always increases the average wage. In addition to the standard gains from trade in a Melitz model, trade liberalization alleviates the distortion arising from the misallocation of resources caused by the presence of monopsony power in the labor market. Additionally, even in the range of parameters where trade liberalization increases inequality, social welfare—measured by an inequality averse welfare function—is increasing for conventional degrees of inequality aversion. Finally, the gains from trade (based on either the average wage or an inequality averse social welfare function) decline with monopsony power.

The remainder of the paper is organized as follows. In section 2 we provide some stylized facts that motivate our theoretical model and are also consistent with the results of our theoretical model. Section 3 provides some discussion of the related theoretical literature. Section 4 provides the theoretical model for the closed economy. Section 5 extends it to the open-economy case. Section 6 provides some concluding remarks.

2 Empirical Background

This section describes four stylized facts that provide the empirical underpinnings of our model: *(i)* individual firms face an upward-sloping labor supply, *(ii)* more monopsony power is associated with lower wages, *(iii)* larger firms set higher wages, and *(iv)* exporters set higher wages than non-exporters.

Stylized Fact 1. *Labor markets are monopsonistic; that is, individual firms face an upward-sloping labor supply.*

A firm participates in a monopsonistic labor market—so that the firm has wage-setting power— if and only if it faces an upward-sloping labor supply. Hence, the key empirical object that indicates the degree of monopsony power in a labor market is the elasticity of firm-level labor supply. As in

the model below, suppose that the labor supply faced by some firm i is given by $L_i^S = \mathbb{B}w_i^\theta$, where w_i is the firm's wage, $\theta > 0$ is the labor-supply elasticity, and $\mathbb{B}$ is a labor-supply shifter. If $\theta \to 0$ we approach the case of a perfectly inelastic labor supply, while if $\theta \to \infty$ we have the case of a perfectly elastic labor supply. The latter case corresponds to a perfectly competitive labor market, with the firm taking the wage as given.

Stylized fact 1 refers to empirical evidence showing that θ is positive and finite. Although for several decades labor economists have discussed the existence of monopsonistic labor markets, the empirical estimation of θ imposes econometric challenges that the literature started to tackle just recently. In particular, even with the availability of firm-level data containing wages and employment, the lack of firm-level labor demand shifters that can serve as instruments for wages makes it hard to estimate the firm-level labor supply elasticity in a log-log regression of employment on wages. As a consequence, empirical studies estimating θ are mainly of two types: *(i) quasi-experimental studies*, which rely on quasi-experiments in special labor markets where certain conditions allow for the identification of the supply curve, and *(ii) job duration studies*, which use data on job spells to implement the estimation procedure proposed by Manning (2003)—developed from the job-search model of Burdett and Mortensen (1998)—to obtain the labor-supply elasticity from the estimation of recruiting and separation elasticities.

The typical labor markets analyzed in quasi-experimental studies are the markets for teachers and nurses and the estimates of labor elasticity in these studies range between 0.1 and 1.4.[3] Given the scarcity of quasi-experimental data from more general labor markets, a large set of studies uses instead Manning's job duration approach. After introducing this estimation technique, Manning (2003) uses U.S. and U.K. data and estimates labor-supply elasticities to individual firms in the range between 0.75 and 1.5. Of the studies that followed—some of which are discussed below as evidence for stylized fact 2—Webber (2015) provides the most comprehensive and disaggregated analysis using Manning's approach. Using the U.S. Census's Longitudinal Employer-Household Dynamics (LEHD) data for 47 states during 1985-2008, Webber (2015) estimates a different labor-supply elasticity for each individual firm in his sample, obtaining a mean of 1.08 and a median of 0.75 (as well, 90 percent of firms have elasticities below 1.73).[4]

---

[3] Falch (2010), for example, estimates a labor-supply elasticity of 1.4 to individual schools in the Norwegian teacher labor market by taking advantage of centralized wage setting combined with a policy-driven mechanism to identify schools with excess labor demand. Similarly, Staiger, Spetz, and Phibbs (2010) rely on an exogenous wage change in Veteran Affairs hospitals and find a labor-supply elasticity of 0.1 (for individual hospitals) in the labor market for nurses. Looking instead at online labor markets, Dube, Jacobs, Naidu, and Suri (forthcoming) build on previous experimental studies for the Amazon MTurk platform—which connects firms ("requesters") and workers for the realization of simple tasks—and calculate a labor supply elasticity for task requesters of about 0.1.

[4] By sector, Webber finds that manufacturing has the highest mean labor-supply elasticity at 1.82, and "admin-

More recently, Berger, Herkenhoff, and Mongey (2019) introduce a new structural method to estimate firm-level labor supply elasticities based on an oligopsony model that delivers measures of labor market power from firm-level market shares and parameters of labor substitutability. Using the U.S. Census's Longitudinal Business Database, they exploit changes in state-level corporate taxes—which serve as shocks to firm-level labor demand—and take advantage of different concentration levels across local labor markets to estimate firm-level labor supply elasticities ranging from 0.76 (for a firm that controls an entire local labor market) to 3.74 for the smallest firms.

Naidu, Posner, and Weyl (2018) also provide an extensive survey of studies that estimate θ and give a range from 1 to 5. To sum up, individual firms face upward-sloping labor supplies, with elasticities being as low as 0.1 and typically less than 5. These low values indicate substantial monopsony power.

Stylized Fact 2. *More monopsony power is related to lower wages.*

Theoretically, more monopsony power—*i.e.*, a lower firm-level labor supply elasticity (θ)—implies a lower wage. Indeed, the profit-maximizing wage for some firm i facing an upward-sloping labor supply with elasticity θ is always given by

$$w_i = \left(\frac{\theta}{1+\theta}\right) MRPL_i = \left(1 - \frac{1}{1+\theta}\right) MRPL_i,$$

where $MRPL_i$ is the firm's marginal revenue product of labor. The term $\frac{1}{1+\theta} \in (0,1)$ is the firm's proportional markdown on the MRPL. If θ is close to zero, the markdown approaches 1 and the firm has almost full monopsony power (it sets a wage close to zero), while if $\theta \to \infty$ the firm has no wage-setting power (its markdown approaches zero) and thus $w_i \to MRPL_i$. Considering a range for θ between 0.1 and 5 implies markdowns ranging between 0.17 (if $\theta = 5$) and 0.91 (if $\theta = 0.1$). Empirically, the emergence of studies that estimate θ has naturally led to papers quantifying the effects of monopsony power on wages.

By being able to estimate a supply elasticity for each individual firm, Webber (2015) is then able to exploit this variation to obtain an estimate for the impact of elasticity differences on log earnings. After controlling for person and firm fixed effects, as well as for demographic and employer controls, he finds that a unit increase in the labor supply elasticity is associated with a 15 percent increase in earnings.

---

istrative support" the lowest at 0.72. Relatedly, Hirsch, Jahn, and Schnabel (2018) use German linked employer-employee quarterly data from 1985 to 2010 to study the evolution through time of the firm-level labor supply elasticity. They find that θ ranges between 1.86 and 2.81, taking lower values—so that firms have more monopsonistic power—during downturns.

Another set of studies providing direct evidence that a lower θ is related to lower wages concerns the estimation of firm-level labor supply elasticities for males and females, which then are used to analyze the gender wage gap. Barth and Dale-Olsen (2009) refer to wage differences between males and females arising from different labor-supply elasticities as *monopsonistic discrimination*, as they are the result of firms' profit-maximization decisions (rather than the result of employers' biases). Using matched employer-employee data from Norway, their preferred job-duration specifications yield elasticities of 0.84 and 1.17 for low- and high-educated females, respectively, and of 1.098 and 1.71 for low- and high-educated males. Moreover, they find that these gender differences in elasticities are related to lower wages for females by about 14.5 percent for the low-educated group, and by about 12.8 percent for the high-educated group—these numbers explain about 90 and 70 percent of the gender wage gap for low- and high-educated workers, respectively.

Similarly, Hirsch, Schank, and Schnabel (2010) use matched employer-employee data from Germany and estimate firm-level labor supply elasticities ranging between 1.87 and 2.59 for females, and between 2.49 to 3.66 for males. These differences imply between 4.6 percent to 17.4 percent higher wages for males, which can explain at least a third of the gender wage gap.[5] Extending his previous work, Webber (2016) uses U.S. LEHD data and finds average firm-level labor supply elasticities of 0.94 for females and 1.09 for males, which are associated with a 3.3 percent higher average wage for males.

Stylized Fact 3. *The employer-size wage effect: on average, larger firms pay higher wages. This effect remains after controlling for worker characteristics.*

A monopsonistic labor market implies that larger firms pay higher wages: for two firms facing a similar upward-sloping labor supply, the firm with the largest labor demand will pay a higher equilibrium wage. Thus, the so-called "employer-size wage effect"—which has strong empirical support—is a natural consequence of the heterogeneous-firm model in this paper. Given our model's homogeneous-labor assumption, the robustness of stylized fact 3 to worker-level controls is important, as it highlights that the employer-size wage effect cannot be explained solely by larger firms hiring more-skilled workers.

The classic reference documenting the prevalence of the employer-size wage effect is the paper by Brown and Medoff (1989), who show using U.S. data that a large positive relationship between employer size and wages appears within groups of similar workers and after including industry-

---

[5]Using data from a U.S. grocery retailer and applying a simpler version of Manning's aproach, Ransom and Oaxaca (2010) obtain female elasticities between 1.5 and 2.5, and male elasticities between 2.4 and 3, with predicted wage gaps ranging between 4 and 20 percent.

level controls. Oi and Idson (1999) describe similar results in their extensive survey of the size-wage premium for several countries. Troske (1999) uses detailed matched employer-employee data of U.S. manufacturing establishments and workers, and finds—after controlling for worker characteristics— a 13 percent higher wage for establishments one standard deviation above the mean when compared to those one standard deviation below.[6] The gap declines to 11 percent after controlling for plant-level variables of workforce skill, indicating that matching of high-skilled workers with larger plants only explains about 20 percent of the employer-size wage premium.

Gibson and Stillman (2009) use worker-level data from a literacy survey for nine OECD countries and after controlling for demographic variables, they obtain statistically significant employer-size wage premiums in seven of them. They further add controls for worker skills (including education and literacy levels), but find that they do nothing or very little to reduce these premiums.[7]

Stylized Fact 4. *The exporter wage premium: on average, exporters pay higher wages than non-exporters. The premium persists even after controlling for worker characteristics.*

In the Melitz (2003) model, firm-level labor demand has two characteristics: first, it is increasing in productivity, and second, given the existence of a cutoff productivity level that separates exporting and non-exporting firms, there is a discrete jump in labor demand at this cutoff as exporters require more workers to produce for the foreign market. The same happens if firms face an upward sloping labor supply (as in our model), but now there are also implications for wages, which show a positive relationship with productivity, and a positive discrete jump at the exporting cutoff level.

Therefore, in our model the exporter wage premium is an enhanced employer-size wage effect: exporters pay higher wages than non-exporters because they are more productive and thus larger, but on top of that, exporting requires a discrete jump in employment that further exacerbates the wage gap. Given the solid empirical evidence showing that exporters are on average larger than non-exporters (Bernard, Jensen, Redding, and Schott, 2007), stylized fact 4 can be seen as an extension of stylized fact 3, but fortunately, it has been independently verified by a large amount of studies.

Bernard and Jensen (1995) present the first evidence showing that U.S. manufacturing exporters

---

[6] Bayard and Troske (1999) expand the sample to include non-manufacturing establishments, finding premiums of 10 and 11 percent in retail trade and services, respectively.

[7] Green, Machin, and Manning (1996) document the employer size-wage premium in the U.K. and importantly, they provide evidence that the effect is larger for women and in the non-union sector (see also the references cited therein for similar evidence in other countries). They interpret the latter results as supporting evidence for a monopsonistic-power channel, as wages of groups with a lower labor supply elasticity at the firm level (such as women and non-unionized workers) should be more responsive to differences in firm-level labor demand.

pay higher wages than non-exporters for both production and non-production workers. Expanding these results, Bernard and Jensen (1999) find that firms that will become exporters are ex-ante more productive, larger, and have higher wages than non-exporters, and once they start exporting their short-run growth rates of employment and wages—after controlling for plant characteristics— are also larger. These findings align well with the mechanisms described in our model. Schank, Schnabel, and Wagner (2007) surveys 22 studies for 21 different countries that obtain similar findings, and more recently, Brambilla, Depetris Chauvin, and Porto (2017) provide evidence of the exporter wage premium using data from 61 non-advanced countries.

Baumgarten (2013) uses linked employer-employee data during 1996-2007 to study the exporter wage gap within and between skill groups for male German workers in manufacturing. Closer to our model, which abstracts from skill differences and within-firm wage heterogeneity, he finds that two thirds of the wage inequality increase during the period occurred within skill groups, 80 percent occurred within industry, and two thirds occurred between establishments. Moreover, conditional on skill levels, he finds that the increase in the exporter wage premium contributed to the rise in inequality, but the effect becomes smaller and statistically insignificant when controlling for establishment size.[8] The last finding is consistent with our model, as in our theory the higher wages for exporters are a consequence of their size.

Similarly, Frías, Kaplan, Verhoogen, and Alfaro-Serrano (2018) use linked employer-employee data from Mexico to study the exporter wage premium while carefully accounting for changes in the skill composition within firms. Using the large Mexican peso devaluation at the end of 1994 as the shock affecting exporting incentives, they find an increase in the exporting wage premium and that this is driven by different wage policies across exporting and non-exporting plants, rather than by changes in the workforce composition.[9]

3 Theoretical Background

Robinson (1933) introduced the first formal discussion of monopsony; however, for several decades it was confined to the fringes of labor economics and to the simple textbook partial-equilibrium case discussing the possibility of a minimum wage raising employment. Since the 1990s, Alan Manning has played a key role in reviving the theoretical and empirical literature on monopsony.

---

[8]Using the same German linked employer-employee data, Schank, Schnabel, and Wagner (2007) also find non-significant effects of exporting status on wages when controlling for establishment size, but find positive significant effects when using instead the export share in total sales.

[9]Macis and Schivardi (2016) perform a similar exercise with linked employer-employee data from Italy and using the 1992 lira depreciation as an identification tool. Keeping worker characteristics constant, they find that exporting causes firms to pay higher wages.

Interestingly, in his book *Monopsony in Motion* (Manning, 2003), he gives credit to Burdett and Mortensen (1998) for reviving his interest in the subject after he encountered an early version of this well known paper. Following Burdett and Mortensen (1998), who introduce a wage-posting model where monopsony power arises due to search frictions, Manning (2003) provides a series of dynamic models to explore the welfare consequences of firms' wage-setting power and to explain several empirical labor-market observations (see Manning, 2011 for a more recent survey of this literature). Unlike these dynamic models of monopsony power, we use a static framework similar to Card, Cardoso, Heining, and Kline (2018) where workers have idiosyncratic preferences for different employers. The static framework keeps the model tractable and allows us to study the interaction between the degree of monopsony power and trade barriers in the open economy.

The closest papers to ours are those presenting Melitz-type heterogeneous-firm models where different wages across firms are the result of rent-sharing mechanisms between firms and workers. These include the fair-wage models of Amiti and Davis (2011) and Egger and Kreickemeier (2009, 2012), the screening model of Helpman, Itskhoki, and Redding (2010), and the search models of Coşar, Guner, and Tybout (2016), Felbermayr, Impullitti, and Prat (2018), and Fajgelbaum (2016). In these models more productive firms are larger and have a larger surplus to share with their workers, and hence they also feature an employer-size wage premium and an exporter wage premium. The key difference with our framework is that we model instead a labor market with monopsonistic competition. By capturing firms' wage-setting power in a single parameter—the firm-level labor supply elasticity, θ—our model delivers clear insights for the wage-inequality consequences of interactions between monopsony power, monopolistic power (as measured by the elasticity of substitution in the product market), productivity dispersion, and trade-barriers parameters.

For example, one of the most important results in the model of Helpman, Itskhoki, and Redding (2010) is that there is an inverted-U relationship between trade barriers and wage inequality: starting from a situation with high trade barriers, trade liberalization initially increases inequality but further reductions in trade costs eventually start to reduce inequality. We find that this result crucially depends on the level of monopsony power, and in particular, it only occurs when monopsony power is very low (*i.e.*, for high levels of the firm-level labor supply elasticity, θ). When the degree of monopsony power is moderate or high, trade liberalization monotonically increases inequality.

4 The Closed-Economy Model

To develop intuition on how monopsony power affects inequality and welfare, we start with a closed-economy model. We then move to an open-economy model with two identical countries.

4.1 Preferences and Demand

The country is populated by $\mathbb{L}$ workers, who are employed by differentiated-good firms. Each worker derives utility from the consumption of differentiated goods and from an idiosyncratic term that indicates the worker's preference for that firm.

A worker's utility from consumption is given by $u(Z) = \ln Z$, where Z is a CES aggregator of differentiated goods,

$$Z = \left(\int_{\omega \in \Omega} q^c(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}},$$

which we call "final good". In Z, $q^c(\omega)$ is the worker's consumption of variety ω, Ω is the set of varieties available for purchase, and $\sigma > 1$ denotes the elasticity of substitution between differentiated-good varieties. A worker's demand for variety ω is then given by $q^c(\omega) = Z \left(\frac{p(\omega)}{P} \right)^{-\sigma}$, where $p(\omega)$ is the price of variety ω, and $P = \left(\int_{\omega \in \Omega} p(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}}$ is the price of the final good, Z. With each firm producing a single variety, ω also identifies the firm.

The final good is the numéraire, and hence $P = 1$. With this normalization, the amount of final good consumed by a worker is simply equal to her wage. Therefore, and following Card, Cardoso, Heining, and Kline (2018), the utility of worker i who works at firm ω receiving wage $w(\omega)$ is

$$v_i(\omega) = \ln w(\omega) + \varepsilon_i(\omega), \tag{1}$$

where $\varepsilon_i(\omega)$ is an idiosyncratic draw from a Type I Extreme Value distribution with mean zero and variance $\pi^2 \eta^2 / 6$, with $\eta > 0$ denoting the distribution's dispersion parameter. As described by Card, Cardoso, Heining, and Kline (2018), the idiosyncratic term $\varepsilon_i(\omega)$ denotes the quality of the worker-firm match (from the worker's perspective), capturing non-pecuniary factors affecting the worker's preference for firm ω, such as distance and personal motives.[10]

4.2 The CELS Monopsonistic Labor Market

Given the utility function in (1), with workers being free to choose where to work, and under the extreme value distribution of the idiosyncratic draws, $\{\varepsilon_i(\omega)\}$, it follows from Ben-Akiva, Litinas,

---

[10]Card, Cardoso, Heining, and Kline (2018) assume a more general utility function, $v_{iS}(\omega) = \psi \ln[w_S(\omega) - b_S] + a_S(\omega) + \varepsilon_{iS}(\omega)$, where where ψ is a parameter that denotes the importance of consumption in overall utility, S denotes skill group (either low or high skilled), b_S is the reservation wage for skill group S, and $a_S(\omega)$ allows for different amenities across firms and skill groups. In this paper we abstract from these other features to focus on the first order effects of monopsony power on welfare and inequality.

and Tsunokawa (1985) (who generalize the discrete-choice logit model of McFadden (1973) to the continuous-choice case) that each worker's logit probability of working at firm ω is

$$\mathbb{P}(\omega) = \frac{\exp\left[\frac{1}{\eta} \ln w(\omega)\right]}{\int_{\nu \in \Omega} \exp\left[\frac{1}{\eta} \ln w(\nu)\right] d\nu}.$$

Thus, the labor supply function faced by firm ω is simply given by $L^S(\omega) = \mathbb{P}(\omega)\mathbb{L}$, which simplifies to

$$L^S(\omega) = \mathbb{B}w(\omega)^{\frac{1}{\eta}} \equiv \mathbb{B}w(\omega)^\theta, \tag{2}$$

where $\theta \equiv \frac{1}{\eta}$ is the wage elasticity of labor supply for every firm, and $\mathbb{B} \equiv \left[\int_{\nu \in \Omega} w(\nu)^\theta d\nu\right]^{-1} \mathbb{L}$ is a labor-supply shifter taken as given by each firm in the market. With $\theta > 0$ being constant across firms, so that all firms face a similar upward-sloping labor supply, this model features a CELS (constant elasticity of labor supply) monopsonistic labor market.

As discussed above, the elasticity θ captures the degree of monopsony power, with a lower value indicating higher wage-setting power. From its definition note that θ is inversely related to η (the dispersion parameter for the distribution of a worker's job preferences over different firms). Intuitively, a large η implies that small wage differences will not induce many workers to change jobs, thereby making the labor-supply elasticity lower and thus granting more monopsony power to firms. Alternatively, a large η can indicate lower labor mobility, which results in a less elastic labor supply. In contrast, in the extreme case of $\eta \to 0$, workers are perfectly mobile—there is no random variation in the utility function in (1)—and the labor market approaches the perfectly competitive case with $\theta \to \infty$.[11]

4.3 Production

Firms are heterogeneous in productivity as in Melitz (2003). After paying a sunk entry cost of f_E units of the final good, a firm draws its productivity φ from a distribution with probability density function $g(\varphi)$ and cumulative function $G(\varphi)$. A firm produces for the market if and only if it can cover a fixed cost of operation, f, also in terms of the final good. After meeting the fixed cost, the production function for a firm with productivity φ is $y(\varphi) = \varphi L$ where L is the amount of labor hired by the firm. In contrast to the Melitz model, however, firms face an upward sloping labor supply curve and thus, producing firms with different productivities pay different wages.

Given that each firm can now be identified by its productivity, we can replace ω with φ, and

---

[11]In our analysis below we perform comparative statics with respect to θ to capture varying degrees of monopsony power, but the underlying variation arises from η.

write the profit-maximization problem for a firm with productivity φ as

$$\max_{w(\varphi)} \left\{ p(\varphi)q(\varphi) - w(\varphi)L(\varphi) \right\},$$

where $q(\varphi) \equiv \mathbb{A}p(\varphi)^{-\sigma}$ is the market demand faced by a firm with productivity φ, with $\mathbb{A}$ denoting the total expenditure on differentiated goods in the economy (which is identical to the aggregate revenue of all firms). In equilibrium the quantity demanded for a firm's variety equals its production $(q(\varphi) = y(\varphi))$, and therefore we can write the inverse demand function as $p(\varphi) = \left[\frac{\varphi L(\varphi)}{\mathbb{A}} \right]^{-\frac{1}{\sigma}}$. Moreover, in equilibrium the labor requirements of a firm with productivity φ equal the labor supply, and thus, we use (2) to substitute for $L(\varphi)$ and rewrite the profit-maximization problem as

$$\max_{w(\varphi)} \left\{ \mathbb{A}^{\frac{1}{\sigma}} \left[\mathbb{B}\varphi w(\varphi)^{\theta} \right]^{\frac{\sigma-1}{\sigma}} - \mathbb{B}w(\varphi)^{1+\theta} \right\}.$$

The profit-maximizing wage of a firm with productivity φ is then given by

$$w(\varphi) = \left(1 - \frac{1}{1+\theta} \right) \underbrace{\left[\left(\frac{1+\theta}{\theta} \right)^{\theta} \left(\frac{\sigma-1}{\sigma} \right)^{\sigma} \frac{\mathbb{A}}{\mathbb{B}} \right]^{\frac{1}{\sigma+\theta}} \varphi^{\frac{\sigma-1}{\sigma+\theta}}}_{MRPL(\varphi)}, \tag{3}$$

where $MRPL(\varphi)$ is the firm's marginal revenue product of labor. Note that each firm sets a wage below its marginal revenue product of labor, with the CELS proportional markdown given by $\frac{1}{1+\theta} \in (0,1)$. More monopsony power (a lower θ) implies a larger markdown, approaching 1 (so that the wage approaches zero) as $\theta \to 0$. Note also that $w'(\varphi) > 0$, and thus, more productive firms pay higher wages. Substituting (3) into (2), we obtain that the amount of labor hired by a firm with productivity φ is

$$L(\varphi) = \left\{ \left[\frac{(\sigma-1)\theta}{\sigma(1+\theta)} \right]^{\sigma\theta} \mathbb{A}^{\theta}\mathbb{B}^{\sigma} \right\}^{\frac{1}{\sigma+\theta}} \varphi^{\frac{(\sigma-1)\theta}{\sigma+\theta}}, \tag{4}$$

which is also increasing in productivity. Therefore, more productive firms not only pay higher wages, but also have more employment.

We can also verify that

$$p(\varphi) = \left\{ \left[\frac{\sigma(1+\theta)}{(\sigma-1)\theta} \right]^{\theta} \frac{\mathbb{A}}{\mathbb{B}} \right\}^{\frac{1}{\sigma+\theta}} \frac{1}{\varphi^{\frac{1+\theta}{\sigma+\theta}}} = \left(\frac{\sigma}{\sigma-1} \right) \left[\left(\frac{1+\theta}{\theta} \right) \frac{w(\varphi)}{\varphi} \right]. \tag{5}$$

The first equality shows that even though more productive firms pay higher wages, they still charge a lower price $(p'(\varphi) < 0)$. The second equality presents the standard CES pricing equation of a fixed markup over marginal cost, with the difference that the marginal cost is now $\left(\frac{1+\theta}{\theta} \right) \frac{w(\varphi)}{\varphi}$: the factor $\frac{1+\theta}{\theta} > 1$ captures the monopsonistic-power feature that producing an extra unit of output

13

increases the wage not only for the extra units of labor needed, but also for all the previously hired labor.

Moreover, we can rearrange (5) to see the double markdown in wages resulting from the monopoly power in the product market and monopsony power in the factor market. Rewriting (5) as

$$w(\varphi) = \left(\frac{\theta}{1+\theta} \right) \left(\frac{\sigma - 1}{\sigma} \right) p(\varphi)\varphi,$$

notice that since φ is the marginal product of labor, $p(\varphi)\varphi$ is the value of the marginal product, which would equal the firm's wage if both the product and the labor markets were competitive. Hence, $\left(\frac{\sigma-1}{\sigma} \right)$ captures the markdown due to the imperfect product market and $\frac{\theta}{1+\theta}$ captures the markdown due to the imperfect labor market.

Lastly, the gross profit (before deducting fixed costs) of a producing firm with productivity φ, $\pi(\varphi) \equiv p(\varphi)q(\varphi) - w(\varphi)L(\varphi)$, can be written as

$$\pi(\varphi) = \left\{ \frac{[(\sigma - 1)\theta]^{(\sigma-1)\theta} \, \mathbb{A}^{1+\theta}\mathbb{B}^{\sigma-1}}{[\sigma(1+\theta)]^{\sigma(1+\theta)}} \right\}^{\frac{1}{\sigma+\theta}} (\sigma + \theta) \, \varphi^{\beta}, \tag{6}$$

where $\beta \equiv \frac{(\sigma-1)(1+\theta)}{\sigma+\theta}$ is the elasticity of $\pi(\varphi)$ with respect to φ. Note that β rises with θ (it approaches $\frac{\sigma-1}{\sigma}$ if $\theta \to 0$, and approaches $\sigma - 1$ when $\theta \to \infty$); therefore, more monopsony power (a lower θ) dampens the effect that firm-productivity differences have in the allocation of markets shares across all firms in the economy. This happens because more productive firms must pay higher wages, which reduces their cost advantage over least-productive firms.

The existence of monopsony power also has implications for the "superstar firms" phenomenon. Autor, Dorn, Katz, Patterson, and Van Reenen (2017) argue that an increase in product market competition—an increase in the elasticity of product demand (σ) in our setting—will lead the high-productivity superstar firms to capture a larger share of the market. To see this clearly, let us compare the output of a high-productivity firm, φ_h, with a low-productivity firm, φ_l. Since the output demand for a firm is $q(\varphi) \equiv \mathbb{A}p(\varphi)^{-\sigma}$, using the expression for $p(\varphi)$ from above we obtain

$$\frac{q(\varphi_h)}{q(\varphi_l)} = \left(\frac{\varphi_h}{\varphi_l} \right)^{\frac{\sigma(1+\theta)}{\sigma+\theta}}.$$

For $\theta \to \infty$ (the competitive-labor-market case) the right-hand side above becomes $\left(\frac{\varphi_h}{\varphi_l} \right)^{\sigma}$, while for $\theta \to 0$ (the absolute-monopsony-power case) the right-hand side approaches $\frac{\varphi_h}{\varphi_l}$. Therefore, any increase in σ increases the relative market share of the high-productivity firm, but the effect is smaller the lower the value of θ and becomes negligible when $\theta \to 0$. Hence, monopsony power reduces the market-share advantage of the high-productivity firm. In other words, an increase

in monopsony power in the labor market dampens the superstar-firm effect. Alternatively, just as an increase in product-market competition enhances the superstar-firm effect, an increase in labor-market competition (an increase in θ) does the same.

4.4 Free-Entry Condition and Aggregate Productivity

As a Melitz-type model, there exists a cutoff productivity level, $\hat{\varphi}$, that determines the tradability of a firm's variety in the market: a firm with productivity φ sells in the market if and only if $\varphi \geq \hat{\varphi}$. The firm at the cutoff has a zero net profit, $\pi(\hat{\varphi}) = f$. Using this zero-cutoff-profit condition, we rewrite the gross profit function for a firm with productivity φ as

$$\pi(\varphi) = \left(\frac{\varphi}{\hat{\varphi}}\right)^{\beta} f. \tag{7}$$

Firms enter up to the point that the expected value of entry is equal to the entry cost, f_E. With firms knowing their productivity only after paying the entry cost, and assuming a death shock that wipes out a fraction δ of firms every period, the free-entry condition is given by $\frac{1}{\delta} \int_{\hat{\varphi}}^{\infty} [\pi(\varphi) - f] g(\varphi) d\varphi = f_E$, which using (7) becomes

$$\frac{1}{\delta} \int_{\hat{\varphi}}^{\infty} \left[\left(\frac{\varphi}{\hat{\varphi}}\right)^{\beta} - 1 \right] f g(\varphi) d\varphi = f_E. \tag{8}$$

This condition pins down the solution for $\hat{\varphi}$. We can then write the following lemma.

Lemma 1. *From* (8), *the elasticity of $\hat{\varphi}$ with respect to θ, $\zeta_{\hat{\varphi},\theta}$, is*

$$\zeta_{\hat{\varphi},\theta} = \left[\frac{\beta\theta}{(1+\theta)^2} \right] \left[\int_{\hat{\varphi}}^{\infty} \varphi^{\beta} g(\varphi) d\varphi \right]^{-1} \left\{ \int_{\hat{\varphi}}^{\infty} \left[\ln\left(\frac{\varphi}{\hat{\varphi}}\right) \right] \varphi^{\beta} g(\varphi) d\varphi \right\} > 0.$$

Therefore, more monopsony power in the labor market (i.e., a lower θ) implies a smaller $\hat{\varphi}$, so that it is easier for low-productivity firms to survive.

To gain tractability, we follow Chaney (2008) and Melitz and Ottaviano (2008) and assume a Pareto distribution for firm-level productivity. In particular, the cumulative distribution function is $G(\varphi) = 1 - \left(\frac{1}{\varphi}\right)^k$, and the probability density function is $g(\varphi) = \frac{k}{\varphi^{k+1}}$, where $\varphi \in [1, \infty)$ and $k > \sigma - 1 > \beta$. The parameter k indicates the degree of productivity dispersion, with a higher k implying less heterogeneity (if $k \to \infty$, firms are homogeneous with productivity equal to 1). Under a Pareto distribution, we solve for $\hat{\varphi}$ from (8) as

$$\hat{\varphi} = \left[\left(\frac{f}{\delta f_E}\right) \frac{\beta}{k-\beta} \right]^{\frac{1}{k}}. \tag{9}$$

15

The average productivity in the economy is then given by

$$\bar{\varphi} = \left[\int_{\hat{\varphi}}^{\infty} \varphi^{\beta} g(\varphi | \varphi \geq \hat{\varphi}) d\varphi \right]^{\frac{1}{\beta}} = \left(\frac{k}{k - \beta} \right)^{\frac{1}{\beta}} \hat{\varphi}, \tag{10}$$

which is also strictly increasing in θ.[12] Thus, monopsony power reduces the economy's aggregate productivity.

4.5 Autarky Equilibrium

Besides the cutoff productivity level, $\hat{\varphi}$, which is uniquely determined by (8), the model has three more endogenous variables: the mass of producing firms, M, the labor-supply shifter, $\mathbb{B}$, and the total expenditure on the final good, $\mathbb{A}$.

The aggregate price is defined as $P = \left[M \int_{\hat{\varphi}}^{\infty} p(\varphi)^{1-\sigma} g(\varphi | \varphi \geq \hat{\varphi}) d\varphi \right]^{\frac{1}{1-\sigma}}$, which conveniently simplifies to $P = M^{\frac{1}{1-\sigma}} p(\bar{\varphi})$. Given our assumption that $P = 1$ (the final good is the numéraire), it follows that $M = p(\bar{\varphi})^{\sigma-1}$, which using (5) can be rewritten as

$$M = \left\{ \left[\frac{\sigma(1+\theta)}{(\sigma-1)\theta} \right]^{\theta} \frac{\mathbb{A}}{\mathbb{B}} \right\}^{\frac{\sigma-1}{\sigma+\theta}} \frac{1}{\bar{\varphi}^{\beta}}. \tag{11}$$

Let M_E denote the mass of entrants every period. In equilibrium, the mass of producing firms that die due to the death shock, δM, is exactly replaced by the mass of successful entrants, $[1-G(\hat{\varphi})]M_E$, and thus, $M_E = \frac{\delta M}{1-G(\hat{\varphi})}$.

The aggregate expenditure on final goods has three components: workers' expenditure on final-good consumption, producing firms' final-good requirements to cover the fixed costs of operation, and entrants' final-good requirements to cover the entry cost. Therefore, the economy's total expenditure on the final good is given by

$$\mathbb{A} = M \int_{\hat{\varphi}}^{\infty} w(\varphi) L(\varphi) g(\varphi | \varphi \geq \hat{\varphi}) d\varphi + Mf + \left[\frac{\delta M}{1 - G(\hat{\varphi})} \right] f_E, \tag{12}$$

where the first term on the right-hand side is the wage bill of producing firms (workers spend all their wages on the final good) with $w(\varphi)$ and $L(\varphi)$ given by (3) and (4), and the last two terms correspond to total fixed and entry costs. Lastly, from section 4.2 we know that the labor-supply shifter is defined as

$$\mathbb{B} = \frac{\mathbb{L}}{M \int_{\hat{\varphi}}^{\infty} w(\varphi)^{\theta} g(\varphi | \varphi \geq \hat{\varphi}) d\varphi}. \tag{13}$$

We can now define the equilibrium in the closed-economy model.

---

[12]We obtain $\zeta_{\bar{\varphi},\theta} = \frac{\theta}{(1+\theta)^2} \left[\frac{2\beta}{k-\beta} - \ln\left(\frac{k}{k-\beta} \right) \right]$, which is always positive because natural log inequalities imply that $\frac{\beta}{k} < \ln\left(\frac{k}{k-\beta} \right) < \frac{\beta}{k-\beta}$.

Definition. *An autarky equilibrium solves for $\hat{\varphi}$ from* (8), *and then solves for M, $\mathbb{A}$, and $\mathbb{B}$ from* (11), (12), *and* (13), *with $w(\varphi)$ and $L(\varphi)$ given by* (3) *and* (4).

Under our Pareto distribution for productivity, from (11), (12), and (13) we solve for the mass of firms as

$$M = \frac{1}{k}\left[\frac{(\sigma-1)\mathbb{L}}{\sigma f (k-\beta)^{\frac{1}{\sigma-1}}}\left(\frac{k}{\beta} - \frac{\theta}{1+\theta}\right)\hat{\varphi}\right]^{\frac{\sigma-1}{\sigma-2}},\tag{14}$$

where $\hat{\varphi}$ is given by (9). The solution in (14) requires that $\sigma > 2$, which we assume to be the case, so that larger markets (a higher $\mathbb{L}$) and lower fixed costs (a lower f) are associated with more firms in the economy. The elasticity of M with respect to θ is then

$$\zeta_{M,\theta} = -\left[\frac{(\sigma-1)\theta\beta}{(\sigma-2)(1+\theta)^2(k-\beta)}\right]\left[\frac{(k-1)(k-\sigma+1)\theta + k\sigma(k-1) - \beta\sigma}{k\sigma + \theta(k-\sigma+1)}\right].$$

A sufficient condition for $\zeta_{M,\theta} < 0$ is that $k \geq 2$, which ensures that $k\sigma(k-1) > \beta\sigma$ for every θ. Hereafter we assume that $k \geq 2$.[13] Hence, in autarky, more monopsony power in the labor market (a lower θ) implies a larger mass of firms: with more productive firms paying higher wages, monopsony power dampens the impact of firm heterogeneity on the allocation of market shares, which translates into higher survival rates and a larger mass of firms.

Finally, the solutions for the aggregate expenditure on final goods, $\mathbb{A}$, and the labor-supply shifter, $\mathbb{B}$, are

$$\mathbb{A} = \frac{Mf\sigma}{\sigma-1}\left(\frac{k\beta}{k-\beta}\right) \quad \text{and} \quad \mathbb{B} = \left[\frac{\mathbb{L}}{Mk}\left(\frac{k}{\beta} - \frac{\theta}{1+\theta}\right)\right]^{1+\theta}\left(\frac{1+\theta}{f\theta}\right)^{\theta}\beta,$$

where M is given by (14). Moreover, the mass of entrants can be written as $M_E = \frac{Mf}{f_E}\left(\frac{\beta}{k-\beta}\right)$, which is proportional to $\mathbb{A}$; indeed, the fraction of aggregate expenditure accounted for by firms covering their entry costs, $\frac{M_E f_E}{\mathbb{A}}$, is constant and equal to $\frac{\sigma-1}{k\sigma}$. Hence, $\mathbb{A}$ and M_E have identical responses to changes in θ, with elasticities given by

$$\zeta_{\mathbb{A},\theta} \equiv \zeta_{M_E,\theta} = \left[\frac{\theta\beta}{(\sigma-2)(1+\theta)^2(k-\beta)}\right]\left[-(k-\sigma+1) + \frac{\sigma(\sigma-1)\beta}{k\sigma + \theta(k-\sigma+1)}\right].$$

In contrast to $\zeta_{M,\theta} < 0$, $\zeta_{\mathbb{A},\theta}$ may be positive for lower levels of θ (so that more monopsony power is associated with less entry and aggregate expenditure) if $k - \sigma + 1$ is sufficiently small, which is a consequence of reductions in the total wage bill when θ approaches zero.

---

[13]Calibrations of heterogeneous-firm models with CES preferences and a Pareto distribution for productivity typically assume values for σ and k that are well above 2 (see, for example, Ghironi and Melitz, 2005 and Chaney, 2008), and therefore, we believe that our assumptions that $\sigma > 2$ and $k \geq 2$ are innocuous.

4.6 Monopsony Power and Inequality

One of the main objectives of this paper is to assess the effects of monopsony power in the labor market on wage inequality. Each firm pays a different wage, and given a precise distribution of firm productivity, we can easily obtain a distribution of wages that will allow us to calculate a measure of wage dispersion.

A firm with productivity φ offers wage $w(\varphi)$, and its share of workers is $\ell(\varphi) \equiv \frac{L(\varphi)}{\mathbb{L}}$. With $g(\varphi|\varphi \geq \hat{\varphi})$ denoting the productivity distribution of active firms, it follows that the productivity-based probability density function of wages is $h(\varphi) \equiv M\ell(\varphi)g(\varphi|\varphi \geq \hat{\varphi})$. The average wage across all workers is then given by $\bar{w}_A = \int_{\hat{\varphi}}^{\infty} w(\varphi)h(\varphi)d\varphi$, where the subscript A denotes "autarky"; note that the total wage bill (the first term on the right hand side of (12)) is equal to $\mathbb{L}\bar{w}_A$.

Applying a change of variables, we can also obtain the direct distribution of wages. Let $f(w)$ denote the probability density function of wages, and let $F(w)$ denote the cumulative distribution function. Hence, it must be the case that $h(\varphi) = f[w(\varphi)]w'(\varphi)$, so that, for example, the average wage can also be calculated as

$$\bar{w}_A = \int_{\hat{w}}^{\infty} wf(w)dw = \int_{\hat{\varphi}}^{\infty} w(\varphi)f[w(\varphi)]w'(\varphi)d\varphi,$$

where $\hat{w} = w(\hat{\varphi})$ is the lowest wage in the economy (paid by the least productive active firm). With $f[w(\varphi)] = \frac{h(\varphi)}{w'(\varphi)}$, under a Pareto distribution for productivity, and using (3) to write φ in terms of w, it follows that

$$f(w) = \frac{\gamma\hat{w}^{\gamma}}{w^{\gamma+1}} \quad \text{and} \quad F(w) = 1 - \left(\frac{\hat{w}}{w}\right)^{\gamma}, \tag{15}$$

which is also a Pareto distribution, where $\gamma \equiv \frac{k\sigma+\theta(k-\sigma+1)}{\sigma-1}$ is the parameter of wage dispersion (a higher γ implies less wage heterogeneity). Thus, the average wage is simply given by $\bar{w}_A = \frac{\gamma\hat{w}}{\gamma-1}$. The following lemma describes the response of the wage dispersion parameter to changes in θ, k, and σ.

Lemma 2. $\frac{d\gamma}{d\theta} > 0$, $\frac{d\gamma}{dk} > 0$, and $\frac{d\gamma}{d\sigma} < 0$.

This lemma states that wage dispersion declines (*i.e.*, γ rises) with either *(i)* less monopsony power in the labor market (a higher θ), *(ii)* less firm heterogeneity (a higher k), or *(iii)* more monopoly power in the goods market (a lower σ). When there is no monopsony power in the labor market ($\theta \to \infty$), then $\gamma \to \infty$ and the distribution collapses toward $\hat{w}$ so that all firms pay the same wage (the same happens if $k \to \infty$). On the other hand, as we approach full monopsony power ($\theta \to 0$), the dispersion parameter γ approaches its lower bound, $\frac{k\sigma}{\sigma-1}$, so that wage heterogeneity is the highest. The following proposition presents this section's main result.

Proposition 1. *(Monopsony power and wage inequality in autarky)*

In autarky, an increase in monopsony power (a reduction in θ) increases wage inequality.

The closed-economy model yields γ as the natural measure of wage inequality (a lower γ implies more wage dispersion, and hence more wage inequality). We can introduce, however, a more traditional measure of inequality: the Gini coefficient, which is always between 0 (perfect equality) and 1 (perfect inequality). Letting $\mathcal{G}_A$ denote the autarky Gini coefficient, we calculate it as

$$\mathcal{G}_A = \frac{1}{\bar{w}_A} \int_{\hat{w}}^{\infty} F(w)[1 - F(w)]dw = \frac{1}{2\gamma - 1}, \tag{16}$$

which is strictly decreasing in θ ($\frac{d\mathcal{G}_A}{d\theta} = -\frac{2}{(2\gamma-1)}\frac{d\gamma}{d\theta} < 0$). When $\theta \to 0$, inequality approaches its upper bound ($\mathcal{G}_A \to \frac{\sigma-1}{2k\sigma-\sigma+1}$), while when $\theta \to \infty$ we achieve perfect equality ($\mathcal{G}_A = 0$; *i.e.*, wages are homogeneous). Intuitively, more monopsony power in the labor market (a lower θ) increases inequality because it allows low-wage less productive firms to survive ($\hat{\varphi}$ is lower). By widening the productivity range of producing firms, the range of wages in the economy also widens.

4.7 Welfare Implications

We use two measures of welfare. First, since all income goes to workers and the number of workers is fixed, the average wage can be used as a measure of welfare as well as a measure of the productive efficiency of the economy. This would keep the results comparable to the models where the focus is on aggregate gains from trade. Second, since our model also has wage inequality, we also look at an alternative measure of social welfare that is sensitive to wage inequality.[14]

4.7.1 Impact of Monopsony Power on Aggregate Productive Efficiency

Looking at the average wage first, note that the wage bill, $\mathbb{W}$, can be written as

$$\mathbb{W} = M \int_{\hat{\varphi}}^{\infty} w(\varphi)L(\varphi)g(\varphi|\varphi \geq \hat{\varphi})d\varphi = \mathbb{L}\bar{w}_A = \frac{Mf\theta}{1+\theta}\left(\frac{k\beta}{k-\beta}\right) = \left(\frac{\sigma-1}{\sigma}\right)\left(\frac{\theta}{1+\theta}\right)\mathbb{A}.$$

where $\bar{w}_A$ is the average wage in the closed economy (subscript A denotes "autarky"). From the previous equation, we obtain that the elasticity of the wage bill (and the average wage) with respect

---

[14]Given the random utility of workers defined in (1), one could use the expected value of maximized utility of a worker as a measure of welfare which in our case is given by

$$E(\max_w v_i) = \ln \hat{w} + \frac{1}{\theta}\ln\left(\frac{\gamma}{\gamma - \theta}\right).$$

However, this measure of welfare does not capture the risk aversion of workers or policymakers appropriately because the workers face a deterministic wage distribution while choosing a job. Also, the random component of utility is not observable to policymakers, and therefore, we focus on a more conventional measure of social welfare which captures inequality aversion.

to the labor-supply elasticity is given by

$$
\begin{aligned}
\zeta_{\mathrm{W},\theta} \equiv \zeta_{\bar{w}_A,\theta} = {} & \zeta_{\mathbb{A},\theta} + \frac{1}{1+\theta} \\
= {} & \frac{\theta\beta}{(\sigma-1)(1+\theta)^2(k-\beta)} \left\{ \frac{\sigma-2}{\theta} \left(\frac{k\sigma-\sigma+1}{\sigma-1} \right) + \frac{\sigma(\sigma-1)\beta}{k\sigma+\theta(k-\sigma+1)} - \frac{k-\sigma+1}{\sigma-1} \right\}.
\end{aligned}
$$

Given that $k > \sigma - 1 > \beta$ and our assumption that $\sigma > 2$, from the term in braces we can see that the previous expression is positive for low levels of θ and is negative for high levels of θ. Therefore, there exists a $\theta < \infty$ that maximizes the wage bill, so that some monopsony power may be desirable to achieve production efficiency. Hence, the average wage is not maximized when $\theta \to \infty$; that is, the competitive labor market does not lead to production efficiency in this model. Below we provide an explanation for this result.

Standard models of monopolistic competition with constant elasticity of demand (with or without firm heterogeneity) are production efficient when the labor market is perfectly competitive. In these models, the fixed/sunk costs are also modeled in terms of labor. Given that we have a heterogeneous wage setting, we model fixed and sunk costs in terms of the final good, similar to Egger and Kreickemeier (2009, 2012) and Felbermayr, Impullitti, and Prat (2018). A consequence of modeling the fixed cost in terms of the numéraire is that the decentralized outcome is not necessarily production efficient. Thus, a social planner interested in maximizing aggregate net output wants to obtain more varieties than what the decentralized market would provide. This is because the final-good CES function has love of variety, so it makes sense for the planner to produce more varieties in smaller quantities and to cover their fixed costs. In a way, when a new firm enters, it confers a positive externality on other firms by lowering their entry cost and fixed cost of production because they are in terms of the final good. The planner internalizes this externality.

The impact of a monopsonistic labor market in this setting is the following. Since wages are marked down by firms, for a given number of firms profits are higher, which in turn induces entry. Therefore, the mass of firms is higher when the labor market is monopsonistic compared to a competitive labor market. As well, given the distortion in the competitive-market case discussed in the previous paragraph, monopsony power can alleviate the distortion by increasing the mass of firms in equilibrium. However, if the monopsony power is too high, there will be too many firms from the point of view of social optimality.

Since we have a heterogeneous-firm setting, there is an additional misallocation effect arising due to monopsony power. Recall our discussion earlier that monopsony power causes more productive firms to pay higher wages as they have to be further up their supply curves. An implication is that compared to a competitive labor market, more productive firms are smaller and less productive

firms are larger. This results in a smaller aggregate output for a given mass of firms. This effect by itself would tend to reduce aggregate output by reallocating labor from more productive firms towards less productive firms.

Because of the misallocation effect associated with firm heterogeneity, the level of θ at which the net aggregate output is maximized is increasing in the degree of firm heterogeneity. While it is theoretically possible for an increase in monopsony power to increase net aggregate output, for conventional parameter values this would occur at an extremely (and unrealistically) high level of θ.[15] For the levels of θ observed in empirical works mentioned in section 2 (which are rarely larger than 5), we expect net aggregate output (and the average wage) to be increasing in θ. That is, for empirically relevant levels of θ, an increase in monopsony power is likely to reduce welfare.

4.7.2 Inequality-Averse Social Welfare

Since the model generates wage inequality, it is worth looking at a more general social welfare function that is sensitive to inequality. Towards this end, we posit the following inequality-averse social welfare function:

$$V = \int_{\hat{w}}^{\infty} u(w)dF(w), \tag{17}$$

where $F(w)$ is the distribution function of wages, $u'(w) > 0$ and $u''(w) \leq 0$. The concavity of $u(\cdot)$ captures the inequality aversion of the social planner. We use the functional form

$$u(w) = \frac{w^{1-\rho} - 1}{1 - \rho}, \tag{18}$$

where $\rho \geq 0$ captures the degree of inequality aversion, and $\rho = 1$ corresponds to $u(\cdot)$ being logarithmic, which is also the functional form for the deterministic component of individual utility in (1).

As in Antràs, de Gortari, and Itskhoki (2017), we convert V into the following tractable form that allows the separation of social welfare into an average-wage component and an inequality component:

$$\mathcal{S}_A = [1 + (1 - \rho)V]^{\frac{1}{1-\rho}},$$

where we use $\mathcal{S}_A$ to denote social welfare in autarky. It follows from (17) and (18) that

$$\mathcal{S}_A = \left[\int_{\hat{w}}^{\infty} w^{1-\rho} f(w)dw \right]^{\frac{1}{1-\rho}}. \tag{19}$$

---

[15]In the open-economy case we do extensive numerical analysis with the following parameters: $\mathbb{L} = 1$, $\delta = 0.1$, $f_E = 0.25$, and $f = 0.1$, $\sigma = 3.8$ and $k = 3.4$. For these parameters, the average wage is maximized at $\theta = 240.6$.

Given the Pareto distribution of wages, equation (19) yields

$$\mathcal{S}_A = \Upsilon \bar{w}_A \quad \text{where} \quad \Upsilon \equiv \frac{\gamma - 1}{\gamma} \left[\frac{\gamma}{\rho + \gamma - 1} \right]^{\frac{1}{1-\rho}}.$$

It follows that $\frac{d\Upsilon}{d\gamma} > 0$ if $\rho > 0$ (and is zero if $\rho = 0$), and $\frac{d\Upsilon}{d\rho} < 0.$[16] Therefore, since wage inequality is inversely related to γ, social welfare is increasing in the average wage and decreasing in inequality. As well, social welfare is decreasing in the degree of inequality aversion, ρ. Also note that if $\rho = 0$, then $\Upsilon \equiv 1$ and $\mathcal{S}_A = \bar{w}_A$, so that the inequality-averse social welfare function nests the average wage as a special case.

From Lemma 2 we know that $\frac{d\gamma}{d\theta} > 0$, and hence, it must be the case that $\frac{d\Upsilon}{d\theta} > 0$ if $\rho > 0$ (and is zero if $\rho = 0$). We know from section 4.7.1 that $\frac{d\bar{w}_A}{d\theta} > 0$ in the empirically plausible range for θ, and therefore $\frac{d\mathcal{S}_A}{d\theta} > 0$ in that range, so that an increase in monopsony power (a decline in θ) reduces welfare. Our key findings on welfare in the closed economy are summarized as follows.

Key Result 1. (*Monopsony power and welfare in autarky*)

In autarky, an increase in monopsony power (a reduction in θ) has a non-monotonic relationship with the average wage. However, in the empirically plausible range of values of θ, an increase in monopsony power reduces the average wage. Using an inequality-averse social welfare function we find that social welfare declines as monopsony power increases. This happens because the average wage declines and inequality in the wage distribution increases.

5 The Open-Economy Model

The closed-economy model shows that monopsony power in the labor market causes more productive firms to pay higher wages, leveling the competition field for less productive firms—who are able to survive thanks to their lower wages—and generating lower aggregate productivity in the economy and an increase in wage inequality. But how do international trade possibilities affect these outcomes? This section extends the previous model to study the effects of monopsony power in a two-country setting, and analyzes how monopsony power and trade liberalization interact to affect wage inequality.

5.1 Preferences, Monopsonistic Labor Market, and Production

There are two identical countries—each of them populated by $\mathbb{L}$ workers—that produce and trade differentiated goods. In each country, preferences and demand are as in section 4.1, with the

---

[16]The result $\frac{\partial \Upsilon}{\partial \rho} < 0$ follows after using the inequality $\ln(x) < x - 1$ for $x > 0$, where x equals $\frac{\gamma}{\rho + \gamma - 1}$ for this case.

difference that the final good, Z, which is the numéraire, is now a CES composite of domestic and imported differentiated-good varieties. As well, the labor supply faced by a firm in each country is given by (2), with the wage elasticity, θ, assumed to be the same in both countries.[17]

In each country, producers of differentiated goods are heterogeneous in productivity. After paying an entry cost of f_E (in terms of the final good), each firm draws its productivity from the same cumulative distribution function, $G(\varphi)$. To access each market, a firm must pay a fixed cost in terms of the final good. For simplicity, we assume that the fixed costs to access the domestic and exports markets are identical and equal to f. In addition, there is an iceberg exporting cost, $\tau \geq 1$, so that an exporting firm must ship τ units of its good for one unit to reach the export market. Based on a firm's ability to cover these fixed costs and the iceberg exporting cost, the firm will either not produce for any market, produce only for the domestic market, or produce for both the domestic and export markets.

Let $s \in \{N, T\}$ indicate firm status, with N identifying *non-trading* firms (which only sell for the domestic market), and T identifying *trading* firms (which sell for the domestic and export markets). For a non-trading firm with productivity φ, its production function is $y_N(\varphi) = \varphi L$ and its profit maximization problem is similar to section 4.3, and thus equations (3), (4), (5), and (6) for firm-level wage, labor, price, and gross profits hold. To indicate firm status, these variables are respectively relabeled as $L_N(\varphi)$, $w_N(\varphi)$, $p_N(\varphi)$, and $\pi_N(\varphi)$.

On the other hand, a trading firm with productivity φ produces for the domestic (D) and export (X) markets, so that its total production equals $y_T(\varphi) \equiv y_D(\varphi) + y_X(\varphi) = \varphi L$, where $y_r(\varphi)$ is the firm's production for market $r \in \{D, X\}$. The profit-maximization problem for this firm is

$$\max\{p_D(\varphi)q_D(\varphi) + p_X(\varphi)q_X(\varphi) - w_T(\varphi)\left[L_D(\varphi) + L_X(\varphi)\right]\},$$

where $q_r(\varphi) \equiv \mathbb{A}p_r(\varphi)^{-\sigma}$ is the demand faced by the firm in market r, with $\mathbb{A}$ denoting the total expenditure in differentiated goods in each country, $p_r(\varphi)$ is the price set by the firm in market r, and $L_r(\varphi)$ is the amount of labor hired by the firm to produce for market r—the total amount of labor hired by the firm is $L_T(\varphi) \equiv L_D(\varphi) + L_X(\varphi)$. From the upward-sloping labor supply in (2), it must hold that $w_T(\varphi) = \left[\frac{L_T(\varphi)}{\mathbb{B}}\right]^{\frac{1}{\theta}}$. Moreover, equilibrium in each differentiated-good's market requires that $q_D(\varphi) = y_D(\varphi)$ and $\tau q_X(\varphi) = y_X(\varphi)$ (a trading firm must produce τ units for each unit of the good that is consumed in the export market), and thus, the inverse demand functions can be written as $p_D(\varphi) = \left[\frac{\varphi L_D(\varphi)}{\mathbb{A}}\right]^{-\frac{1}{\sigma}}$ and $p_X(\varphi) = \left[\frac{\varphi L_X(\varphi)}{\tau \mathbb{A}}\right]^{-\frac{1}{\sigma}}$.

---

[17]To present the most parsimonious model, we abstract from asymmetries in the degrees of monopsony power across countries. We leave the study of this and other asymmetries for future work.

After rewriting the profit-maximization problem of a trading firm with productivity φ in terms of $L_D(\varphi)$ and $L_X(\varphi)$, we obtain as solution

$$L_D(\varphi) = \left(1 + \frac{1}{\tau^{\sigma-1}}\right)^{-\frac{\sigma}{\sigma+\theta}} L_N(\varphi) \quad \text{and} \quad L_X(\varphi) = \frac{1}{\tau^{\sigma-1}}\left(1 + \frac{1}{\tau^{\sigma-1}}\right)^{-\frac{\sigma}{\sigma+\theta}} L_N(\varphi),$$

where $L_N(\varphi)$ is given by (4). The total amount of labor employed by the firm is then $L_T(\varphi) = \left(1 + \frac{1}{\tau^{\sigma-1}}\right)^{\frac{\theta}{\sigma+\theta}} L_N(\varphi)$. Notice that $1 + \frac{1}{\tau^{\sigma-1}} \in (1,2)$, so that it is always true that

$$L_X(\varphi) < L_D(\varphi) < L_N(\varphi) < L_T(\varphi).$$

Hence, a firm with productivity φ employs more labor if it exports, but the amount of labor it hires to produce for its domestic market is smaller than if it were a non-trading firm. This illustrates the impact of an upward labor supply: exporting makes a firm bigger, but hiring more labor implies a higher equilibrium wage for the firm, which makes it less competitive in the domestic market as it must set a higher price.

Verifying that exporting makes domestic operations costlier, we obtain that the wage offered by a trading firm with productivity φ and its domestic price are given by

$$w_T(\varphi) = \left(1 + \frac{1}{\tau^{\sigma-1}}\right)^{\frac{1}{\sigma+\theta}} w_N(\varphi) \quad \text{and} \quad p_D(\varphi) = \left(\frac{\sigma}{\sigma-1}\right)\left[\left(\frac{1+\theta}{\theta}\right)\frac{w_T(\varphi)}{\varphi}\right],$$

where $w_N(\varphi)$ is given by (3), so that it is always the case that $w_T(\varphi) > w_N(\varphi)$ and $p_D(\varphi) > p_N(\varphi)$, with $p_N(\varphi)$ given by (5). The price set by the trading firm in the export market is then $p_X(\varphi) = \tau p_D(\varphi)$. Lastly, the gross profit of a trading firm with productivity φ can be written as

$$\pi_T(\varphi) = \left(1 + \frac{1}{\tau^{\sigma-1}}\right)^{\frac{1+\theta}{\sigma+\theta}} \pi_N(\varphi),$$

where $\pi_N(\varphi)$ is given by (6). Therefore, in spite of a higher wage and lower domestic sales, the gross profits of a firm with productivity φ are always larger if it exports.

5.2 Indifference Conditions, Free-Entry, and Aggregate Productivity

The open-economy model solves for two cutoff productivity levels, $\hat{\varphi}_N$ and $\hat{\varphi}_T$, that determine firm status: given $\hat{\varphi}_N < \hat{\varphi}_T$, a firm with productivity φ does not produce if $\varphi < \hat{\varphi}_N$, produces only for the domestic market if $\varphi \in [\hat{\varphi}_N, \hat{\varphi}_T)$, and produces for the domestic and export markets if $\varphi \geq \hat{\varphi}_T$. These cutoff levels satisfy the indifference conditions

$$\pi_N(\hat{\varphi}_N) = f,$$
$$\pi_T(\hat{\varphi}_T) - 2f = \pi_N(\hat{\varphi}_T) - f,$$

where the first condition is similar to the zero-cutoff-profit condition in the closed-economy model—the net profit of a non-trading firm at $\hat{\varphi}_N$ must be zero—while the second condition indicates that a firm at the cutoff $\hat{\varphi}_T$ is indifferent between being a trading firm or a non-trading firm (*i.e.*, its net profits if it is a trading firm are the same as the net profits if it were a non-trading firm).[18] Notice that the determination of the exporting cutoff is different from the standard Melitz model, in which a firm exports if the gross profits obtained in the export market are greater or equal than the fixed cost of exporting. Here, since firm size affects wages, the impact of exporting on domestic sales will enter into the calculation as well.

Using the gross profit equation for trading firms at the end of section 5.1, we rewrite the second indifference condition as $\left[\left(1 + \frac{1}{\tau^{\sigma-1}}\right)^{\frac{1+\theta}{\sigma+\theta}} - 1\right]\pi_N(\hat{\varphi}_T) = f$. With $\pi_N(\varphi)$ given by (6), we can then combine the two indifference conditions to obtain

$$\hat{\varphi}_T = \lambda\hat{\varphi}_N, \quad \text{where} \quad \lambda \equiv \frac{\tau}{\left[(1+\tau^{\sigma-1})^{\frac{\beta}{\sigma-1}} - \tau^\beta\right]^{\frac{1}{\beta}}} > \tau. \tag{20}$$

This is one of the two equations we need to solve for $\hat{\varphi}_N$ and $\hat{\varphi}_T$. The term λ is strictly decreasing in θ, $\frac{d\lambda}{d\theta} < 0$, and approaches τ as $\theta \to \infty$. Thus, an increase in monopsony power in the labor market (*i.e.*, a lower θ), widens the productivity gap between the least efficient non-trading firm and the least efficient trading firm.

Similar to (7), we use the indifference conditions to rewrite $\pi_N(\varphi)$ and $\pi_T(\varphi)$ as

$$\pi_N(\varphi) = \left(\frac{\varphi}{\hat{\varphi}_N}\right)^\beta f \quad \text{and} \quad \pi_T(\varphi) = \left(1 + \lambda^\beta\right)\left(\frac{\varphi}{\hat{\varphi}_T}\right)^\beta f. \tag{21}$$

The free-entry condition is then given by $\frac{1}{\delta}\left\{\int_{\hat{\varphi}_N}^{\hat{\varphi}_T}[\pi_N(\varphi) - f]g(\varphi)d\varphi + \int_{\hat{\varphi}_T}^\infty[\pi_T(\varphi) - 2f]g(\varphi)d\varphi\right\} = f_E$, which using (21) is written as

$$\frac{1}{\delta}\left\{\int_{\hat{\varphi}_N}^{\hat{\varphi}_T}\left[\left(\frac{\varphi}{\hat{\varphi}_N}\right)^\beta - 1\right]fg(\varphi)d\varphi + \int_{\hat{\varphi}_T}^\infty\left[\left(1 + \lambda^\beta\right)\left(\frac{\varphi}{\hat{\varphi}_T}\right)^\beta - 2\right]fg(\varphi)d\varphi\right\} = f_E. \tag{22}$$

Using (20) and (22), and under the assumed Pareto distribution for productivity, the equilibrium cutoff productivity levels are

$$\hat{\varphi}_N = \frac{1}{\lambda}\left[\left(1 + \lambda^k\right)\left(\frac{f}{\delta f_E}\right)\frac{\beta}{k - \beta}\right]^{\frac{1}{k}} \quad \text{and} \quad \hat{\varphi}_T = \left[\left(1 + \lambda^k\right)\left(\frac{f}{\delta f_E}\right)\frac{\beta}{k - \beta}\right]^{\frac{1}{k}}. \tag{23}$$

We can now see how more monopsony power (as captured by a lower θ) affects the cutoff productivity levels. From the solution in (23), the elasticities of the cutoff levels with respect to θ

---

[18]On the left-hand side of the second condition we subtract $2f$ from the gross profits because a trading firm must pay a fixed cost f to access each market.

are

$$\zeta_{\hat{\varphi}_N,\theta} = \frac{\theta}{(1+\theta)^2}\left\{\frac{\beta}{k-\beta} + \frac{1}{1+\lambda^k}\left[\beta\ln\lambda + (1+\lambda^\beta)\ln\left(1+\frac{1}{\lambda^\beta}\right)\right]\right\},$$

$$\zeta_{\hat{\varphi}_T,\theta} = \frac{\theta}{(1+\theta)^2}\left\{\frac{\beta}{k-\beta} - \frac{\lambda^k}{1+\lambda^k}\left[\beta\ln\lambda + (1+\lambda^\beta)\ln\left(1+\frac{1}{\lambda^\beta}\right)\right]\right\}.$$

While $\zeta_{\hat{\varphi}_N,\theta} > 0$, so that $\hat{\varphi}_N$ is strictly increasing in θ, we obtain that $\zeta_{\hat{\varphi}_T,\theta}$ is negative for low levels of θ and positive for high levels, so that $\hat{\varphi}_T$ has a U-shaped relationship with θ. On the other hand, and as in the conventional Melitz model, trade liberalization increases $\hat{\varphi}_N$ and decreases $\hat{\varphi}_T$: $\zeta_{\hat{\varphi}_N,\tau} = -\left(\frac{1}{1+\lambda^k}\right)\left[\frac{\lambda}{\tau(1+\lambda^\beta)^{\frac{1}{1+\theta}}}\right]^{\sigma-1} < 0$ and $\zeta_{\hat{\varphi}_T,\tau} = \left(\frac{\lambda^k}{1+\lambda^k}\right)\left[\frac{\lambda}{\tau(1+\lambda^\beta)^{\frac{1}{1+\theta}}}\right]^{\sigma-1} > 0$. We summarize the results in the following lemma.

Lemma 3. *The survival cutoff, $\hat{\varphi}_N$, is decreasing in monopsony power but the exporting cutoff, $\hat{\varphi}_T$, has a non-monotonic relationship with monopsony power. Trade liberalization increases the survival cutoff, $\hat{\varphi}_N$, and reduces the exporting cutoff, $\hat{\varphi}_T$.*

Figure 1 presents a numerical representation of these results. Without loss of generality, we set $\mathbb{L} = 1$, $\delta = 0.1$, $f_E = 0.25$, and $f = 0.1$. From Ghironi and Melitz (2005), we set $\sigma = 3.8$ and $k = 3.4$. Most of the figures below look at the relationship between θ and variables of interest for different levels of iceberg trade costs: $\tau^L = 1$, $\tau^M = 1.2$, and $\tau^H = 1.5$, where superscripts indicate low, middle, or high levels. When looking at the relationship between τ and a variable of interest, we set low, middle, and high levels for θ as $\theta^L = 0.15$, $\theta^M = 1.8$, $\theta^H = 6$.

The behavior of $\hat{\varphi}_N$ is similar to $\hat{\varphi}$ in the closed-economy model: more monopsony power in the labor market (a lower θ) allows less productive firms to survive. The most interesting result in Figure 1, however, is the U-shape behavior of $\hat{\varphi}_T$. As firms gain monopsony power (moving from $\theta \to \infty$ to the left), exporting becomes easier, but for low levels of θ (a high level of monopsony power), further declines in θ start to inhibit exporting. We can understand the intuition for this result in terms of two counteracting effects that an increase in monopsony power has on exporting.

First, as monopsony power increases, exporting becomes easier because as mentioned earlier, monopsony power dampens the advantage of high-productivity firms. This by itself induces more firms to export. Second, an increase in monopsony power can also make exporting more difficult because as the labor supply curve becomes steeper, exporting requires larger wage increases, reducing a firm's profitability in the domestic market and thereby making exporting less attractive. If the degree of monopsony power is not too high, the first effect dominates and as a result, an increase in monopsony power reduces $\hat{\varphi}_T$. But beyond a point the second effect starts dominating, and hence $\hat{\varphi}_T$ starts increasing with further increases in monopsony power. When we are close to
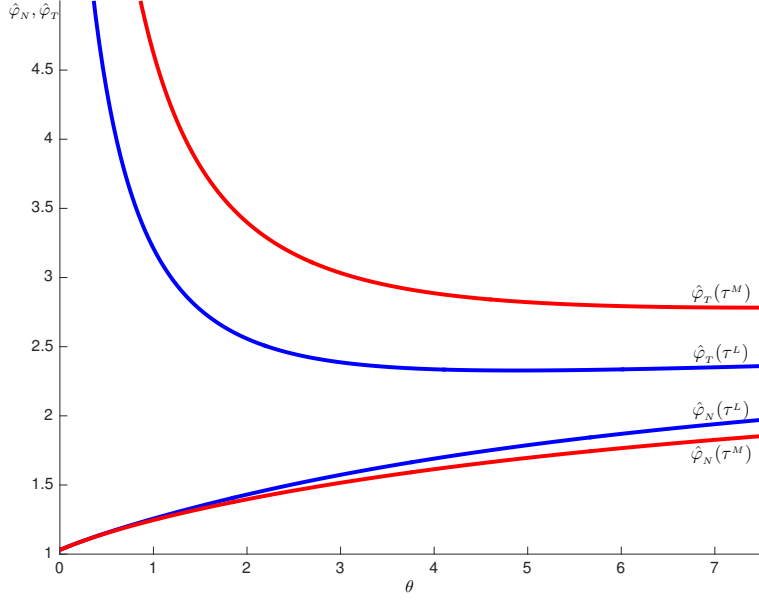
Figure 1: Monopsony power and $\hat{\varphi}_N$ and $\hat{\varphi}_T$

absolute monopsony power ($\theta \to 0$), $\hat{\varphi}_T$ is at its highest and there is almost no exporting. This is a consequence of all firms having near zero marginal costs, so that any productivity advantage of the most productive firms will not be sufficiently large to allow the firm to cover the fixed costs of exporting—recall from the closed-economy model that monopsony power in the labor market dampens the effect that firm heterogeneity has in the allocation of market shares.

As in Melitz (2003), trade liberalization (a decline in τ) wipes out the least productive firms. However, the elasticity is smaller (in absolute value) the lower θ is, $\frac{d|\zeta_{\hat{\varphi}_N,\tau}|}{d\theta} > 0$, and hence, monopsony power inhibits the response of $\hat{\varphi}_N$ to trade liberalization. A reduction in τ raises $\hat{\varphi}_N$, which reduces wage dispersion and thus would push for lower inequality, but this effect is dampened when there is monopsony power in the labor market. As in Melitz, the cutoff for trading firms, $\hat{\varphi}_T$, declines after a reduction in τ, but in contrast to $\hat{\varphi}_N$, the response is magnified (the elasticity is larger) with more monopsony power: the lower θ is, the larger $\zeta_{\hat{\varphi}_T,\tau}$ is ($\frac{d\zeta_{\hat{\varphi}_T,\tau}}{d\theta} < 0$). To sum up, with a high level of monopsony power (a low θ), a reduction in τ increases $\hat{\varphi}_N$ only a little but reduces $\hat{\varphi}_T$ a lot.

As in the closed-economy model, the average productivity in each country is given by

$$\bar{\varphi} = \left[\int_{\hat{\varphi}_N}^{\infty} \varphi^{\beta} g(\varphi|\varphi \geq \hat{\varphi}_N) d\varphi \right]^{\frac{1}{\beta}} = \left(\frac{k}{k-\beta} \right)^{\frac{1}{\beta}} \hat{\varphi}_N,$$

27

which is increasing in θ. We can also obtain the average productivity for each type of firm as

$$\bar{\varphi}_N = \left[\int_{\hat{\varphi}_N}^{\hat{\varphi}_T} \varphi^\beta g(\varphi | \varphi \in [\hat{\varphi}_N, \hat{\varphi}_T)) d\varphi \right]^{\frac{1}{\beta}} = \left[\frac{k}{k-\beta} \left(\frac{\lambda^k - \lambda^\beta}{\lambda^k - 1} \right) \right]^{\frac{1}{\beta}} \hat{\varphi}_N,$$

$$\bar{\varphi}_T = \left[\int_{\hat{\varphi}_T}^{\infty} \varphi^\beta g(\varphi | \varphi \geq \hat{\varphi}_T) d\varphi \right]^{\frac{1}{\beta}} = \left(\frac{k}{k-\beta} \right)^{\frac{1}{\beta}} \hat{\varphi}_T,$$

with the overall average productivity calculated as $\bar{\varphi} = \left[\frac{M_N}{M} \bar{\varphi}_N^\beta + \frac{M_T}{M} \bar{\varphi}_T^\beta \right]^{\frac{1}{\beta}}$, where M_s denotes the mass of firms with status $s \in \{N, T\}$ in each country, and $M \equiv M_N + M_T$ is the total mass of producers in each country.

5.3 Equilibrium in the Open-Economy Model

Every period in each country, the mass of producing firms that die due to the death shock, δM, is exactly replaced by the mass of successful entrants, $[1 - G(\hat{\varphi}_N)]M_E$, and thus, we can write M as $M = [1 - G(\hat{\varphi}_N)]\frac{M_E}{\delta}$. By firm status, the masses of non-trading and trading firms in terms of M_E are respectively given by $M_N = [G(\hat{\varphi}_T) - G(\hat{\varphi}_N)]\frac{M_E}{\delta}$ and $M_T = [1 - G(\hat{\varphi}_T)]\frac{M_E}{\delta}$. With identical countries, the total mass of varieties consumed in each country is $M_N + 2M_T$, where $M_N + M_T$ are domestic varieties, and M_T varieties are imported.

The aggregate price P—which is a weighted average of non-trading firms' prices, trading firms' prices, and import prices—can be conveniently written as

$$P = \left\{ M_N p_N (\bar{\varphi}_N)^{1-\sigma} + M_T \left[p_D (\bar{\varphi}_T)^{1-\sigma} + p_X (\bar{\varphi}_T)^{1-\sigma} \right] \right\}^{\frac{1}{1-\sigma}}$$

$$= \left[M_N p_N (\bar{\varphi}_N)^{1-\sigma} + M_T \left(1 + \frac{1}{\tau^{\sigma-1}} \right)^{\frac{1+\theta}{\sigma+\theta}} p_N (\bar{\varphi}_T)^{1-\sigma} \right]^{\frac{1}{1-\sigma}},$$

where the second equality uses the relationships between $p_N(\varphi)$, $p_D(\varphi)$, and $p_X(\varphi)$ described in section 5.1. Using the expression for $p_N(\varphi)$ in (5), the previous expression for M_N and M_T in terms of M_E, and given our assumption that $P = 1$ (the final good is the numéraire), it follows that we can write the mass of entrants as

$$M_E = \left\{ \left[\frac{\sigma(1+\theta)}{(\sigma-1)\theta} \right]^\theta \frac{\mathbb{A}}{\mathbb{B}} \right\}^{\frac{\sigma-1}{\sigma+\theta}} \left\{ \frac{\delta}{[G(\hat{\varphi}_T) - G(\hat{\varphi}_N)]\bar{\varphi}_N^\beta + [1 - G(\hat{\varphi}_T)](1 + \tau^{1-\sigma})^{\frac{1+\theta}{\sigma+\theta}} \bar{\varphi}_T^\beta} \right\}, \qquad (24)$$

with average productivities $\bar{\varphi}_N$ and $\bar{\varphi}_T$ defined as at the end of section 5.2.

Similar to the closed economy case, the aggregate expenditure on final goods in each country is the sum of workers' expenditure on final-good consumption and firms' final-good requirements to cover the fixed and entry costs. In an open economy, consumption expenditure on each country's

final good comes from both countries workers. This difference, however, is inconsequential because exports of a country exactly cancel out with its imports, so that consumption expenditure in each country's final good continues to be equal to the wage bill, $\mathbb{W}$. Therefore, the total expenditure on each country's final good is $\mathbb{A} = \mathbb{W} + M_N f + 2M_T f + M_E f_E$, which considers that trading firms must pay f for each market they access. Using the expressions for M_N and M_T above, we rewrite $\mathbb{A}$ as

$$\mathbb{A} = \frac{M_E}{\delta} \int_{\hat{\varphi}_N}^{\infty} w(\varphi)L(\varphi)g(\varphi)d\varphi + \frac{M_E}{\delta}\left[2 - G(\hat{\varphi}_N) - G(\hat{\varphi}_T)\right]f + M_E f_E. \tag{25}$$

where

$$L(\varphi) = \begin{cases} L_N(\varphi) \ \text{if} \ \varphi \in [\hat{\varphi}_N, \hat{\varphi}_T) \\ L_T(\varphi) \equiv L_D(\varphi) + L_X(\varphi) \ \text{if} \ \varphi \geq \hat{\varphi}_T \end{cases} \quad \text{and} \quad w(\varphi) = \begin{cases} w_N(\varphi) \ \text{if} \ \varphi \in [\hat{\varphi}_N, \hat{\varphi}_T) \\ w_T(\varphi) \ \text{if} \ \varphi \geq \hat{\varphi}_T, \end{cases} \tag{26}$$

with $L_s(\varphi)$, $L_r(\varphi)$, and $w_s(\varphi)$ for $s \in \{N, T\}$ and $r \in \{D, X\}$ described as in section 5.1.

Finally, the labor-supply shifter, $\mathbb{B} \equiv \left[\int_{\nu \in \Omega} w(\nu)^{\theta} d\nu\right]^{-1} \mathbb{L}$, of the firm-level labor supply in (2) can also be written in terms of M_E as

$$\mathbb{B} = \left[\frac{M_E}{\delta} \int_{\hat{\varphi}_N}^{\infty} w(\varphi)^{\theta} g(\varphi)d\varphi\right]^{-1} \mathbb{L}. \tag{27}$$

We can now define the equilibrium in the open-economy model.

Definition. *An open-economy equilibrium solves for $\hat{\varphi}_N$ and $\hat{\varphi}_T$ from (20) and (22), and then solves for M_E, $\mathbb{A}$, and $\mathbb{B}$ from (24), (25), and (27), with $L(\varphi)$ and $w(\varphi)$ defined as in (26).*

Under the assumed Pareto distribution for productivity, the solution for $\hat{\varphi}_N$ and $\hat{\varphi}_T$ is given in (23). Once we obtain these and the solutions for M_E, $\mathbb{A}$, and $\mathbb{B}$, we can solve for the distinct masses of firms, average productivities, average wages, and average prices.

Figure 2 shows how the masses of entrants, non-trading, and trading firms respond to changes in the labor-supply elasticity, θ, for various configurations of the iceberg trade cost ($\tau^L = 1$, $\tau^M = 1.2$, and $\tau^H = 1.5$). Note from Figure 2a that entry is the lowest when $\theta \to 0$, and thus, a high level of monopsony power in the labor market is associated with low levels of entry in the economy. This, however, does not imply that the amount of producers in each country is small. Figure 2b shows a large number of non-trading firms when θ is small, which reflects the higher rates of survival due to monopsony power shielding the least productive firms from competition. On the other hand, Figure 2c shows that the number of trading firms approaches zero as $\theta \to 0$: with almost absolute monopsony power in the labor market (so that firms pay almost zero wages), firms' *effective* productivities are almost identical—recall that monopsony power dampens the effect of

(a) The mass of entrants

(b) The mass of non-trading firms

(c) The mass of trading firms

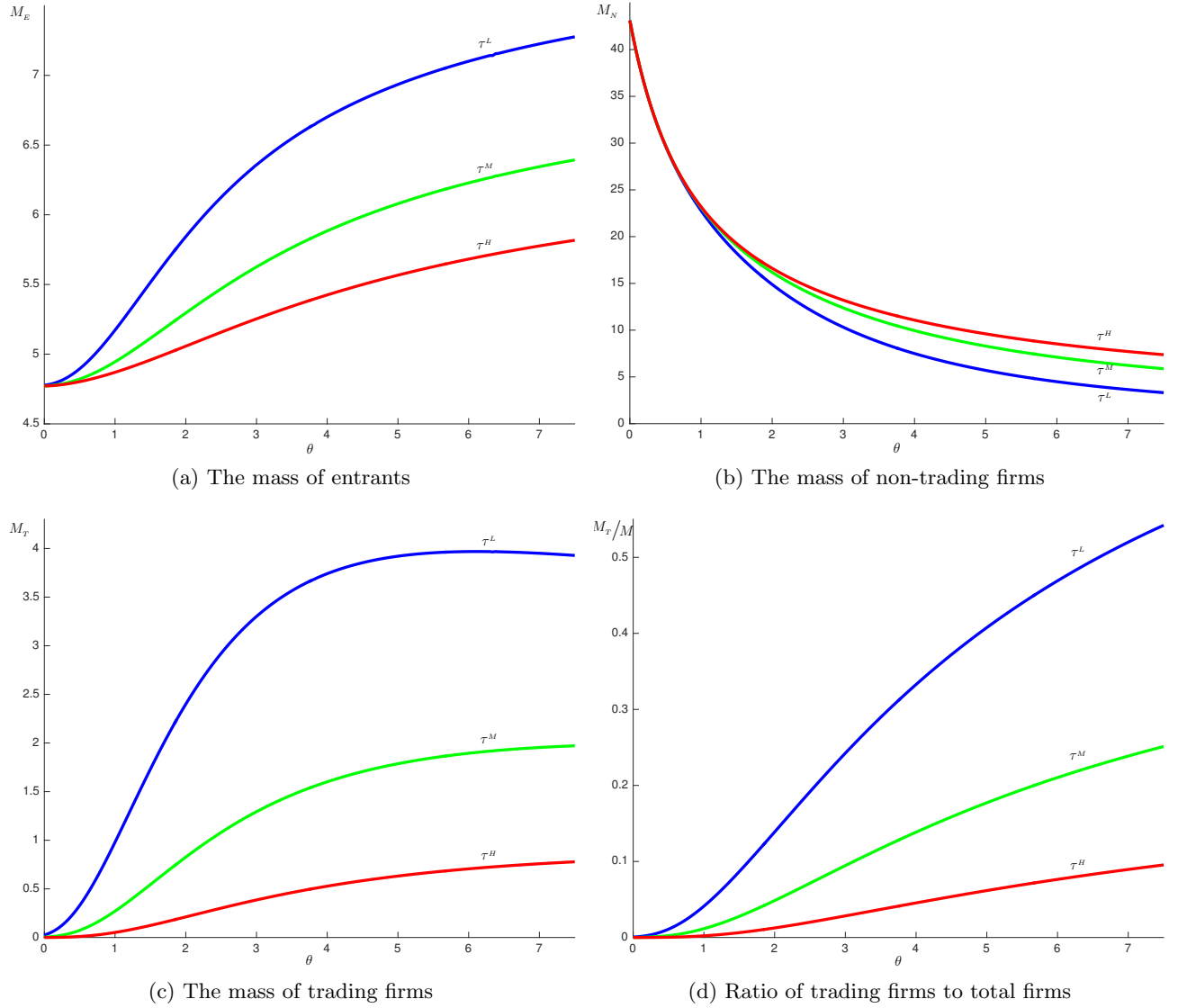(d) Ratio of trading firms to total firms

Figure 2: Monopsony power and the masses of firms

idiosyncratic productivity differences across firms—so that gross profit opportunities from exporting for the most productive firms are not large enough to cover the fixed costs of exporting.

Lastly, Figure 2d shows that, no matter the level of trade costs, more monopsony power reduces the share of trading firms in the economy. Analytically, $\frac{M_T}{M} = \frac{1-G(\hat{\varphi}_T)}{1-G(\hat{\varphi}_N)} = \frac{1}{\lambda^k}$, where λ is defined as in (20). From section 5.3 we know that $\frac{d\lambda}{d\theta} < 0$, which then implies that $\frac{d(M_T/M)}{d\theta} > 0$.[19] Therefore, in spite of the U-shaped relationship between $\hat{\varphi}_T$ and θ, in this model it is always true that monopsony power reduces the share of exporters in each country.

---

[19] An easy way to verify that $\frac{d\lambda}{d\theta} < 0$ is by noticing from $\hat{\varphi}_T = \lambda \hat{\varphi}_N$ that $\zeta_{\lambda,\theta} = \zeta_{\hat{\varphi}_T,\theta} - \zeta_{\hat{\varphi}_N,\theta}$. From the expressions for $\zeta_{\hat{\varphi}_T,\theta}$ and $\zeta_{\hat{\varphi}_N,\theta}$ in section 5.2, it easily follows that $\zeta_{\lambda,\theta} < 0$.

5.4 Monopsony Power and Inequality in the Open Economy

In the closed-economy model, the Pareto parameter of the wage distribution, $\gamma \equiv \frac{k\sigma + \theta(k-\sigma+1)}{\sigma-1}$, summarizes the relationship between monopsony power in the labor market and inequality. More monopsony power (a lower θ), reduces γ, which increases wage dispersion and thus inequality. In terms of productivity, this effect is captured by a reduction in the cutoff productivity level, so that low-productivity firms (which pay lower wages) are more likely to survive after a reduction in θ. In an open economy the mechanisms of action are richer, as a decline in θ reduces the cutoff level for non-trading firms, $\hat{\varphi}_N$, but may reduce or increase the cutoff level for trading firms, $\hat{\varphi}_T$. Given that the degree of inequality is not solely captured by γ, the Gini coefficient takes a greater relevance in an open-economy setting.

Let us first obtain the wage distribution. Each non-trading firm with productivity φ offers wage $w_N(\varphi)$, so that the fraction of workers receiving this wage is $\ell_N(\varphi) \equiv \frac{M_N L_N(\varphi)}{\mathbb{L}}$. With $g(\varphi|\varphi \in [\hat{\varphi}_N, \hat{\varphi}_T))$ denoting the productivity distribution of non-trading active firms, it follows that the productivity-based probability density function of wages is $h_N(\varphi) \equiv \ell_N(\varphi)g(\varphi|\varphi \in [\hat{\varphi}_N, \hat{\varphi}_T))$ for $\varphi \in [\hat{\varphi}_N, \hat{\varphi}_T)$. Similarly, the productivity-based probability density function of wages is $h_T(\varphi) \equiv \ell_T(\varphi)g(\varphi|\varphi \geq \hat{\varphi}_T)$ for $\varphi \geq \hat{\varphi}_T$, with $\ell_T(\varphi) \equiv \frac{M_T L_T(\varphi)}{\mathbb{L}}$. Therefore, the average wage across all workers in each country is $\bar{w} = \int_{\hat{\varphi}_N}^{\infty} w(\varphi)h(\varphi)d\varphi$ where $h(\varphi) = h_N(\varphi)$ if $\varphi \in [\hat{\varphi}_N, \hat{\varphi}_T)$, $h(\varphi) = h_T(\varphi)$ if $\varphi \geq \hat{\varphi}_T$, and $w(\varphi)$ is given by (26).

Similar to section 4.6, we apply a change of variables to obtain the direct distribution of wages. Letting $F(w)$ denote the cumulative distribution function for wages, we obtain

$$
F(w) = \begin{cases} \alpha_N \left[\frac{1}{1-\lambda^{-\beta\gamma/(1+\theta)}} \right] \left[1 - \left(\frac{\hat{w}_N}{w} \right)^{\gamma} \right] & \text{if } w \in \left[\hat{w}_N, \lambda^{\frac{\beta}{1+\theta}} \hat{w}_N \right) \\ \alpha_N & \text{if } w \in \left[\lambda^{\frac{\beta}{1+\theta}} \hat{w}_N, \hat{w}_T \right) \\ \alpha_N + \alpha_T \left[1 - \left(\frac{\hat{w}_T}{w} \right)^{\gamma} \right] & \text{if } w \geq \hat{w}_T, \end{cases}
\tag{28}
$$

where $\alpha_N \equiv \frac{\lambda^k - \lambda^{\beta\theta/(1+\theta)}}{\lambda^k + (1+\lambda^\beta)^{\theta/(1+\theta)} - \lambda^{\beta\theta/(1+\theta)}}$ is the fraction of workers in non-trading firms, $\alpha_T \equiv 1 - \alpha_N$ is the fraction of workers in trading firms, $\hat{w}_N = w_N(\hat{\varphi}_N)$, $\hat{w}_T = w_T(\hat{\varphi}_T)$, $\hat{w}_T = (1+\lambda^\beta)^{\frac{1}{1+\theta}} \hat{w}_N$, and γ is defined as above. Recall that at productivity $\hat{\varphi}_T$ firms have a jump in their wages as they become exporters; equation (28) shows that the wage jump is from $\lambda^{\frac{\beta}{1+\theta}} \hat{w}_N$ to $\hat{w}_T$.

After deriving $f(w)$ from (28), we get an average wage of

$$
\bar{w} = \alpha_N \left[\frac{1}{1 - \lambda^{-\beta\gamma/(1+\theta)}} \right] \left(1 + \frac{1}{\lambda^k} \right) \left(\frac{\gamma \hat{w}_N}{\gamma - 1} \right).
\tag{29}
$$

Moreover, letting $\mathcal{G} = \frac{1}{\bar{w}} \int_{\hat{w}}^{\infty} F(w)[1 - F(w)]dw$ denote the open-economy Gini coefficient, we use
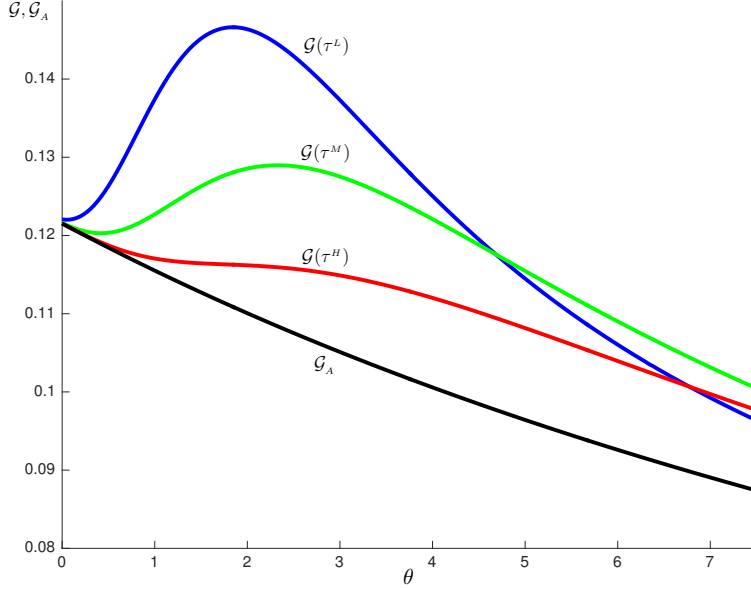
Figure 3: θ and the Gini coefficient

(28) and (29) to obtain

$$
\mathcal{G} = \frac{1}{2\gamma - 1} \left[1 + \frac{2(\gamma - 1)\alpha_N}{1 + \lambda^k} \left\{ 1 - \lambda^\beta \left[\left(\frac{1}{\lambda^\beta} + 1 \right)^{\frac{\theta}{1+\theta}} - 1 \right] \left[\frac{\gamma(\lambda^{\frac{\beta(\gamma-1)}{1+\theta}} - 1)}{(\gamma - 1)(\lambda^{\frac{\beta\gamma}{1+\theta}} - 1)} \right] \right\} \right], \tag{30}
$$

which approaches $\mathcal{G}_A = \frac{1}{2\gamma - 1}$ as $\tau \to \infty$. Comparing $\mathcal{G}$ with $\mathcal{G}_A$ yields the following proposition.

Proposition 2. *(Inequality in the open economy and in autarky)*

For every τ and θ, it is always true that $\mathcal{G} > \mathcal{G}_A$: inequality in the open economy is always higher than inequality in the closed economy.

To prove this proposition, it is sufficient to show that the term in braces in (30) is between 0 and 1, which follows from $\lambda^\beta \left[\left(\frac{1}{\lambda^\beta} + 1 \right)^{\theta/(1+\theta)} - 1 \right] \in (0, 1)$ and $\frac{\gamma(\lambda^{\beta(\gamma-1)/(1+\theta)}-1)}{(\gamma-1)(\lambda^{\beta\gamma/(1+\theta)}-1)} \in (0, 1)$.[20] In contrast to the closed-economy model, however, inequality does not necessarily increase as θ declines. Figure 3 shows the relationship between θ and the Gini coefficient for different levels of τ, as well as for the autarky case ($\mathcal{G}_A$). For high levels of τ (the $\mathcal{G}(\tau^H)$ line) we get closer to the closed-economy case and we can observe $\frac{d\mathcal{G}}{d\theta} < 0$ for all levels of θ. For lower levels of τ, the Gini coefficient reaches a maximum that is higher than the upper bound in the closed-economy case. Note also that there are reductions in inequality when we move from the $\mathcal{G}(\tau^M)$ line to the $\mathcal{G}(\tau^L)$ line, when θ is sufficiently high. That is, in that range of θ, trade liberalization reduces inequality.

---

[20]To prove the last part, we use the log inequality $\ln x^\gamma > \frac{x^\gamma - 1}{x^\gamma}$ for $x > 0$, to show that $\frac{\lambda^{\beta\gamma/(1+\theta)}-1}{\gamma}$ is increasing in θ.
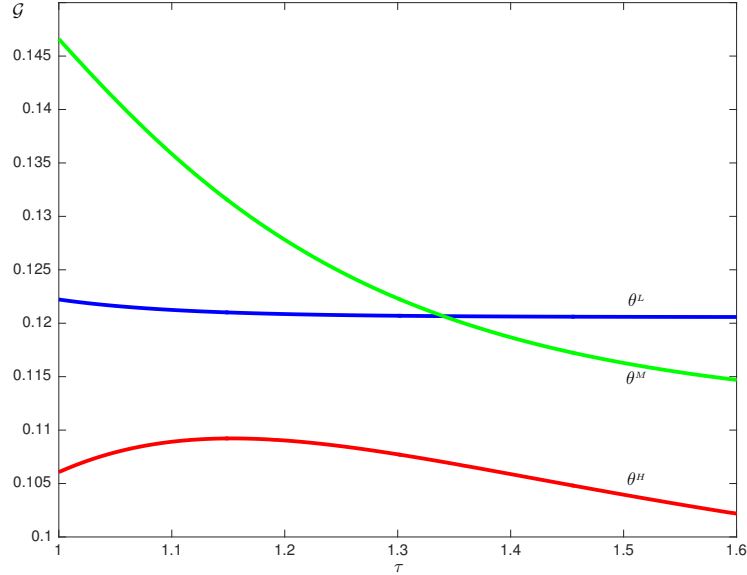
Figure 4: τ and the Gini coefficient

To look further into this, we analyze the relationship between τ and $\mathcal{G}$ for different levels of θ. Figure 4 presents that relationship for low, medium and high levels of θ ($\theta^L = 0.15$, $\theta^M = 1.8$, and $\theta^H = 6$). When θ is low, the Gini coefficient increases as τ declines, but it does not move that much (the θ^L line is very flat): with almost full monopsony power, there are almost no exporters and hence trade liberalization has little effect on inequality. For θ^M, which is high enough so that exporting is no longer a very rare activity, trade liberalization causes large drops in $\hat{\varphi}_T$. The θ^M line in Figure 4 shows that this leads to large increases in inequality, and here is where we reach inequality levels that are higher than the upper bound in the closed economy. Intuitively, under θ^M trade liberalization causes a large increase in the number of exporters, which drives a large increase in inequality as a consequence of higher wages on the right side of the wage distribution (recall that exporters pay higher wages due to the upward-sloping labor supply).

The results for θ^L and θ^M in Figure 4 stand in sharp contrast to the finding of Helpman, Itskhoki, and Redding (2010) of an inverted-U relationship between inequality and trade costs. In our setting, the inverted-U relationship only emerges if the labor-supply elasticity is high (see the θ^H line in Figure 4). To understand these results, it is useful to remember from section 5.2 that *(i)* trade liberalization increases $\hat{\varphi}_N$ and reduces $\hat{\varphi}_T$, and *(ii)* that a lower θ (more monopsony power) reduces the impact of trade liberalization on $\hat{\varphi}_N$ but magnifies it on $\hat{\varphi}_T$. While *(i)* is present in Helpman, Itskhoki, and Redding (2010), they do not consider monopsony power and thus abstain from *(ii)*. Notice that *(i)* implies counteracting effects of trade liberalization on inequality: while the increase in $\hat{\varphi}_N$ would reduce inequality by compressing the distribution of wages, the decline in
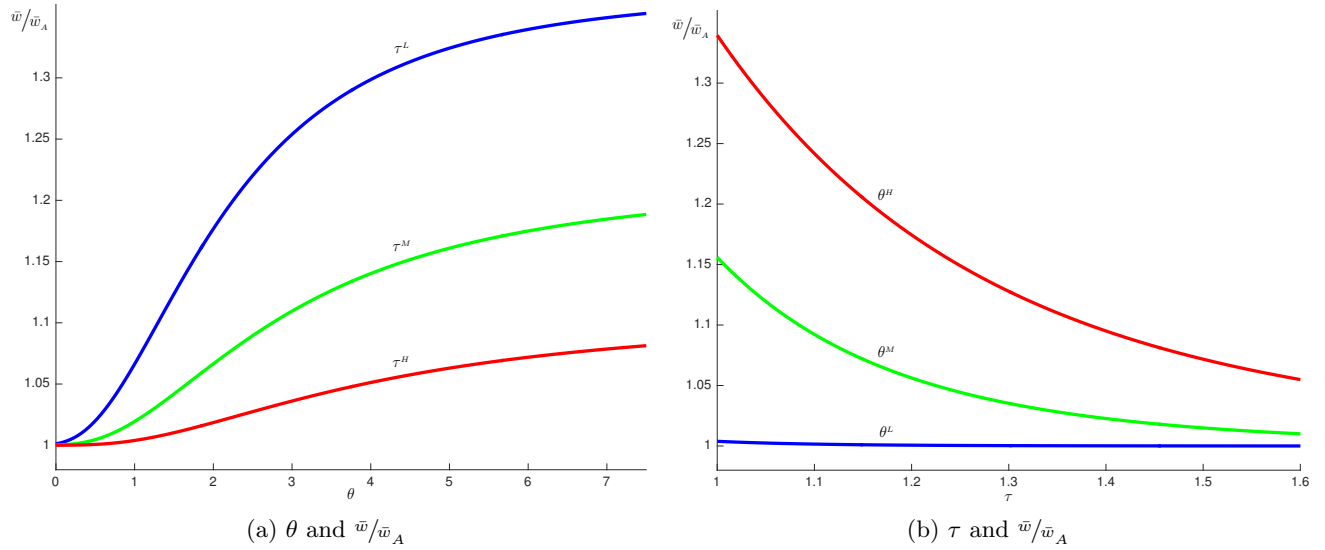
33

Figure 5: Gains from trade

$\hat{\varphi}_T$ would tend to increase inequality because exporting firms pay higher wages than purely domestic firms. In Helpman, Itskhoki, and Redding (2010) or when θ is high, the inequality-reducing effect of an increase in $\hat{\varphi}_N$ dominates for low levels of τ, causing net reductions in inequality after trade liberalization. However, from *(ii)* we know that an increase in monopsony power reduces the impact of trade liberalization on $\hat{\varphi}_N$, and thus dampens its associated inequality-reducing force. With low and moderate levels of θ, the inequality-reducing force is so weak that trade liberalization always increases inequality.

5.5   Welfare in the Open Economy

Similar to the closed economy case, we first look at the average wage as a measure of welfare. In Figure 5 we plot the ratio of the average wage in the open economy to the average wage in autarky, $\bar{w}/\bar{w}_A$, as a measure of gains from trade. It is apparent from Figure 5 that trade liberalization provides gains from trade for all levels of monopsony power and trade costs. Note, however, that the gains from trade are lower the greater the monopsony power (*i.e.* the lower θ is). For example, in Figure 5a when the iceberg trade cost declines from τ^H to τ^L the gains from trade captured by $\bar{w}/\bar{w}_A$ are larger for higher values of θ. Similarly, Figure 5b shows that for each level of τ, the ratio $\bar{w}/\bar{w}_A$ is higher the larger θ is.

Intuitively, an increase in monopsony power reduces the gains from trade through two channels. First, as shown before (see Figure 2d), an increase in monopsony power leads to a decline in the share of exporters. Since exporters pay a higher wage than non-exporters, this is a source of
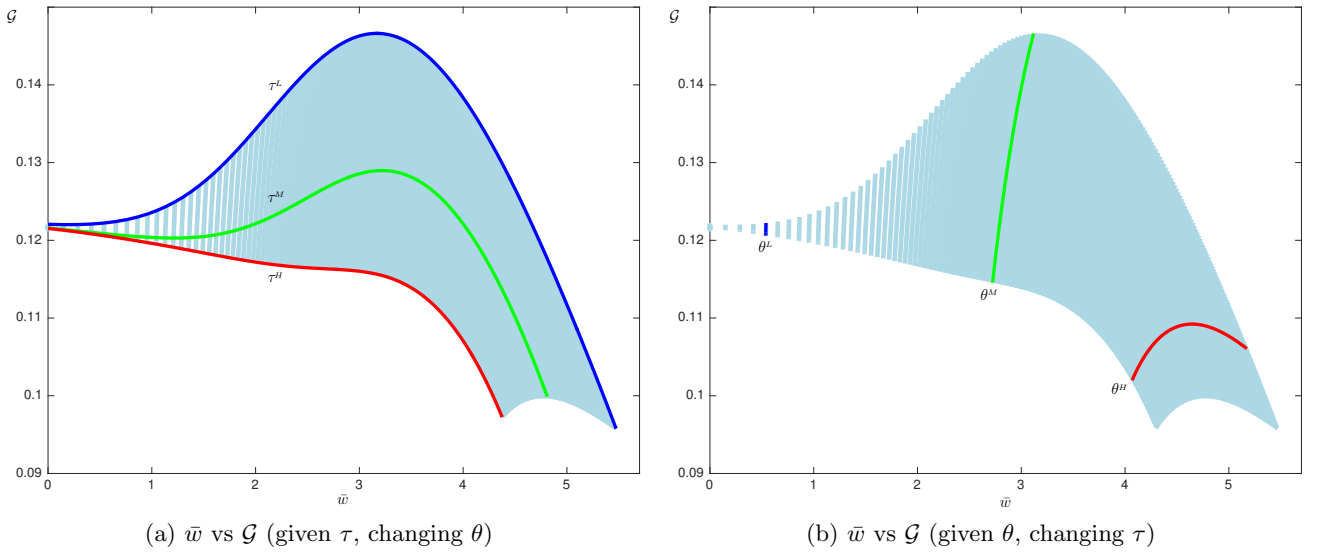
Figure 6: Average wage and the Gini coefficient

(a) $\bar{w}$ vs $\mathcal{G}$ (given τ, changing θ) (b) $\bar{w}$ vs $\mathcal{G}$ (given θ, changing τ)

decline in the average wage. Second, as mentioned in the closed-economy case, more monopsony power increases the misallocation of resources across heterogeneous firms; as a consequence, the Melitz-type market-share reallocations toward more productive firms after trade liberalization are dampened.

As discussed in section 4.7.2, a social welfare function exhibiting inequality aversion should increase in the average wage but decrease in inequality. Figures 3 and 4 capture the impact of θ and τ on inequality, while Figure 5 shows their impact on the average wage. To learn about possible trade-offs between inequality and the average wage, Figure 6 describes the relationship between the Gini coefficient and the average wage for different levels of θ and τ. In Figure 6a we draw solid lines for three different levels of τ to show how the pair $(\bar{w}, \mathcal{G})$ varies as we change θ. For example, θ is declining as we move from right to left along the τ^L line, which initially causes $\bar{w}$ to decrease and $\mathcal{G}$ to increase, leading to an unambiguous reduction in welfare; beyond a point, however, a trade-off arises as further declines in θ continue to reduce the average wage but also reduce inequality (as shown in Figure 3).

Figure 6b draws solid lines for different levels of θ and shows the responses of $(\bar{w}, \mathcal{G})$ as τ changes (τ is declining as we move from left to right in each solid line). In that case, the θ^M line indicates that for a moderate degree of monopsony power, a reduction in τ causes a slight increase in the average wage but a sharp increase in inequality (as was seen in Figure 4). On the other hand, the θ^H line in Figure 6b shows that as trade is liberalized the average wage increases, but inequality increases first before starting to decrease.

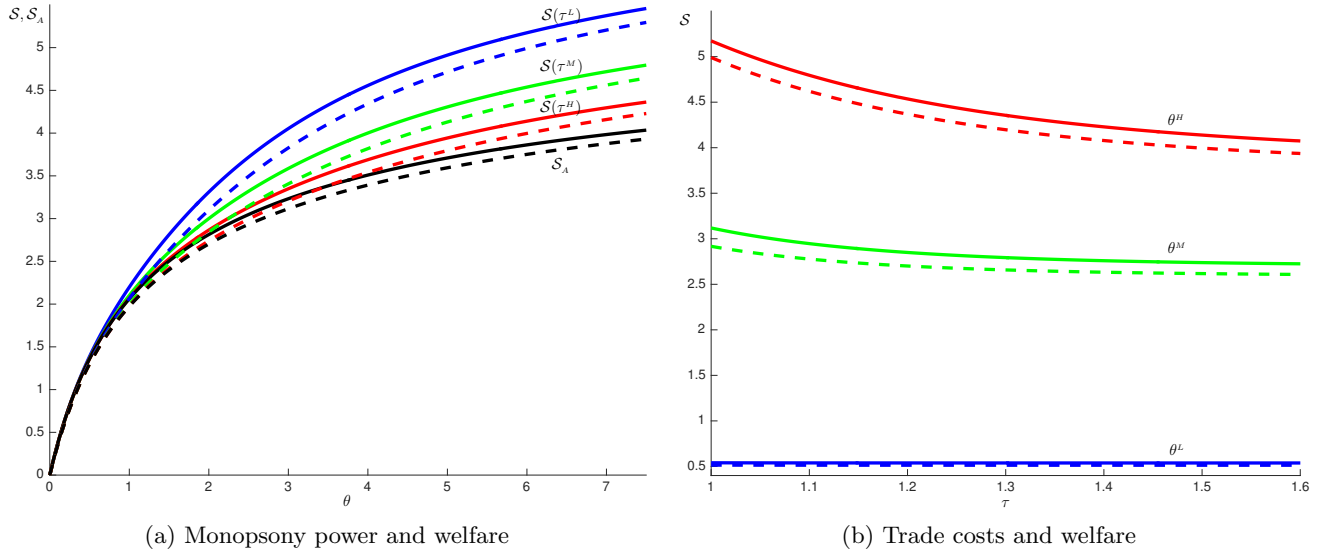(a) Monopsony power and welfare   (b) Trade costs and welfare

Figure 7: Social welfare: $\rho = 0$ (solid lines), $\rho = 2$ (dashed lines)

In parameter zones where the average wage and inequality are both increasing, the impact of monopsony power and trade liberalization on social welfare is ambiguous. To resolve the ambiguity, we use the same inequality-averse social welfare function that we used in the closed-economy case. From equation (19), we obtain that social welfare in the open-economy case is given by

$$\mathcal{S} = \left[\frac{\gamma}{\rho + \gamma - 1}\right]^{\frac{1}{1-\rho}} \left\{\alpha_N \left[\frac{1}{1 - \lambda^{-\beta\gamma/(1+\theta)}}\right]\left[1 - \frac{1}{\lambda^{\beta(\rho+\gamma-1)/(1+\theta)}}\right]\hat{w}_N^{1-\rho} + \alpha_T \hat{w}_T^{1-\rho}\right\}^{\frac{1}{1-\rho}}. \quad (31)$$

Using two levels of inequality aversion ($\rho = 0$ and $\rho = 2$), Figure 7 presents the relationship between social welfare and monopsony power (for different levels of τ), and the relationship between social welfare and trade costs (for different levels of θ). As in the closed-economy case, social welfare declines with ρ and collapses to the average wage when $\rho = 0$.

Figure 7 shows that in both cases ($\rho = 0$ and $\rho = 2$), social welfare declines with monopsony power and increases with trade liberalization. Figure 7a shows that as we approach the case of absolute monopsony power ($\theta \to 0$), social welfare approaches zero. Figure 7b indicates that the effects of trade liberalization on social welfare are smaller the higher the level of monopsony power, becoming negligible for the θ^L case.

In spite of our previous finding that trade liberalization monotonically increases inequality for sufficiently low levels of θ, Figure 7a shows that social welfare is always higher in the open economy than under autarky. To clearly see the gains from trade in terms of our inequality-averse social welfare measure, Figure 8 presents the ratio of social welfare in the open economy to social welfare in autarky, $\mathcal{S}/\mathcal{S}_A$, for various levels of θ and τ. The results in Figure 5 are reproduced in Figure 8

(a) Monopsony power and welfare ratio   (b) Trade costs and welfare ratio
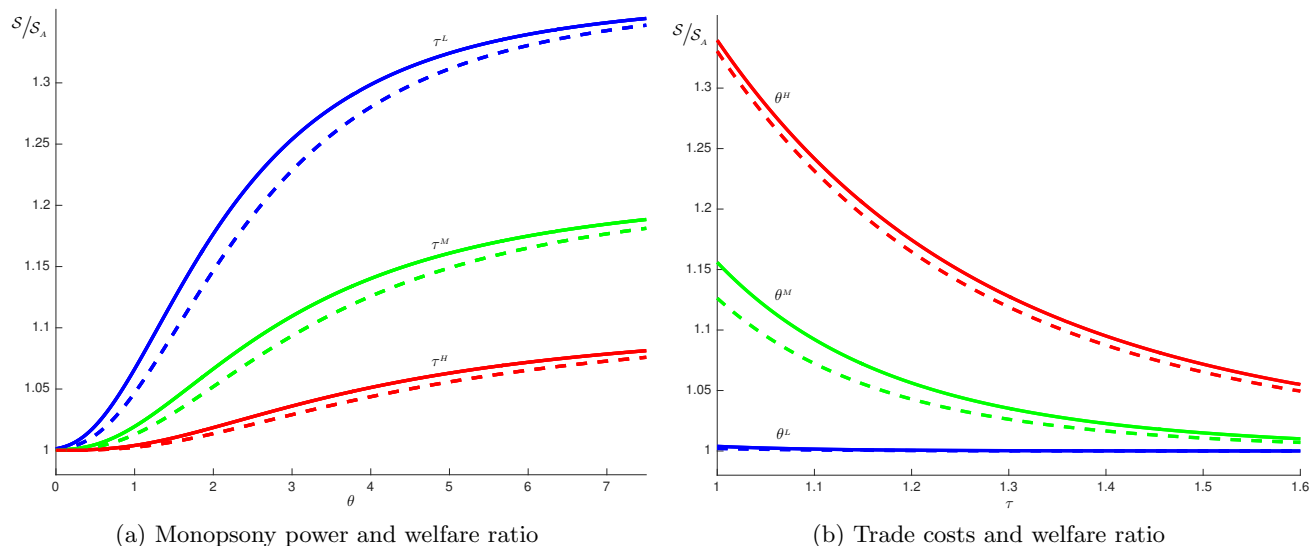
Figure 8: Social welfare ratio to autarky: $\rho = 0$ (solid lines), $\rho = 2$ (dashed lines)

and are represented by the solid lines (if $\rho = 0$, then $\mathcal{S}/\mathcal{S}_A = \bar{w}/\bar{w}_A$). The dashed lines, which capture the results for the inequality-aversion case ($\rho = 2$), show a similar pattern as the solid lines, with increases in monopsony power reducing the gains from trade. In addition to the effects discussed in the context of Figure 5, now gains from trade are further reduced whenever trade liberalization increases inequality.[21] Our welfare results are summarized as follows.

Key Result 2. *(Welfare in the open economy)*

The average wage is higher in the open economy than in the closed economy, and trade liberalization increases the average wage. Social welfare increases with trade liberalization—even in the range of parameter values where trade liberalization increases inequality—and declines with monopsony power. In addition, monopsony power dampens the gains from trade.

6 Concluding Remarks

We constructed a theoretical model to study the implications of globalization on inequality and welfare in the presence of a monopsonistic labor market. The model allows for rich interactions between the degree of monopsony power and the level of trade barriers in determining the impact of trade liberalization on inequality and welfare. Monopsony power dampens the effects of firm heterogeneity on the allocation of market shares by leveling the competition field for low-productivity firms, which pay lower wages than more productive firms. In a closed economy, this effect translates

---

[21]Although our inequality-averse social welfare function always yields a welfare increase after trade liberalization (even for high levels of ρ), it is also conceivable that another more inequality-averse social welfare function could exhibit losses in welfare from trade liberalization.

into a wider distribution of wages and hence more inequality. In an open economy, however, an increase in monopsony power may reduce inequality as the number of exporters decline, driving a compression in the right side of the wage distribution. However, we find that inequality in the open economy is always higher than in the closed economy. In spite of this, social welfare is always higher in the open economy due to the positive impact of trade liberalization on the average wage.

Our framework provides a simple yet powerful structure to study the effects of monopsony power in the labor market on trade-related outcomes. It assumes a constant elasticity of labor supply to provide clear intuition on the first-order effects of labor market power. However, our model can be easily enriched to incorporate endogenous markdowns for wages, with each firm facing a different elasticity of labor supply. In such a setting, a minimum wage can serve as a "choke" wage that would generate markdown heterogeneity in the labor market in a similar way that a choke price causes variable markups in the product market (as in, for example, Melitz and Ottaviano, 2008), with low-productivity firms facing higher labor-supply elasticities and thus setting lower markdowns. We leave this venue of research for the future.

References

AMITI, M., AND D. R. DAVIS (2011): "Trade, Firms, and Wages: Theory and Evidence," *Review of Economic Studies*, 79(1), 1–36.

ANTRÀS, P., A. DE GORTARI, AND O. ITSKHOKI (2017): "Globalization, Inequality and Welfare," *Journal of International Economics*, 108, 387–412.

AUTOR, D., D. DORN, L. F. KATZ, C. PATTERSON, AND J. VAN REENEN (2017): "The Fall of the Labor Share and the Rise of Superstar Firms," Working Paper 23396, National Bureau of Economic Research.

BARTH, E., AND H. DALE-OLSEN (2009): "Monopsonistic Discrimination, Worker Turnover, and the Gender Wage Gap," *Labour Economics*, 16(5), 589–597.

BAUMGARTEN, D. (2013): "Exporters and the Rise in Wage Inequality: Evidence from German Linked Employer-Employee Data," *Journal of International Economics*, 90(1), 201 – 217.

BAYARD, K., AND K. R. TROSKE (1999): "Examining the Employer-Size Wage Premium in the Manufacturing, Retail Trade, and Service Industries Using Employer-Employee Matched Data," *American Economic Review: Papers & Proceedings*, 89(2), 99–103.

BEN-AKIVA, M., N. LITINAS, AND K. TSUNOKAWA (1985): "Continuous spatial choice: The continuous logit model and distributions of trips and urban densities," *Transportation Research Part A: General*, 19(2), 119 – 154.

BERGER, D. W., K. F. HERKENHOFF, AND S. MONGEY (2019): "Labor Market Power," Working Paper 25719, National Bureau of Economic Research.

BERNARD, A. B., AND J. B. JENSEN (1995): "Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987," *Brookings Papers on Economic Activity*, 1995, 67–119.

BERNARD, A. B., AND J. B. JENSEN (1999): "Exceptional Exporter Performance: Cause, Effect, or Both?," *Journal of International Economics*, 47(1), 1–25.

BERNARD, A. B., J. B. JENSEN, S. J. REDDING, AND P. K. SCHOTT (2007): "Firms in International Trade," *Journal of Economic Perspectives*, 21(3), 105–130.

BRAMBILLA, I., N. DEPETRIS CHAUVIN, AND G. PORTO (2017): "Examining the Export Wage Premium in Developing Countries," *Review of International Economics*, 25(3), 447–475.

BROWN, C., AND J. MEDOFF (1989): "The Employer Size-Wage Effect," *Journal of Political Economy*, 97(5), 1027–1059.

BURDETT, K., AND D. T. MORTENSEN (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39(2), 257–273.

CARD, D., A. R. CARDOSO, J. HEINING, AND P. KLINE (2018): "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 36(S1), S13–S70.

CHANEY, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98(4), 1707–21.

COȘAR, A. K., N. GUNER, AND J. TYBOUT (2016): "Firm Dynamics, Job Turnover, and Wage Distributions in an Open Economy," *American Economic Review*, 106(3), 625–63.

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

EGGER, H., AND U. KREICKEMEIER (2009): "Firm Heterogeneity and the Labor Market Effects of Trade Liberalization," *International Economic Review*, 50(1), 187–216.

——— (2012): "Fairness, Trade, and Inequality," *Journal of International Economics*, 86(2), 184–196.

FAJGELBAUM, P. D. (2016): "Labor Market Frictions, Firm Growth, and International Trade," Working Paper 19492, National Bureau of Economic Research.

FALCH, T. (2010): "The Elasticity of Labor Supply at the Establishment Level," *Journal of Labor Economics*, 28(2), 237–266.

FELBERMAYR, G., G. IMPULLITTI, AND J. PRAT (2018): "Firm Dynamics and Residual Inequality in Open Economies," *Journal of the European Economic Association*, 16(5), 1476–1539.

FRÍAS, J. A., D. S. KAPLAN, E. A. VERHOOGEN, AND D. ALFARO-SERRANO (2018): "Exports and Wage Premia: Evidence from Mexican Employer-Employee Data," Discussion paper, Columbia University.

GHIRONI, F., AND M. J. MELITZ (2005): "International Trade and Macroeconomic Dynamics with Heterogeneous Firms," *Quarterly Journal of Economics*, 120(3), 865–915.

GIBSON, J., AND S. STILLMAN (2009): "Why Do Big Firms Pay Higher Wages? Evidence from an International Database," *The Review of Economics and Statistics*, 91(1), 213–218.

GREEN, F., S. MACHIN, AND A. MANNING (1996): "The Employer Size-Wage Effect: Can Dynamic Monopsony Provide an Explanation?," *Oxford Economic Papers*, 48(3), 433–455.

HELPMAN, E., O. ITSKHOKI, AND S. REDDING (2010): "Inequality and Unemployment in a Global Economy," *Econometrica*, 78(4), 1239–1283.

HIRSCH, B., E. J. JAHN, AND C. SCHNABEL (2018): "Do Employers Have More Monopsony Power in Slack Labor Markets?," *ILR Review*, 71(3), 676–704.

HIRSCH, B., T. SCHANK, AND C. SCHNABEL (2010): "Differences in Labor Supply to Monopsonistic Firms and the Gender Pay Gap: An Empirical Analysis Using Linked Employer-Employee Data from Germany," *Journal of Labor Economics*, 28(2), 291–330.

KRUEGER, A., AND E. POSNER (2018): "A Proposal for Protecting Low-Income Workers from Monopsony and Collusion," *The Hamilton Project Policy Proposal*, 5.

MACIS, M., AND F. SCHIVARDI (2016): "Exports and Wages: Rent Sharing, Workforce Composition, or Returns to Skills?," *Journal of Labor Economics*, 34(4), 945–978.

MANNING, A. (2003): *Monopsony in Motion: Imperfect Competition in Labor Markets.* Princeton University Press.

——— (2011): "Imperfect Competition in the Labor Market," in *Handbook of Labor Economics*, ed. by D. Card, and O. Ashenfelter, vol. 4, chap. 11, pp. 973–1041. Elsevier.

MCFADDEN, D. L. (1973): "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, ed. by P. Zarembka, pp. 105–142. Academic Press: New York.

MELITZ, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.

MELITZ, M. J., AND G. I. P. OTTAVIANO (2008): "Market Size, Trade, and Productivity," *Review of Economic Studies*, 75(1), 295–316.

NAIDU, S., E. A. POSNER, AND E. G. WEYL (2018): "Antitrust Remedies for Labor Market Power," *Harvard Law Review*, forthcoming.

OI, W. Y., AND T. L. IDSON (1999): "Firm Size and Wages," in *Handbook of Labor Economics*, ed. by D. Card, and O. Ashenfelter, vol. 3, chap. 33, pp. 2165 – 2214. Elsevier.

RANSOM, M., AND R. OAXACA (2010): "New Market Power Models and Sex Differences in Pay," *Journal of Labor Economics*, 28(2), 267–289.

ROBINSON, J. (1933): "The Theory of Imperfect Competition," *Quarterly Journal of Economics*.

SCHANK, T., C. SCHNABEL, AND J. WAGNER (2007): "Do Exporters Really Pay Higher Wages? First Evidence from German Linked Employer–Employee Data," *Journal of International Economics*, 72(1), 52–74.

STAIGER, D., J. SPETZ, AND C. PHIBBS (2010): "Is There Monopsony in the Labor Market? Evidence from a Natural Experiment," *Journal of Labor Economics*, 28(2), 211–236.

STARR, E., J. J. PRESCOTT, AND N. D. BISHARA (2019): "Noncompetes in the US Labor Force," Working Paper 18-013, University of Michigan Law & Econ Research Paper.

TROSKE, K. R. (1999): "Evidence on the Employer Size-Wage Premium from Worker-Establishment Matched Data," *Review of Economics and Statistics*, 81(1), 15–26.

WEBBER, D. A. (2015): "Firm Market Power and the Earnings Distribution," *Labour Economics*, 35, 123–134.

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