Arndt-Corden Department of Economics
Crawford School of Public Policy
ANU College of Asia and the Pacific

The dynamics of labor share decline in manufacturing: Evidence from Indonesia

Riandy Laksono

Arianto A. Patunru*

Arndt-Corden Department of Economics

Crawford School of Public Policy, Australian National University

Email: arianto.patunru@anu.edu.au

Riandi.laksono@anu.edu.au

* Corresponding author

May 2024

Working Papers in Trade and Development

No. 2024/03

The dynamics of labor share decline in manufacturing: Evidence from Indonesia

Riandy Laksono and Arianto A. Patunru

Abstract

Labour share of income in developing economies has generally declined with increased engagement in international trade, raising concern about adverse distributional consequences of trade for workers. Using a panel dataset of Indonesian manufacturing firms from 1990 to 2015, we evaluate how trade affects the dynamics of labor share from a micro-level perspective. We find that trade liberalization contributes to declining labor share, mainly by shifting market share towards better-performing firms with already-low labor share. While this is in line with the superstar firm framework, such model fails to characterize the labor share dynamics in a developing economy like Indonesia where aggregate markups and concentration do not rise. Instead, this study supports a trade-based explanation for labor share decline.

JEL codes: F61, F63, F66, J30

Keywords: labor share, trade liberalization, superstar firms, Indonesia, manufacturing sector

1. Introduction

Developing economies have kept their trade policy relatively open and free in recent times. However, the labor share of income has shown a declining trend in many parts of the developing countries, (Ahsan and Mitra, 2014, Xu et al., 2018, Dao et al., 2019, Kamal et al., 2019, Leblebicioğlu and Weinberger, 2021). The declining labor share (DLS) trend has raised concerns that trade may worsen distributional outcomes for workers in developing economies.

We evaluate how trade liberalization affects the dynamics of labor share in a developing economy using the framework provided by a class of heterogeneous firm models (Melitz, 2003, Melitz and Ottaviano, 2008, Autor et al., 2020). In particular, the superstar firm model (Autor et al., 2020) has demonstrated that increasing competition could lead to two simultaneous forces: it could change the dynamics of labor share within firms, and it could induce a compositional shift in an industry. Better-performing firms that tend to have lower labor share expand while less productive firms with higher labor share shrink amidst a competitive market environment, depressing the labor share of income at the aggregate, through the reallocation of market share among the surviving firms. Meanwhile, a few other firms could exit the market altogether, while some new ones could enter the market in the background. The superstar firms model shows that the reallocation channel is the dominant driver of DLS in the case of the United States and European economies. However, the model does not offer insight into the source of shock that triggers such reallocation. Using newly constructed panel data of Indonesia's manufacturing firms from 1990 to 2015, we decompose aggregate labor share change into its various micro-level drivers and show that trade liberalization induces the rise of superstar firms, contributing to the DLS. Nevertheless, this study uncovers a different version of the superstar firm model where it does not result in increasing aggregate markup and concentration in the industry.

This paper has three main contributions. First, it helps explain why trade may worsen the distributional outcome for workers in the context of developing economies (Pavcnik, 2017). It has been found that greater trade liberalization and import competition can significantly impact income inequality (Pavcnik, 2017, Goldberg and Pavcnik, 2007, Vadila and Resosudarmo, 2020), poverty (Topalova, 2010, Kis-Katos and Sparrow, 2015), and employment (Autor et al., 2013, Dix-Carneiro and Kovak, 2019). Our findings add to this literature by showing that trade liberalization could also have a crucial effect on the distribution of income across various factors of production. This confirms the view that trade liberalization and openness might be beneficial in general, but the gain may not be shared proportionally across factors or society in general. In this paper we document a labor share-reducing effect of trade liberalization works mainly through reallocation of market share towards low labor share firms in the industry. To the best of our knowledge, this is the first evidence from a developing economy of the importance of intra-industry reallocation channel in explaining the link between trade liberalization and the DLS. Previous studies in developing countries have relied heavily on firm-level evidence, thus capturing only one particular side of the DLS story: the within-firm channel (Ahsan and Mitra, 2014, Kamal et al., 2019, Leblebicioğlu and Weinberger, 2021). Our study also exploits the composition of firms within an industry and discovers that the shift in the firms' composition within an industry constitutes an important channel that explains the role of trade liberalization in the DLS trend.

Second, it shows that in the case of a developing economy, superstar firms may arise following trade liberalization, but the aggregate markup and concentration do not necessarily increase. In a developed economy like Finland or France, greater export participation can trigger reallocation mechanism a-la superstar firms model that leads to the decline in aggregate labor share (Böckerman and Maliranta, 2012; Panon, 2022). In contrast, our study shows that despite that market share reallocates into more productive firms, the reduction of markup within firms is more substantial that it exceeds the rate of reallocation, hence creating downward pressure for aggregate markup in net terms.

Third, from policy perspective, this study contributes to the debate on the impact of trade on workers. Trade protection has often been devised to protect workers' welfare. We argue, however, that trade protection is less likely to work in workers' interests: firms with lower labor costs and higher productivity perform better following trade liberalization.. This implies that less productive firms with higher labor share will be more likely to survive under a more protective environment, thus lowering the aggregate industry's productivity. As a result, the aggregate industry's productivity will be lower in a protective trade environment. Therefore, the use of trade protection will potentially come at the expense of industries' and firms' productivity. Less productive industries and firms are unlikely to align with workers' interests as high-paying jobs tend to come from more productive industries and firms (Bernard et al., 2007).

In this study we first aggregate the manufacturing firm data into a narrowly-defined industry-level data and then break down the source of aggregate labor share change into four components: (i) firm's specific changes (within-firm channel), (ii) change in the compositional structure of firms within industry (reallocation channel), (iii) firms entry, and (iv) firms exit. Some stylized facts emerge. Aggregate labor share in Indonesia's manufacturing sector has fallen substantially from around 13 percent in 1990 to 6 percent in 2015. This decline does not always come from the firm-specific reduction in labor share (within-firm channel). Rather, between 1990 and 2000, the DLS in the manufacturing sector mainly came from the changing firms' composition in the industry (reallocation channel). The market share was increasingly shifting towards low labor share firms during that period, thus pushing down the aggregate labor share in the industry. Meanwhile, the contribution of firms' entry and exit to the overall change in labor share is relatively small.

Next, we analyze whether industry-level aggregate labor share changes are affected by trade liberalization policy, as proxied by reduction in the effective rate of protection (ERP). We find that trade liberalization contributes to the DLS in Indonesia's manufacturing sector. However, the impact of trade liberalization varies across the different micro-level channels. Trade liberalization does not reduce labor share within firms. Rather, it induces reallocation of market share towards low labor share firms in the industry. This intra-industry reallocation effect outweighs that of the within-firm channel and firms' entry and exit, thus pushing the aggregate labor share downward. Therefore, accounting for all channels, the total effect of trade liberalization is labor share-reducing. This result is robust to alternative specifications, models, and choice of tariff measures.

We further investigate why trade liberalization induces reallocation into low labor share firms, by assessing different dimensions of firms' heterogeneity. More specifically, we test the relationship between firms' labor share and other indicators, such as productivity, markup, and capital intensity. We confirm that low labor share firms tend to be more productive and have higher capital intensity and higher markups. Equipped with higher productivity, the low labor share firms have a better chance of surviving and expanding under a liberalized trade environment. This result suggests that the trade-induced DLS in Indonesia's manufacturing sector is inevitable outcome as trade selects the better-performing, more productive firms (Melitz, 2003, Melitz and Ottaviano, 2008). Although this reaffirms the superstar mechanism as in Autor et al. (2020), where the better performing-low labor share firms rise due to increasing market competitiveness, we also find that aggregate markup does not rise in Indonesia's manufacturing sector due to substantial decrease in average markup across individual firms. Furthermore, aggregate concentration tends to be steady as well. Thus, the superstar firms model is not a perfect characterization of labor share dynamics in Indonesia's manufacturing sector. Rather, this paper shows that the role of trade liberalization could not be ignored in explaining the DLS in a developing economy like Indonesia even after controlling for the changes in markup, concentration, and productivity at the industry level.

The following section discusses the conceptual framework. Section 3 describes data and variables used, while empirical strategy will be outlined in Section 4. The result will be presented and discussed in Section 5, and Section 6 concludes.

2. Conceptual Framework

The superstar firms model (Autor et al., 2020) predicts that following increased competition, aggregate labor share would decline as market shares are reallocated disproportionally towards the better-performing, 'superstar' firms. These firms tend to be large, have lower labor share, higher productivity, and charge greater markups. Thus, when the industry's gravity moves towards superstar firms, the aggregate labor share falls as a consequence. Although the model acknowledges that the change in market competitiveness can also influence aggregate labor share via other channels, such as within-firm movement, firms' exit, and entry, the contribution of the reallocation term is predicted to dominate any other channels. Not only does labor share decline, but aggregate markup and concentration are also expected to increase due to the reallocation toward a few dominant superstar firms. However, the model does not offer insight into sources of shock leading to the rise of superstar firms in the industry.

The superstar firm model shares a similar heterogeneous firms assumption with Melitz's trade model (Melitz, 2003, Melitz and Ottaviano, 2008), but the latter focuses on the role of trade and does not evaluate the dynamics of labor share. Both models share similar prediction about the centrality of the intra-industry reallocation effect arising from any shock that increases market competitiveness. As competitiveness rises, aggregate performance in the industry changes. First, it could change the dynamics of labor share within firms. The rising competitiveness in the market could force firms to adjust their labor cost share. The result will depend on how firms respond to trade liberalization. Firms may reduce labor share to improve efficiency. However, in a labor-abundant economy, firms' labor share may increase as in the standard Heckscher-Ohlin-Samuelson model. The share may also increase due to a larger markup which is eventually passed on to workers (Kamal et al., 2019; Leblebicioğlu and Weinberger, 2021).

Trade liberalization also generate a compositional shift within industry. High-performing firms expand while the mediocre ones stay but shrink, forcing the reallocation of market shares across the surviving firms. The least efficient firms exit the market altogether, while new firms can enter. This compositional change influences the change in the aggregate labor share in the industry.

Various studies have shown that reallocation improves the aggregate productivity in the industry. However, they differ on how it would affect aggregate markups. By treating markup endogenously, Melitz and Ottaviano (2008) argue that the pro-competitive gain of trade within individual firms is so substantial that it exceeds the increasing markup originating from reallocation of market share toward high-markup, better-performing firms. In contrast, the superstar firms model developed by Autor et al. (2020) predicts a stronger reallocation effect than the within-firm one, thus giving rise to the aggregate markups within an industry. The increase in the aggregate markup will also be accompanied by rising industry concentration in the model, indicating larger market power of a few dominant firms in the industry.

The total effect of trade on aggregate labor share is not clear a priori. It will depend on which channel dominates: (i) firm-specific change (within-firm channel), (ii) the reallocation of market share across firms (reallocation channel), or (iii) firms' turnover (exit- and entry channels). Therefore, the net effect of trade liberalization on labor share is an empirical matter that should be assessed in totality across various channels.

3. Data and Variables Construction

3.1. Dataset sources and construction

We use four main datasets: panel data for manufacturing firms, commodity-level inputs, price deflators, and most-favored-nation (MFN) tariffs. The manufacturing firm panel dataset covers the period of 1990 to 2015, which is constructed from the annual manufacturing survey of medium and large establishments (*Statistik Industri,* SI). The survey is conducted by Statistics Indonesia (*Badan Pusat Statistik*, BPS) which targets all firms in the formal sector with 20 or more workers that are listed in the BPS's manufacturing industry directory. The SI dataset records detailed plant-level information on production, export, input, ownership, location, and sectoral affiliation. Although it is a plant-level database, most of the plants in the dataset belong to a single-plant firm (Putra, 2021).

The commodity-level inputs dataset is compiled by BPS alongside the plant-level data and will be used to appraise the input structure of a given industry in the manufacturing sector. This becomes the basis for estimating the effective rate of protection (ERP) in a particular industry when combined with tariff information.

In the deflator dataset the wholesale price index (WPI) is used to deflate nominal variables in the SI, such as value-added, raw material expenses, fuel, electricity, and other expenses as well as capital stock. The WPI for production value and material expenses from 1990 to 2015 used in this paper is the extended version of the published WPI data on the BPS website.[1] In addition, a dedicated WPI for capital goods is used to deflate the capital stock. The WPI for capital goods is downloaded from the CEIC database. All of the WPI data use the year 2000 as the base year.[2]

The most-favored-nation (MFN) tariff dataset is retrieved from the World Bank WITS (World Integrated Trade Solution) database. This paper uses both ad-valorem and specific rates at the tariff line level. The specific rates are further transformed into their ad-valorem equivalence following the

[1] We would like to thank [redacted for review process] for sharing the plant-level and input-level SI datasets as well as the more detailed WPI data, respectively. The sectoral weight from BPS is used to extend the aggregate WPI data. The growth of WPI is preserved when changing the base year.

[2] The original data from the CEIC database is stored in various base years depending on the period of observation. It is then rebased to the year 2000 by using the original growth of the capital goods WPI as the basis for extrapolation. Since there is a disconnected series from 1998 to 2001 in the capital goods WPI data, we use the growth rate of gross capital formation (GCF) deflator for the period of 1999 to 2000 as a basis for extrapolation from 1999 to 2000. Furthermore, WPI for capital goods for the period 1990 to 1992 is not available in the CEIC, so the growth rate for 1990 to 1993 are borrowed from the overall WPI growth rate as provided by an extension performed by Professor [redacted].

UNCTAD methodology that is available in the WITS database.[3] These tariff rates are then aggregated at the 4-digit ISIC level (based on the Revision 2 version) using the simple average formula to rule out any possible endogeneity related to changing trade structure. There is an issue of missing tariff data in WITS. We address this problem using linear interpolation, considering that tariffs tend to show a linear downward trend over time.[4]

Between 1990 to 2015, the SI dataset uses three different sectoral classifications. They are the *Klasifikasi Baku Lapangan Usaha Indonesia* (KBLI) 2009 for 2010-2015 data, KBLI 2005 for 1999-2009 data[5], and *Klasifikasi Lapangan Usaha Indonesia* (KLUI) 1990 for 1990-1998 data. These three classifications are coded on a 5-digit basis, which is basically an extension of the 4-digit ISIC revision 4, revision 3, and revision 2, respectively. To get a consistent classification, we convert the sectoral codes for 1999-2015 data into KLUI 1990 using an unpublished concordance table from BPS. Furthermore, those industries that are not always part of the manufacturing sector definition over time are removed. To allow for dynamic decomposition to run smoothly, industries with only a few observations are removed or, if possible, lumped together with other similar industries under the same 3-digit ISIC classification. This strategy removes around 0.3 percent of observations in the original dataset. As a result, the final number of observations is 574,988, spread into 78 industries in an unbalanced panel setting (see Table A1 in Appendix A for the list of industries).[6]

3.2. Data treatment

Given the raw nature of the SI database, some issues emerge. First, a few firms do not have a full 5-digit KBLI or KLUI code. In treating this issue, we follow the strategy similar to that of Putra (2021), where the incomplete digit is replaced by the historical mode value of the 5-digit sectoral code for that particular firm, but only when the rounded-down digit of the potential replacement value equals the incomplete case.[7]

---

[3] The purpose of including the AVE rate in the average tariff rate is to capture the actual rate of protection afforded to a given industry. This is because the exclusion of specific rates can make the average tariff rate biased downward, thus underestimating the real protection in a particular industry.

[4] The median of interpolated tariff data for the period of 1991, 1992, and 1994 has a comparable trend with the truly observed tariff data of the same period used in Amiti and Konings (2007). Comparison is available upon request.

[5] There are two different KBLIs used during this period, namely KBLI 2000 and KBLI 2005. There are only slight differences between the two, in particular, a minor change in the classification for dairy industry and textile. This has been adjusted accordingly in this analysis.

[6] This number already excludes manufacturing firms located in Timor Leste. BPS included it in the SI dataset during the period when Timor Leste was still part of Indonesia.

[7] For example, the incomplete case is recorded as KBLI code 31. If the mode value is 31121, then the code 31 is replaced by 31121. However, if the mode value is 32111 then the incomplete case is left as it is. This treatment

Second, and more importantly, a large number of firms reported non-positive capital stock data. This includes negative, zero, and missing data. These are well-known measurement errors and non-response issues that have traditionally plagued studies using the SI database (Blalock and Gertler, 2004, Amiti and Konings, 2007, Pane and Patunru, 2021). To overcome this problem, we apply a multiple imputation method (see Appendix B).

The remaining issues are data outliers and missing variables other than capital stock data. We follow Autor et al. (2020) in winsorizing the lowest and highest one percentile of various variables in the SI database, such as value-added, capital stock, raw material, and other expenses. This treatment will also be applied for the constructed variables such as labor share, TFP, and markup. In addition, firm's value-added data often contains negative or missing values. To address this issue, we assign a very low weight to these data points to minimize bias associated with a simple deletion strategy.

3.3. Variables Construction

We construct several firm-level variables for the analysis. The first is the firm's labor share. We follow the standard definition in the literature, which expresses labor share of firm i at time t (S_{it}) as the ratio of nominal labor cost to nominal value-added: $S_{it} = \frac{C_{it}^{L}}{NVA_{it}}$. We use both narrow and broader definitions of labor cost in this paper. The former includes only salary or wage, while the latter also covers other kinds of compensation, such as insurance, pension, and other allowance. The labor share based on wage (wage share) will be the focus of analysis in this paper, and that of compensation (compensation share) will be used as a comparison.

To complement the analysis of labor share, we estimate firm-level productivity and markups. These characteristics matter in determining which firms lose or win from trade openness, thus defining the dynamics in aggregate industry performance. These variables will help explain why trade shock propagates through a particular mechanism. In addition, they will also be used to assess the relevance of competing arguments put forward by the literature in explaining the decline of aggregate labor share. Markups and productivity variables are recovered after estimating the production function for manufacturing sector. To control for the well-known endogeneity bias, we estimate the production function in two stages. There are many variations within this control function approach, particularly those developed by Olley and Pakes (1996) (OP estimator), Levinsohn and Petrin (2003) (LP estimator), and Ackerberg et al. (2015) (ACF estimator). Given the limited availability of proxy variables and stability of the estimator, we opt for the LP method for the main analysis, while ACF

---

assumes continuity of business activity of particular firms but at the same time also avoid shifting business activity dramatically.

estimator is reserved for comparison.[8] The production function is estimated separately for each of the 2-digit ISIC industries to allow for variation in production technology across industries. Total Factor Productivity (TFP) is obtained as the residual from the production function estimation, while markup is recovered using labor input coefficient from production function estimation (De Loecker and Warzynski, 2012, De Loecker et al., 2020, Autor et al., 2020). Appendix C gives details on the adopted methodology in estimating production function, TFP, and markup.

3.4. Aggregation and decomposition

Since this paper focuses on the industry-level analysis to decompose trade's effect into various micro-level components, the firm-level variables must be aggregated. When an industry comprises of firms with different level of labor share, the industry's aggregate labor share at time t (S_t) can then be expressed as the weighted sum of firm i's labor shares within that industry (S_{it}), where the weight is the firm's share of value-added (w_{it}). Equally, it can also be defined as the ratio of labor cost ($\sum_i C_{it}^L$) to the value-added of that industry ($\sum_i NVA_{it}$). Thus:

$$S_t = \frac{\sum_i C_{it}^L}{\sum_i NVA_{it}} = \sum_i \left(\frac{C_{it}^L}{NVA_t}\right) = \sum_i \left(\frac{C_{it}^L}{NVA_{it}} \frac{NVA_{it}}{NVA_t}\right) = \sum_i w_{it} S_{it}. \tag{1}$$

Olley and Pakes (1996) develop a way to decompose aggregate labor share at a particular time t into two different components:

$$S_t = \bar{S}_t + \sum_i (w_{it} - \bar{w}_t)(S_{it} - \bar{S}_t) = \bar{S}_t + cov(w_{it}, S_{it}). \tag{2}$$

The first term on the right-hand side ($\bar{S}_t$) is the unweighted average of firms' labor share at particular time t, $\bar{w}_t$ is the average firms' market share in the industry, while $cov(w_{it}, S_{it})$ is the covariance (joint distribution) of firms' market share and labor share. This last term is also often referred to as the reallocation or Olley-Pakes (OP) term, where it can be used to measure market share reallocation or compositional shift across different types of firms in an industry over time. The lower the covariance (reallocation) term, the more market share is allocated to firms with lower-than-average labor share.

Melitz and Polanec (2015) modified the OP decomposition to account for the contribution of firms' entry into and exit from the industry (hence, Dynamic Olley Pakes Decomposition (DOPD)). It is argued that the DOPD method can correct the over-measurement bias of entrants' contributions and

[8] We use the Stata prodest package in estimating production function coefficients (Rovigatti and Mollisi (2018)

the under-measurement bias of surviving firms (see Melitz and Polanec (2015) for detailed discussion). In any particular industry, firms can be part of one of these three groups (G): surviving firms (A), entrants (E), or exiters (X). For example, suppose there are two periods of $t = 1$ and $t = 2$. Then, surviving firms are those that stay at both $t = 1$ and $t = 2$, exit firms are only active at $t = 1$, while entrants become operational at $t = 2$. Therefore, aggregate labor share in period 1 is $S_1 = \sum_{i\in A,X} w_{it}S_{it}$, while in period 2 it is $S_2 = \sum_{i\in A,E} w_{it}S_{it}$. Defining $w_{Gt} = \sum_{i\in G} w_{it}$ as the aggregate market share of a particular group G of firms in an industry and $S_{Gt} = \sum_{i\in G}(w_{it}/w_{Gt})S_{it}$ as that group's aggregate labor share, we can then rewrite the aggregate labor share in both periods into:

$$S_1 = w_{A1}S_{A1} + w_{X1}S_{X1} = S_{A1} + w_{X1}(S_{X1} - S_{A1}),$$

$$S_2 = w_{A2}S_{A2} + w_{E2}S_{E2} = S_{A2} + w_{E2}(S_{E2} - S_{A2}),$$

where $w_{A1} + w_{X1} = w_{A2} + w_{E2} = 1$. By substituting the survivors' contribution with the components of OP decomposition in equation (2), the change in aggregate labor share from t =1 to t = 2 can now be expressed as:

$$\begin{aligned}\Delta S = S_2 - S_1 &= (S_{A2} - S_{A1}) + w_{E2}(S_{E2} - S_{A2}) + w_{X1}(S_{A1} - S_{X1}) \\ &= \Delta\bar{S}_A + \Delta cov_A + w_{E2}(S_{E2} - S_{A2}) + w_{X1}(S_{A1} - S_{X1}).\end{aligned} \tag{3}$$

The first two terms in the second row of equation (3) are similar to OP decomposition, only that it is now applied to staying firms in the industry between the periods. The third and fourth terms, on the other hand, measure the change in labor share attributed to firms' entry and exit, respectively. The linear decomposition means that the aggregate labor share change on the left-hand side will be equal to the micro-level components on the right-hand side, which comes from firm-level dynamics. The decomposition based on heterogeneous firms' framework implies that the total effect of trade on aggregate labor share is not clear a priori. It will depend on which channel is operational and dominates, as trade liberalization could affect each component differently.

These aggregation and decomposition procedures can also be applied to other indicators, such as markup (M_{jt}) and productivity (Φ_{jt}).[9] To maintain consistency, value-added is used as the weight in the aggregation process of all performance indicators, especially labor share, markups, and productivity. Other aggregate variables, particularly those used as covariates in the main model, are

---

[9] The underlying firm-level indicators are measured in percentage for labor share and in log form for TFP and markup. Thus, the aggregate labor share change is in percentage points; while the aggregate markup ($\Delta\mathrm{M}_{jt}$) and productivity change ($\Delta\Phi_{jt}$) are in percentage.

obtained from firm-level data using simple aggregation to properly reflect the industry-level characteristics.

We can further decompose the change in labor share into changes in wages, employment and value added as well as productivity. Aggregate labor share (S_t) in an industry can be defined as total labor cost (aggregate wages (ω_t) times employment (L_t)) over aggregate value added (NVA_t). Then, the aggregate labor share and the change of it can be expressed as follows:

$$S_t = \frac{\omega_t L_t}{NVA_t} = \frac{(\omega_t/P_t)L_t}{NVA_t/P_t} = \frac{(\omega_t/P_t)}{{(NVA_t/P_t)}/{L_t}}$$

$$ln(S_t) = ln(\omega_t/P_t) + ln(L_t) - ln(NVA_t/P_t) = ln(\omega_t/P_t) - ln\left({(NVA_t/P_t)}/{L_t}\right), \tag{4}$$

Where P_t is industry-level deflator at time *t*, while ω_t/P_t means real aggregate wage, NVA_t/P_t is real value-added, and ${(NVA_t/P_t)}/{L_t}$ represents real labor productivity in an industry.[10]

Table 1 provides the summary statistics of all the variables (see Table A2 in Appendix A for detailed description of each variable).

Table 1. Summary statistics

Variables	N	Mean	SD
Value added (Rp million, constant 2000)	574988	131.80	1657.13
Raw material (total) expenses (Rp million, constant 2000)	574988	166.59	1762.22
Fuel expenses (Rp million, constant 2000)	574988	7.15	155.64
Electricity expenses (Rp million, constant 2000)	574988	6.91	160.13
Auxiliary expenses (Rp million, constant 2000)	574988	27.56	515.38
Total number of workers: all type	574988	192.47	713.99
Number of workers: non-production (skilled)	574988	31.47	157.20
Foreign ownership (%)	574988	6.27	22.69
Export status (1 = exporter, 0 = non-exporter)	574988	0.13	0.34
Capital stock, original and imputed (Rp million, constant 2000)	574988	144.79	2317.41

[10] This decomposition means that the change in aggregate labor share and its components is measured in terms of percentage change. This is different than in the DOPD technique that measures labor share's change in terms of percentage points change.

Labor share: wage or salary relative to value added	574988	0.49	0.51
Labor share: total compensation relative to value added	574988	0.54	0.62
Markups – LP (ratio)	574643	2.49	3.96
Markups – ACF Cobb-Douglas (ratio)	574643	3.10	5.30
Markups – Simple accounting (ratio)	574988	1.49	0.97
TFP – LP (log)	574988	5.45	1.20
TFP – ACF Cobb-Douglas (log)	574988	4.67	1.53
Capacity utilization (%)	550701	64.79	32.51
Effective Rate of Protection (ERP), input weight: average 2000-2015 (%)	2028	32.37	429.28
Nominal Rate of Protection (NRP): output tariff including ad-valorem estimation (%)	2028	18.61	80.58
Input tariff, input expenses to value added weight: average 2000-2015 (%)	2028	4.44	4.20
Change in aggregate labor share (percentage points change)	1950	-0.28	17.08
Percentage change in aggregate labor share (%)	1950	-2.20	54.16
Percentage change in aggregate real wage (%)	1950	3.29	37.68
Percentage change in aggregate employment (%)	1950	3.41	25.30
Percentage change in aggregate real labor productivity (%)	1950	5.49	47.31
Percentage change in aggregate real value-added (%)	1950	8.90	53.21

Note: Capital stock data are summarized following Rubin's rule for pooling multiple imputation data (Enders, 2010).

Source: Authors' estimation.

4. Empirical Framework

4.1. Effective Rate of Protection as Exogenous Shock

Establishing a causal link between trade policy and outcomes of interest has been a challenging avenue in empirical studies. One of the primary reasons is that trade policy is not determined in a vacuum. The government may set most favored nation (MFN) tariffs with pre-determined goals or in response to industry's performance and characteristics. For example, it could deliberately protect the infant industries, the highly organized ones (Grossman and Helpman, 1994), or selectively liberalize the readiest sectors, such as those with higher productivity (Brandt et al., 2017). These issues have made it difficult to interpret the impact of trade liberalization on certain outcome since firms and industries could self-select into certain policy regimes.

Addressing this issue, some studies opt for using the preferential tariffs committed in the free trade agreement (FTA) as a trade policy shock. However, this kind of tariff has its own problem as well. Preferential tariff is typically negotiated under a specific modality that secures a gradual trade liberalization commitment for many years to come. Hence, the modality feature of preferential tariff makes it highly likely to be anticipated by private sectors. This will potentially bias the estimation of the treatment variable of interest.

We propose using the effective rate of protection (ERP) at the industry level as a treatment variable to minimize the endogeneity problem. ERP is a prominent concept in international trade literature for measuring the overall degree of protection afforded to the industry (Corden, 1966, Balassa, 1971). It is distinct from nominal protection, which only covers output tariff in its measurement. ERP, instead, takes into account the overall protective effect of tariffs imposed on the output and input of a particular industry. While output tariff raises the price of a particular imported product, thus increasing the protection in the domestic industry, in contrast, input tariff raises the price of imported raw materials, hence reducing the level of protection of the user industry. The interest of the domestic output-producing industry then contradicts that of the domestic input-producing industry. In this case, it is highly likely that the input-producing industry will counteract any lobbying effort from the output-producing sector to improve the level of protection via lowering the input tariff (Grossman and Helpman, 1994). This counterbalancing act will then limit the influence of the industry's lobby on its own sector and makes it less likely for industries or firms to self-select into certain ERP level.

In addition, ERP is rather unexpected from the perspective of the industry. The government does not deliberately choose a particular ERP level to respond to a lobby, anticipate future events, or achieve any development goals. ERP is also not used to inform policymakers about the next course of action. Instead, it is utilized to advocate for a more uniform tariff structure in the economy (Athukorala, 2006, Greenaway and Milner, 2003). More often than not, the government fails to understand the interindustry implication of imposing a tariff in a particular sector (Balassa, 1971). Thus, the movement of ERP protection could plausibly be unanticipated and unexpected from the point of view of the corresponding industry. It has been argued that unanticipated and unexpected policies tend to work better for identification purposes (Autor et al., 2016, Autor, 2018).

We find support for the exogeneity assumption of ERP. Since industry's influence will likely be confined to its own tariff rate, then the deviation of ERP from the nominal protection (own tariff) can be expected as exogenously determined. Table A3 in Appendix A shows that the deviation of ERP from its nominal protection is not driven by various industry performance indicators, including the industry's initial condition. In addition, the exogeneity assumption for ERP is more plausible than that of output tariff as most industry characteristics do not predict the movement of ERP over time (see Table A4 in Appendix A). Section 5 further shows that there is no anticipatory effect of ERP on industry's outcome, suggesting that it fulfils strict exogeneity assumption.

4.2. Constructing ERP

We follow Corden's procedure for constructing the ERP (Corden, 1966) and only cover the MFN import tariff rates.[11] The ERP rate in a particular period is estimated using the following formula:

$$g_j = \frac{\tau_j - \sum_{r=1}^{n} a_{rj}\tau_r}{1 - \sum_{r=1}^{n} a_{rj}}, \tag{5}$$

where τ_j is the tariff on the main product of industry j, τ_r is the tariff on a set of tradable inputs *r* used in the production of industry j, and a_{rj} is the share of input *r* in the production of industry j, which is defined as $a_{rj} = \frac{M_{rj}}{Y_j}$ where M_{rj} is the value of input *r* in the production of industry j, while Y_j is the the value of production of output j.

We use the commodity-level input SI dataset (SI-input) and the standard firm-level SI dataset in order to compute a_{rj}. The SI-input is structured in a firm-commodity-year dimension, while the standard SI dataset is at firm-year level. Since this paper aims to construct effective protection at sectoral level, namely at 4-digit ISIC level, both datasets need to be aggregated at that level for every time period. First, firm i's input use at commodity level (c), that operate in the production of sector j, is accumulated at sectoral level (r): $M_{irj} = \sum_{c=1}^{n} M_{icrj}$. Then, it is further aggregated so that it shows how intensive output-sector j uses input from sector r: $M_{rj} = \sum_{i=1}^{n} M_{irj}$. Finally, input share is calculated by dividing M_{rj} with sectoral output j which comes from aggregation of firm i output belonging to sector j $\left(Y_j = \sum_{i=1}^{n} Y_{ij}\right)$. It is worth noting that the input and firm-level SI datasets might have a slightly different composition of firms due to incomplete records, mainly in input datasets. To construct the a_{rj}, we use a consistent set of data where firms must exist in both the SI-output and SI-input databases.

[11] Due to data limitation, we do not include non-tariff barriers. As such, our results might underestimate the impact of trade liberalization on aggregate labour share. The World Bank released a new database on non-tariff measures in Indonesia (Montfaucon et al., 2023), but it only spans from 2008, whereas our manufacturing dataset starts from 1990. Another option is to include a dummy variable indicating the presence of NTB at industry level; however, this will be cancelled out by the industry fixed effect that we use in the regression.

The technical coefficient a_{rj} is calculated using the share of both imported and domestic raw materials *r* used in the production of industry *j*. This is because the tariff on input will impact both materials bought from import or domestic suppliers as domestic input producers will also have the option to raise their price in response to the tariff levied. Therefore, regardless of the source of materials, the user industry will suffer from an input tariff (Balassa, 1971). The inclusion of domestic raw materials in the calculation of ERP will thus show a more accurate picture of the overall protection in the industry.

We further opt for a fixed technical coefficient over time to avoid additional endogeneity issues.[12] Although it may seem restrictive, the use of a time-varying technical coefficient in the construction of ERP might reflect changing pattern of material used rather than the change in the policy environment, thus making it difficult to ascertain the causal connection between trade on labor share. We construct a fixed technology coefficient (a_{rj}^0) based on the industry *j*'s average use of input *r* from the period 2000 to 2015: $a_{rj}^0 = 1/n \sum_t^n a_{rj}^t$.[13] This strategy will allow for isolating only the effect of changing tariffs and expunging any contribution of changing technology. The ERP level for every period of *t*, then, can be rewritten as follows:

$$g_{jt} = \frac{\tau_{jt} - \sum_{r=1}^{n} a_{rj}^0 \tau_{rt}}{1 - \sum_{r=1}^{n} a_{rj}^0}. \quad (6)$$

Figure 1 shows that various trade protection measures in the manufacturing sector generally trended downward, except for a few years around the period of the Asian financial crisis. The movement of median ERP over time resembles very closely that of output tariff. The co-movement of industries' output tariff and ERP also occur in other studies, such as in Kamal et al. (2019). Despite the similarity, the two-way *t*-tests in Table A5 (Appendix A) show that they have statistically different values in terms of level or change. This reflects the nature of ERP, which has a larger variation and wider range (see the shaded grey area in Figure 1) than the output tariff (the dashed black line). While the change in output tariff can only come from itself, the change in ERP, on the other hand, can come from either the change in output tariff, input tariff, or both. Therefore, while nominal protection for a given industry may remain constant for several years due to the infrequent nature of trade policy reforms, it does not imply that overall protective rates are also unchanged, as the movement in ERP could come from the change in input tariffs, absent the change in output tariff. In addition, the greater

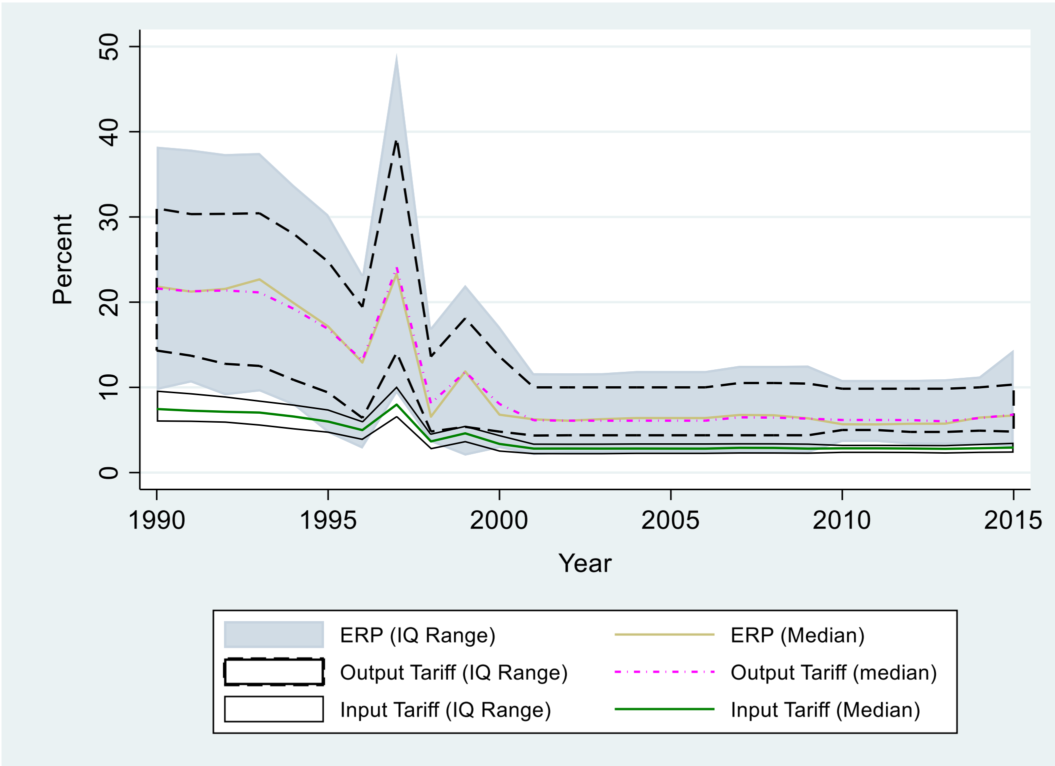
---

[12] The endogeneity problem from using a time-varying production technology is also highlighted by other studies, such as Amiti and Konings (2007).

[13] As a comparison, we will also present ERP estimation using a fixed technology coefficient constructed only from 2006 data. We use the 2006 as a comparison based on the reason that it is the census year, hence increasing the likelihood that the data capture more complete information on industry's input structure.

variation in ERP has the added benefit of reducing the likelihood of finding a statistically significant effect when such effect does not exist, thus making inference more reliable.

Figure 1. ERP, output tariffs, and input tariffs in manufacturing sector



Note: Input tariffs are weighted by the share of input expenses in the production value.

Source: Authors' estimation.

4.3.Main Estimating Equation and Identifying Assumption

To evaluate the effect of trade liberalization on the aggregate labor share as well as on various micro-level components that drive its change, we adopt the empirical framework similar to the one used by Böckerman and Maliranta (2012), Brandt et al. (2017), and Autor et al. (2020):

$$\Delta S_{jt} = \alpha + \beta_1 \Delta ERP_{jt} + \Delta Z'_{jt}\beta_2 + \delta_j + \theta_t + \varepsilon_{jt}. \tag{7}$$

It is a first-difference model where j and t denote 4-digit industry by ISIC revision 2 and year, respectively. The first difference model is preferred to the level model because the former allows for linearly decomposing the effect of ERP on labor share into the four components: within-firm, reallocation, and contribution from firms' entry and exit. The first difference model could also minimize bias due to varying initial level of protection (Vadila and Resosudarmo, 2020). ΔS_{jt} is the dependent variable which consists of aggregate labor share change and its decomposed components as in equation (3). Meanwhile, ΔERP_{jt} is the change of effective rate of protection (ERP) in an industry and ΔZ_{jt} is a set of covariates, measured in change, that are chosen to control for potential confounding factors to the effect of ERP.

Literature informs the choice of covariates, and it shows that there is a multitude of factors that could potentially drive the decline of aggregate labor share other than trade. We address these potential confounders by controlling them in the regression as covariates. Three groups of covariates are used in the main model. First is the covariate that controls any biased technical change that favors factors other than labor. Here, we use real capital stock relative to workers at the aggregate industry level $\left(KINT_{jt} = \sum_i CAP_{ijt} / \sum_i LTL_{ijt}\right)$ to control for the change in labor share that is driven by the change in capital intensity in the industry. Second, factors that reflect skill composition and the demographic profile of firms in the industry. Some industries might differentially reduce their labor share faster than others due to more intensive use of skilled labor. In addition, older firms could be more capital intensive, thus having a lower labor cost component in their value-added (Hopenhayn et al., 2018). Therefore, in the main model, we also control for the share of skilled workers $\left(SKILL_{jt} = \sum_i LNP_{ijt} / \sum_i LTL_{ijt}\right)$ and the average cumulative age of firms in the industry $\left(AAGE_{jt} = 1/n \sum_i^n cum\_age_{ijt}\right)$. The share of skilled workers is defined as the number of non-production workers (LNP_{ijt}) relative to total workers (LTL_{ijt}), while the cumulative age of firms in the industry is constructed assuming the year 1990 or the earliest available year as the starting age for each firm $\left(cum\_age_{ijt} = \sum_1^t age_{ijt}\right)$. Third, covariates that control for production-related aspects across industries. These include the share of exporters to total firms $\left(EXPORT_{jt} = 1/n \sum_i^n export_{ijt}\right)$, the average share of foreign investment in total investment $\left(DFO_{jt} = 1/n \sum_i^n dfo_{ijt}\right)$, average capacity utilization $\left(UTIL_{jt} = 1/n \sum_i^n util_{ijt}\right)$, and the share of auxiliary expenses in value-added $\left(OIEXPSH_{jt} = \sum_i OIEXP_{ijt} / \sum_i NVA_{ijt}\right)$ within the industry.[14]

[14] There could be multicollinearity issues across the independent variables. However, Table A6 in Appendix A indicates that it is less likely to pose major issue for the main estimating equation. This is because (i) there is no strong correlation among the covariates, and (ii) there is no correlation between the treatment variable (change of ERP) and covariates, except for the change in average firms' age in the industry but the degree of correlation is not strong.

Exporter share, skilled worker share, foreign investment share, capacity utilization, and auxiliary expense share in value-added are all scale-free. For these covariates, the first difference is applied straight from the level variable, thus having a percentage point change interpretation. Meanwhile, for capital intensity and average cumulative age of firms in the industry, they are first transformed into natural logarithmic values and then are differenced annually. Hence, the interpretation for these two covariates is percentage changes.

The change in industry performance, including aggregate labor share in the manufacturing sector, could be driven by policy reforms other than trade liberalization. This issue is noteworthy in the context of Indonesia since trade policy reforms often come in the package with other policies in many developing economies. Democratization (Abeberese et al., 2021) and minimum wage (Alatas and Cameron, 2008), for example, have been found to affect manufacturing firms' performance in Indonesia which eventually feed into the change in industry performance in aggregate. Another source of concern is that the initial characteristics of the industry and industry-specific policies could also confound the effect of trade. To the extent possible, this paper accounts for these other confounding shocks, especially the time-specific and industry-specific shocks, in the model by using the two-way fixed effects.[15] Regression with and without covariates are also presented to assess whether the confounders threaten the causal relation between ERP and labor share change.

Identification rests on the assumption that industry does not self-select into a certain ERP level. The deviation of ERP from the nominal protection rate can be thought of as exogenously determined from the perspective of that industry or workers' association. By using the two-way fixed effects, this paper isolates trade liberalization, measured by the change in ERP, as the main shock influencing the industry's performance. Import share in value-added is around 70% in Indonesia's manufacturing sector. Moreover, the Government of Indonesia has also undertaken significant reform in its trade policy regimes during the last few decades. Median output tariff in the manufacturing sector fell from 21% in 1990 to merely 8% in 2000, yet the rate at which trade reform has taken place varies across industries. Hence, the change in ERP likely serves as an important source of shock that could affect the overall industry's performance, including labor share.

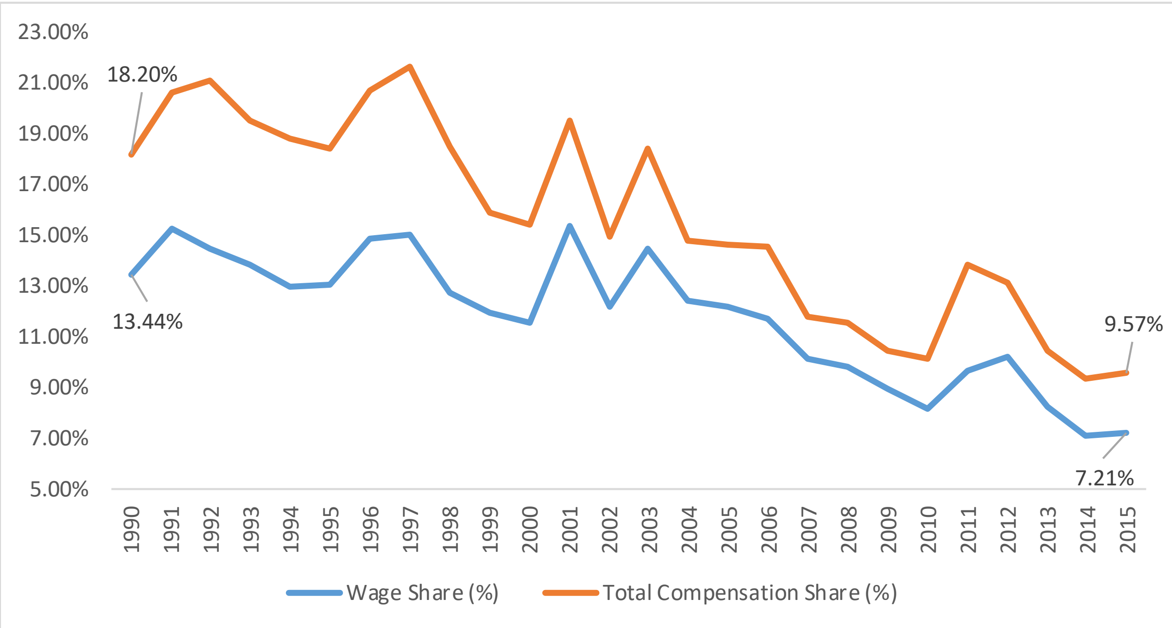
5. Results and Discussion

5.1. The Nature of Labor Share Decline in Indonesia's Manufacturing Sector

[15] As a sensitivity check, we provide estimation that control for industry-specific time trend to account for potentially differential trend in labor share change across industries. This is to address potentially spurious relation between ERP and labor share as labor share might have changed differentially across industries regardless of any trade policy changes.

Aggregate labor share in Indonesia's manufacturing sector dropped significantly between 1990 and 2015. It started at more than 13% in 1990, then fell gradually over time until it reached as low as 7% in 2015 (Figure 2). The decline occurs in most industries of the manufacturing sectors, albeit with some degree of variations (Figure 3).

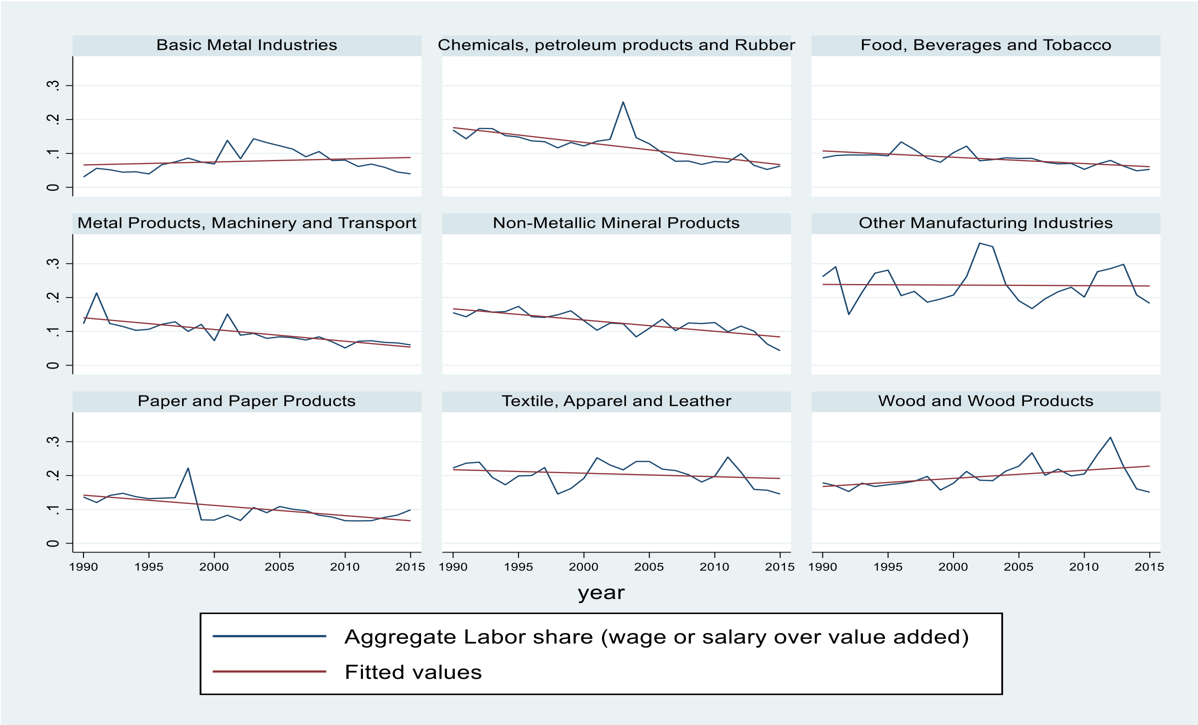
Figure 2. Aggregate labor share in Indonesia's manufacturing sector



Note: Aggregate labor share is calculated as the weighted average of firms' labor share, where each firm's share in total value added in manufacturing sector is used as weight.

Source: Authors' estimation.

Figure 3. Aggregate labor share across manufacturing industries at two-digit ISIC



Note: Aggregate labor share is calculated as the weighted average of firms' labor share within the two-digit ISIC rev.2 category. Each firm's value-added within that category is used as weight.

Source: Authors' estimation.

Using the dynamic decomposition (DOPD) method we break down the micro-level sources of the DLS in Indonesia's manufacturing sector. Several key observations emerge from the five-years decomposition (see Table 2 and Figure 4).[16] First, the change in aggregate labor share in Indonesia's manufacturing sector mainly originated from the labor share's change among the surviving firms, while the contribution of firms' entry and exit is small, although non-negligible. During 1990-2015, 92% out of the total 6.24 percentage point reduction in the aggregate labor share in Indonesia's manufacturing sector comes from the decrease in labor share among stayers, combining reallocation and within-firm channels. Only 8% comes from firms' exit (-2.30) and entry (1.76). Table 3 shows that, in general, the withdrawal of firms reduces aggregate labor share, while the entry of firms increases it. These findings suggest that both exiters and entrants tend to have a relatively higher labor share than the incumbents. This churning pattern, especially the firms' exit, is consistent with heterogeneous firms' literature. Since labor share is inversely related to productivity (Berkowitz et al., 2015), it is logical to expect that the less productive firms (with higher labor share) will withdraw

[16] The descriptive analysis in this section mainly focuses on the five-year period in order to isolate the overall big picture in the dynamics of labor share changes over the longer time. The five-year split is also the time interval that optimally uses all the available data, which is from 1990 to 2015.

from the market. The relatively higher labor share among entrants could mean that these new firms start in the industry with a low level of productivity. Other studies, such as Autor et al. (2020), have confirm this finding.

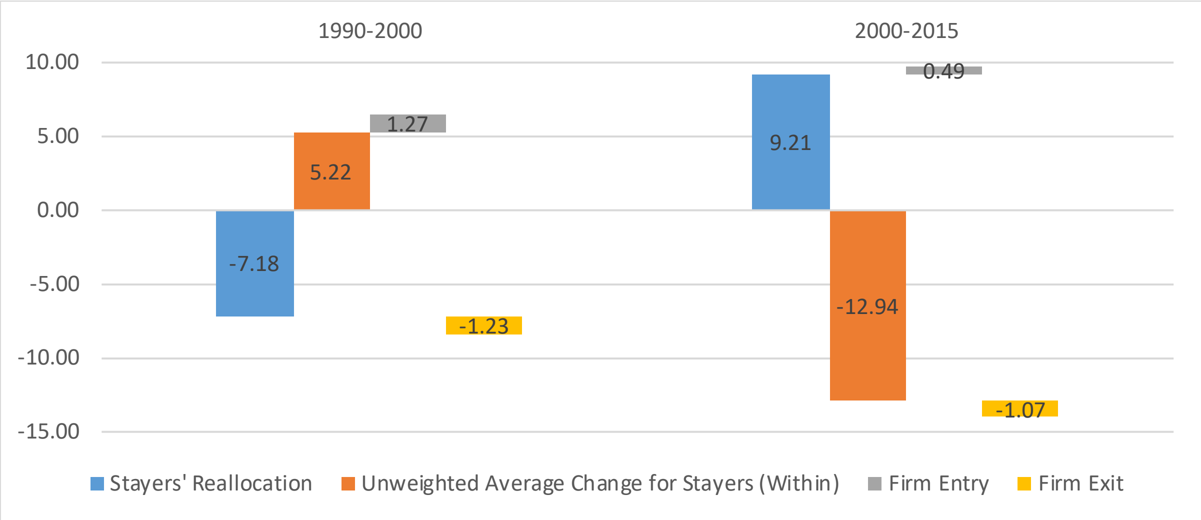
Table 2. Dynamic Olley-Pakes Decomposition (DOPD) of labor share changes in manufacturing sector (percentage point), 5-year interval

	Stayers' reallocation (1)	Unweighted average change for stayers (within) (2)	Firm entry (3)	Firm exit (4)	Total change (5)	Percentage changes relative to 1990 (%) (6)
A. Wage or salary share of value-added						
1990-1995	-4.55	3.70	1.17	-0.68	-0.37	
1995-2000	-2.62	1.52	0.10	-0.55	-1.55	
2000-2005	3.30	-2.71	0.25	-0.14	0.69	
2005-2010	4.17	-7.50	0.04	-0.81	-4.09	
2010-2015	1.74	-2.73	0.20	-0.12	-0.91	
1990-2000 (cumulative)	-7.18	5.22	1.27	-1.23	-1.92	-14.29
2000-2015 (cumulative)	9.21	-12.94	0.49	-1.07	-4.31	-32.09
1990-2015 (cumulative)	2.03	-7.72	1.76	-2.30	-6.24	-46.38
B. Total compensation share of value-added						
1990-1995	-4.52	4.62	1.00	-0.86	0.24	
1995-2000	-3.97	1.57	-0.12	-0.54	-3.06	
2000-2005	4.90	-5.83	0.25	-0.04	-0.72	
2005-2010	3.72	-7.33	-0.02	-0.90	-4.53	
2010-2015	1.77	-2.44	0.25	-0.14	-0.55	
1990-2000 (cumulative)	-8.50	6.19	0.88	-1.40	-2.82	-15.51
2000-2015 (cumulative)	10.40	-15.60	0.48	-1.08	-5.81	-31.92
1990-2015 (cumulative)	1.90	-9.41	1.36	-2.48	-8.63	-47.44

Note: Aggregate labor share is calculated as the weighted average of firms' labor share, where each firm's value added in manufacturing sector is used as weight. Decomposition of labor share changes is conducted in a five-year interval. It means the analysis uses information from the start and the end of the period while ignoring variations in between. Therefore, survivors, entrants, and exiters are all categorized within that five-year interval. All change is measured in percentage point changes unless specified otherwise.

Source: Authors' estimation.

Figure 4. DOPD of labor share changes in manufacturing sector (percentage point change)



Note: Aggregate labor share is calculated as the weighted average of firms' labor share, where each firm's value added in manufacturing sector is used as weight. Decomposition of labor share changes is conducted in a five-year interval. It means the analysis uses information from the start and the end of the period while ignoring variations in between. Therefore, survivors, entrants, and exit-ers are all categorized within that five-year interval. Labor share includes salary or wages component only. The period split accumulates the value derived from the five-year interval decomposition. For example, it means that the value of 1990-2000 represents the sum over 1990-1995 and 1995-2000 periods.

Source: Authors' estimation.

The joint importance of within-firm and reallocation channels differs from the previous findings in the literature. Studies that decompose the sources of aggregate labor share changes, such as by Böckerman and Maliranta (2012), Karabarbounis and Neiman (2014), and Autor et al. (2020), tend to find only one leading source. It was either the within-firm or reallocation channel. This study shows that the leading source of aggregate labor share decline could be non-monotonous depending on the firms' environment. Variations in the trade environment could potentially explain how aggregate labor share evolves over time. Next, we analyze whether trade liberalization contributes to the aggregate labor share decline and through what channel the former affects the latter.

Second, the decline of aggregate labor share in Indonesia's manufacturing sector does not always come from firm-specific reduction in labor share (within-firm channel). Both the reallocation channel and the within-firm changes among surviving firms play a crucial role in the decline of aggregate labor share. When trade liberalization took off rapidly in Indonesia's manufacturing sector between 1990 and 2000, within-firm labor share (unweighted average change of labor share for stayers as in column (2) of Table 2) was actually growing by more than five percentage points in cumulative term.

However, at the aggregate level, labor share has declined by 1.92 percentage points.[17] The DOPD exercise indicates that this decline mainly came from the changing firms' composition in the industry, where the market share was increasingly shifting towards low labor share firms during that period. On the other hand, from 2000 to 2015, when tariffs were relatively stable at a low level, the firm-specific (within-firm) reduction in labor share was the main driver behind the aggregate labor share fall (see column (1) and (2) of Table 2). Decomposition based on annual data (see Figure A2 in Appendix A) and a more disaggregated industry-level (see Table A7 in Appendix A) shared a similar story.

5.2.The Effect of Trade Liberalization on Aggregate Labor Share

5.2.1. Main Results

Table 3 provides the main estimates of the effect of ERP change on aggregate labor share change and its components. Linear decomposition provided by the DOPD method, as given by formula (3), allows for breaking down the total trade effect into four components. Hence, the sum of columns (2)-(5) is approximately close to column (1) $(0.000517 \approx 0.000758 + (-0.000190) + (-0.000181) + 0.000141)$.

[17] This fact is confirmed using overall firm-level distribution for the year of 1990 and 2000. Figure A1 of Appendix A shows that the cumulative distribution function of manufacturing firms shifted to the right in 2000 compared to 1990.

Table 3. The effect of ERP on aggregate labor share and its components

	(1)	(2)	(3)	(4)	(5)
	Total change	Stayers' reallocation	Unweighted average change for stayers (within)	Firm entry	Firm exit
Change of ERP	0.000517**	0.000758*	-0.000190	-0.000181***	0.000141**
	[0.000164]	[0.000291]	[0.000357]	[0.0000509]	[0.0000416]
Change of ERP (standardized coefficient)	0.00609**	0.00877*	-0.00172	-0.00596***	0.00465**
	[0.00193]	[0.00336]	[0.00324]	[0.00168]	[0.00137]
N	1948	1948	1948	1879	1865
R^2	0.068	0.308	0.312	0.017	0.013

Note: All estimations are based on TWFE regression at 4-digit industry categories, where each observation is weighted by its value-added share at the initial observation period. Dependent variables include the total change in aggregate labor share and its decomposed components. Standard errors are in brackets and clustered at the industry level. Standardized coefficient reflects how many standard deviation the dependent variable will change due to one unit increase in the standard deviation of the independent variable. The following sign shows significance level: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Source: Authors' estimation

The result in column (1) shows that lower ERP reduces aggregate labor share. Specifically, industries that are exposed to greater trade liberalization face a larger reduction in their labor share. A one percentage point cut in ERP rate leads to a 0.000517 percentage point reduction in aggregate labor share. In a standardized format, this result can be interpreted as one standard deviation decrease in the change of ERP leading to a 0.006 standard deviation decrease in aggregate labor share change. The small magnitude of the coefficient reflects the nature of the large reduction in tariff but a small and gradual decrease in aggregate labor share over the years. Other studies also found a small coefficient where one percentage point change in tariff leads to a much lower percentage point change in labor share (Ahsan and Mitra, 2014, Kamal et al., 2019, Leblebicioğlu and Weinberger, 2021). Despite the small magnitude, the effect is significant at the one percent level.

The results in columns (2)-(5) indicates that the labor share-reducing effect of trade liberalization works mainly by inducing the reallocation of market share towards low labor share firms in the industry (0.000758). This intra-industry reallocation effect of trade outweighs other channels in driving the decline of aggregate labor share. We also find that trade liberalization does not reduce labor share within firms.

Table 4 also shows that aggregate labor share driven by firms' turnover (entry and exit) is more prominent in the industries that face greater trade liberalization. Particularly, a reduction in ERP enhances aggregate labor share through the entry of new firms into the industry (-0.000181). This is due to the entrants' characteristics that tend to have a relatively higher labor share than incumbents. On the other hand, trade liberalization contributes to aggregate labor share decline through the exit of high labor share firms from the market (0.000141). Note that the coefficient of ERP in these two channels should be interpreted with caution. The turnover of firms in the SI dataset does not always represent true exit and entry in the industry but might have to do with inconsistent participation in the SI census. However, the contribution of trade's effect via firms' entry and exit is relatively tiny and essentially cancels each other out. Hence, it should not alter the overall result.[18]

Overall, this paper discovers that trade liberalization can produce the superstar firm effect a-la Autor et al. (2020), where market share is reallocated towards firms with low labor share following the increase in market toughness. Here, trade liberalization measured by the decline in ERP can be interpreted as the force that raises competitiveness in the market. Not only does this result is consistent with the superstar firms model, but it also resembles the prediction made in other classes of heterogeneous firms' models that highlight the importance of the intra-industry reallocation effect of trade, such as Melitz (2003) and Melitz and Ottaviano (2008).

The approach in this study has also allowed reconciling the seemingly disconnected story between the labor share-enhancing effect of trade liberalization in some studies (Ahsan and Mitra, 2014, Kamal et al., 2019, Leblebicioğlu and Weinberger, 2021) and the actual aggregate labor share trend in the countries under investigation. Those studies exploit only firm-level variation in labor share and largely ignore the role of compositional shift of firms within an industry. They have found that trade liberalization increases firms' labor share in the case of China and India. However, aggregate labor share shows a declining trend amidst significant trade reform in those countries. Our paper has found that other mechanism is at work and have been ignored by those studies, namely the intra-industry reallocation or compositional shift. Leblebicioğlu and Weinberger (2021) have emphasized the importance of accounting for the intra-industry reallocation effect in explaining the fall of aggregate labor share in India, but they do not formally test it in their paper. This study provides the first evidence in the case of developing economies that the intra-industry reallocation effect is a critical channel that links trade to labor share a-la superstar firms model.

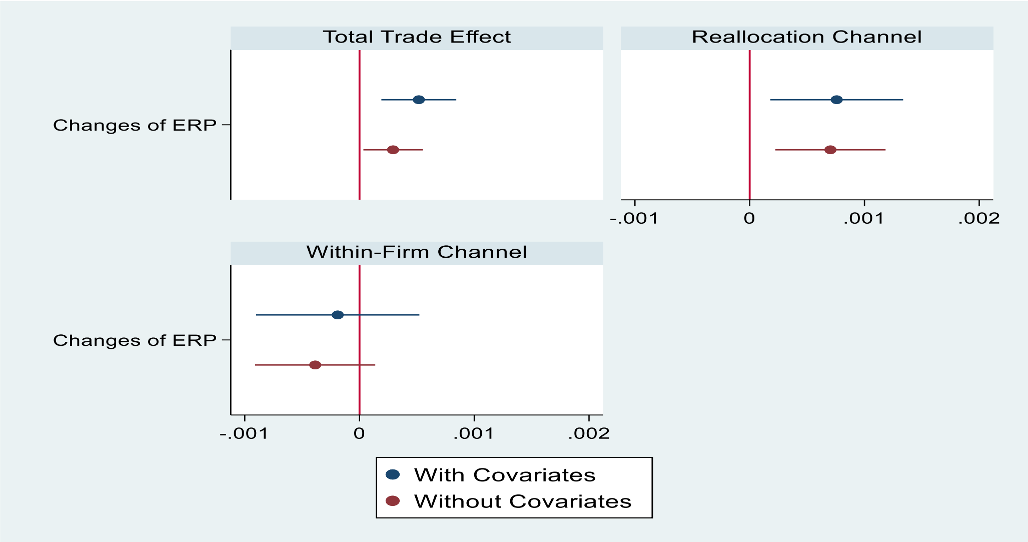
[18] See Table A8 of Appendix A for the full result that includes all covariates.

5.2.2. Robustness Check

This paper further assesses several potential threats that could derail the credibility of the main results. The first relates to the key identifying assumption around the ERP. We assume that ERP is unanticipated from the industry's perspective, which allows for the direction of causality to go from ERP to labor share and not the other way around. In other words, the future change in ERP should not be anticipated by current industry's outcome for the assumption to work. We test the plausibility of this strict exogeneity assumption following the strategy suggested by Wing et al. (2018), where the standard regression is augmented by two years leads and lags variables. The result in column (5) of Table 4 shows that this model has no anticipation effect, as shown by statistically insignificant forward treatment variables either individually or jointly. This finding supports the exogeneity assumption of ERP change.

Next, we assess whether the results are sensitive to different specifications, econometric models, and choice of policy and outcome variables. Figure 5 shows that the effect of trade is not sensitive to the choice of covariates. The labor share-reducing effect of lower trade protection is still maintained using the model with and without covariates. The importance of the reallocation effect of trade on labor share also holds in both models, with no significant within-firm channel regardless of the model choice. The magnitudes of the trade effect from the two models are also not wildly different, with the point of estimates still lying within each other's confidence interval.

Figure 5. Comparison of estimates with and without covariates



Note: All estimations are based on TWFE regression at 4-digit industry categories, where each observation is weighted by its value-added share at the initial observation period. The dot represents the point of estimate, while the line is the 95% confidence interval. The confidence interval that crosses the red reference line indicates a statistically insignificant effect at the five percent level.

Source: Authors' estimation.

There is also a possibility that aggregate labor share across industries changed differentially over time regardless of the trade liberalization exposure, hence capturing the spurious relation between treatment and outcome variables of interest. In addition, labor share movement may also be influenced by s shock that occurs specifically at a particular time for certain industries, hence cannot be adequately captured by standard TWFE model that we use in our baseline estimation. To address the former, we add an industry-specific time trend to the main model to check if the secular change in aggregate labor share could explain away the treatment effect. For the latter issue, we incorporate the interaction of year and 2-digit industry fixed effects. Column (2) and (3) of Table 4 shows that the statistically significant effect of ERP on labor share is still intact despite controlling for the secular change trend. The reallocation channel underpinning the effect of ERP on labor share is also still maintained (see Table A9 of Appendix A).

Table 4. Robustness check: alternative measures and specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dep: Total change in labor share	Baseline	With industry-specific time trend	With year-industry (2 digit) FE	Total compensation	Leads & lags	OLS	OLS with initial control	ERP with technical coefficient of 2006	Original (missing) tariff data	Ad-valorem tariff only
Change of ERP	0.000517**	0.000850***	0.000596***	0.000997***	0.00128+	0.000505***	0.000520***	0.00113*	0.0605**	0.0343+
	[0.000164]	[0.000231]	[0.000163]	[0.000277]	[0.000662]	[0.000139]	[0.000138]	[0.000511]	[0.0207]	[0.0183]
Lag (1)-Change of ERP					-0.000826					
					[0.00135]					
Lag (2)-Change of ERP					-0.000596					
					[0.00130]					
Lead (1)-Change of ERP					-0.00969					
					[0.0213]					
Lead (2)-Change of ERP					0.0229					
					[0.0207]					
N	1948	1948	1948	1948	1638	1948	1948	1948	1404	1948
R^2	0.068	0.070	0.199	0.067	0.081	0.068	0.068	0.068	0.087	0.069
Industry FE	✓	✓	✓	✓	✓					
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year-Industry (2-digit) FE			✓							
Covariates	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline & initial level	Baseline	Baseline	Baseline
F-test *p*-value		0.0000 (industry			0.0043 (lags)					

time trend)	0.3068 (forwards)

Note: All estimations are based on TWFE regression at 4-digit industry categories, where each observation is weighted by its value-added share at the initial observation period. Standard errors are in brackets and clustered at the industry level. The following sign shows significance level: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Source: Authors' estimation

Changing the econometric model only has little impact on the overall result. In columns (6)-(7) of Table 4, OLS models are used instead of TWFE. They produce estimates that are relatively close to the main results. The inclusion of industries' initial characteristics in the OLS model also does not matter much for the treatment variable of interest. These initial time-invariant covariates are absorbed by the industry fixed effect in the main estimates.

The labor share-reducing effect of trade liberalization is also maintained when using: (i) total compensation, instead of only salary, as the outcome variable (column (4)) and (ii) different measurement techniques in constructing the ERP variable. The latter includes using the technology coefficient of the year 2006 instead of the average of 2000 to 2015 (column (8)), original tariff data with missing information (column (9)), as well as ad-valorem only tariff data (column (10)).[19] However, the choice of tariff data could be rather sensitive to the decomposed result, not the total effect (see Table A9 of Appendix A). This might have to do with how close the tariff estimation is to the actual protective condition in the industry. In the main model, this paper prefers to use the average tariff that includes the AVE estimation of the specific tariffs in ERP construction because it is arguably closer to the actual protection given to the industry.

In summary, the robustness check shows that the negative effect of trade liberalization on aggregate labor share does not wash away by changing model specifications and altering the choice of variables. The primacy of the intra-industry reallocation effect of trade on labor share also survives most of the alternative models (see Table A9 of Appendix A).

[19] The ad-valorem only tariff data exclude the ad-valorem equivalent estimates of specific tariff.

5.2.3. *Effect Heterogeneity*

The extent of trade liberalization varies over time, and industries vary in characteristics. Therefore, it is interesting to see if the effect of trade liberalization on labor share is particularly more substantial during specific periods and in certain groups of industries. We check the heterogeneous effect of trade on labor share by splitting observations into different groups. The first group is based on the period where observations are divided into the period 1990-2000 and 2000-2015. Import tariffs experienced more rapid reduction pre-2000, while it was relatively stable post-2000. We would like to assess if the trade-induced labor share reduction varies depending on the degree of tariff cuts that are heterogeneous across time. The next two groups classify industries according to technology intensity and whether the sector constitutes part of network trade operated under the global production network (GPN) scheme.[20]

As Table 5 shows, there are considerable differences across the groups. In terms of period heterogeneity, the effect is significant only during the big tariff changes in the 1990s but not in the 2000s, where the tariff was relatively stable at a low level. This helps explain the small result in the main estimates as the entire period mixes the significant effect in the 1990s with the insignificant one (2000s). This result is broadly consistent with the tendency of the initial effect

[20] We use OECD classification to categorize high technology and low technology sectors, while the GPN sectors are defined following classification suggested by Athukorala (2014). The GPN sectors include textile (SITC 656 and 657), general machinery (SITC 74), automated data processing machines (SITC 75), telecommunication products (SITC 76), electrical machinery (SITC 77), road vehicles (SITC 78), other transport equipment (79), travel goods (SITC 83), apparel and clothing (SITC 84), footwear (SITC 85), professional and scientific instruments (SITC 87), photographic apparatus (SITC 88), toys and sport goods (SITC 894).

to dissipate over time, as shown by the negative sign in the lagged treatment coefficients (see column (5) of Table 4).

Table 5. Heterogeneous effect of ERP on aggregate labor share

	(1) Total change	(2) Stayers' reallocation	(3) Unweighted average change for stayers (within)	(4) Firm entry	(5) Firm exit
			By period		
1990-2000	0.000735**	0.000958***	-0.0000773	-0.000277***	0.000221***
	[0.000268]	[0.000263]	[0.000422]	[0.0000563]	[0.0000455]
2000-2015	0.0601	-0.0934	0.0956	0.0527*	0.0182
	[0.0719]	[0.245]	[0.207]	[0.0202]	[0.0197]
			By organizational characteristics in the global market		
GPS Sectors	0.119+	-0.00149	0.0940	0.00586	0.0181
	[0.0612]	[0.0819]	[0.0825]	[0.00626]	[0.0112]
Non-GPS Sectors	0.000199***	0.000596*	-0.000378+	-0.000180***	0.000164***
	[0.0000548]	[0.000261]	[0.000217]	[0.0000366]	[0.0000172]
			By technology intensity		
Hi-Tech Sectors	-0.0516	0.0687	-0.139	-0.00315	0.0238
	[0.0607]	[0.0495]	[0.101]	[0.0184]	[0.0264]
Low-Tech Sectors	0.000433**	0.000682*	-0.000215	-0.000178**	0.000145**
	[0.000154]	[0.000305]	[0.000335]	[0.0000541]	[0.0000458]

Note: All estimations are based on TWFE regression at 4-digit industry categories, where each observation is weighted by its value-added share at the initial observation period. Standard errors are in brackets and clustered at the industry level. The following sign shows significance level: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Source: Authors' estimation.

With regards to sectoral heterogeneity, the results show that the labor share-reducing effect of trade liberalization appears particularly strong in low-technology sectors. This result is somewhat surprising considering that industries operating in high-technology environments constantly face potential disruption in their sectors. So, intuitively, the pressure to reduce labor costs is more considerable for high-technology sectors than the low-technology ones. This paper does, in fact, document a more considerable fall in labor share for the high-technology sectors compared to low-tech ones (see Table 6). However, the high-tech sectors are virtually operating in a free trade environment, as they face a very low ERP, to begin with. The insignificant effect of domestic trade liberalization on labor share for these high-tech sectors suggests that the pressure to reduce labor costs may come from somewhere else.

Meanwhile, trade liberalization reduces labor share both in GPN and non-GPN sectors, with a much stronger significance level in non-GPN sectors but a larger magnitude for GPN sectors. There is no apparent significant channel that drives the effect of trade on labor share in the GPN sector, while the intra-industry reallocation is the strongest driver for trade-induced labor share decline in the non-GPN sectors. The significant effect of trade liberalization in driving the decline of labor share in the GPN sectors signifies the role of cost-efficiency (Athukorala, 2014) and liberal trade policies (Fernandes et al., 2022) in the sectors organized within the networked trade. On the other hand, the non-GPN sectors are exposed to a more significant reduction in ERP compared to the GPN sectors. This helps explain the significant effect of trade liberalization on labor share within the non-GPN sectors.

Table 6. Mean comparison of labor share and ERP changes by group of sectors and periods

	(1) Changes in aggregate labor share (annual average, p.p. change)	(2) Change in ERP (annual average, p.p. change)	(3) Initial level of ERP (%)
1990-2000	-0.416	-20.245	
	[0.759]	[11.366]	
2000-2015	-0.190	-0.049	
	[0.400]	[0.362]	
t-test p-value	*0.775*	*0.030*	
GPN Sectors	-0.408	-0.556	23.518
	[0.730]	[0.261]	[3.128]
Non-GPN Sectors	-0.196	-13.122	345.312
	[0.425]	[7.556]	[328.225]
t-test p-value	*0.788*	*0.177*	*0.429*
Hi-Tech Sectors	-0.435	-0.039	5.916225
	[0.794]	[0.279]	[10.742]
Low-Tech Sectors	-0.211	-11.722	311.4208
	[0.433]	[6.577]	[285.477]
t-test p-value	*0.790*	*0.237*	*0.479*

Note: Mean estimation is based on the unweighted average across observations within particular groups. Standard errors of mean estimation are in brackets.

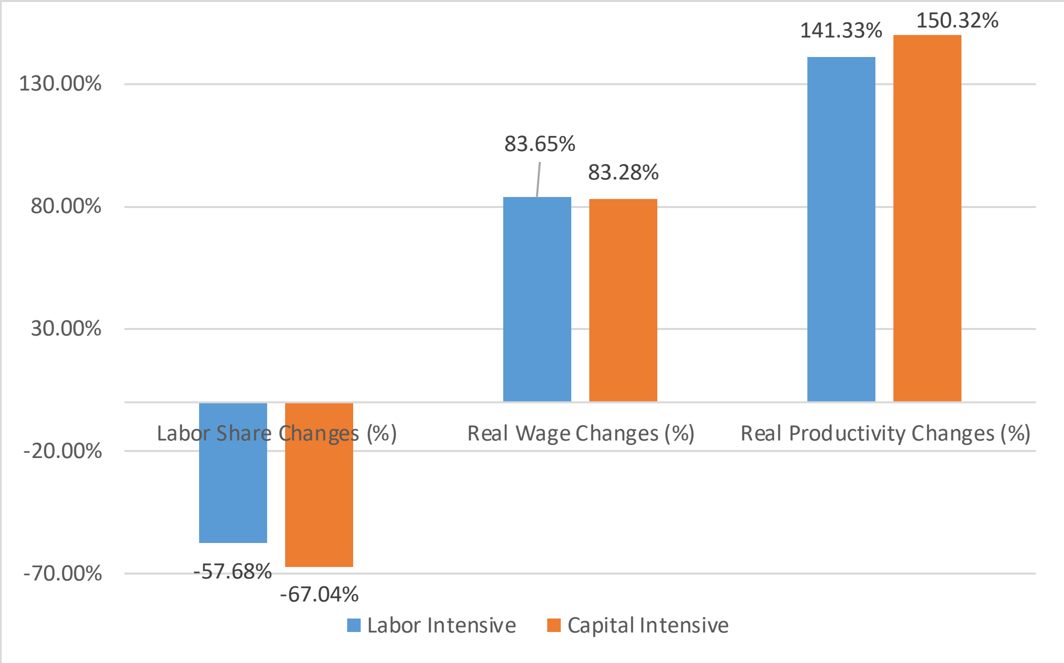
Source: Authors' estimation.

The results in all groups are consistent with the main estimates, in which the intra-industry reallocation into low-labor-share superstar firms is the main channel driving trade liberalization's effect on aggregate labor share.

5.3. Further discussion on main findings

So far, the findings have indicated that trade liberalization contributes to the decline of aggregate labor share in Indonesia’s manufacturing sector by inducing reallocation into superstar firms in the industry. Although trade liberalization reduces labor share in the aggregate, further decomposition, as in equation (3), reveals how workers might benefit from trade liberalization. Table 7 shows that trade liberalization measured by lower ERP raises aggregate real wages and labor productivity across industries. However, it is found to reduce aggregate employment. The real wages in labor-intensive industries also increased markedly from 1990 to 2015 (see Figure 6). These results suggest that trade liberalization in Indonesia’s manufacturing sector has induced specialization in labor-intensive industries, further raising wages across the board, in particular for labor-intensive sectors. However, this trade-induced expansion seems to have been achieved through reduced employment but with larger labor productivity (see negative sign in columns (3) and positive sign in column (4) of Table 7). Capital intensity has also increased in aggregate as the compositional shift towards low labor share firms means moving towards more capital-intensive firms (see Figure 7). These processes have resulted in the DLS in Indonesia’s manufacturing sector.

Figure 6. Changes in aggregate labor share, real wages, and real productivity (%), 1990-2015 (cumulative), by labor intensity



Note: Labor intensity follows Aswicahyono et al. (2010) classification. We further group the electronics sector into labor-intensive industries due to likely heavy assembling activities in the sector.

Source: Authors' estimation.

Table 7. Decomposed effects of ERP on aggregate labor share and its elements

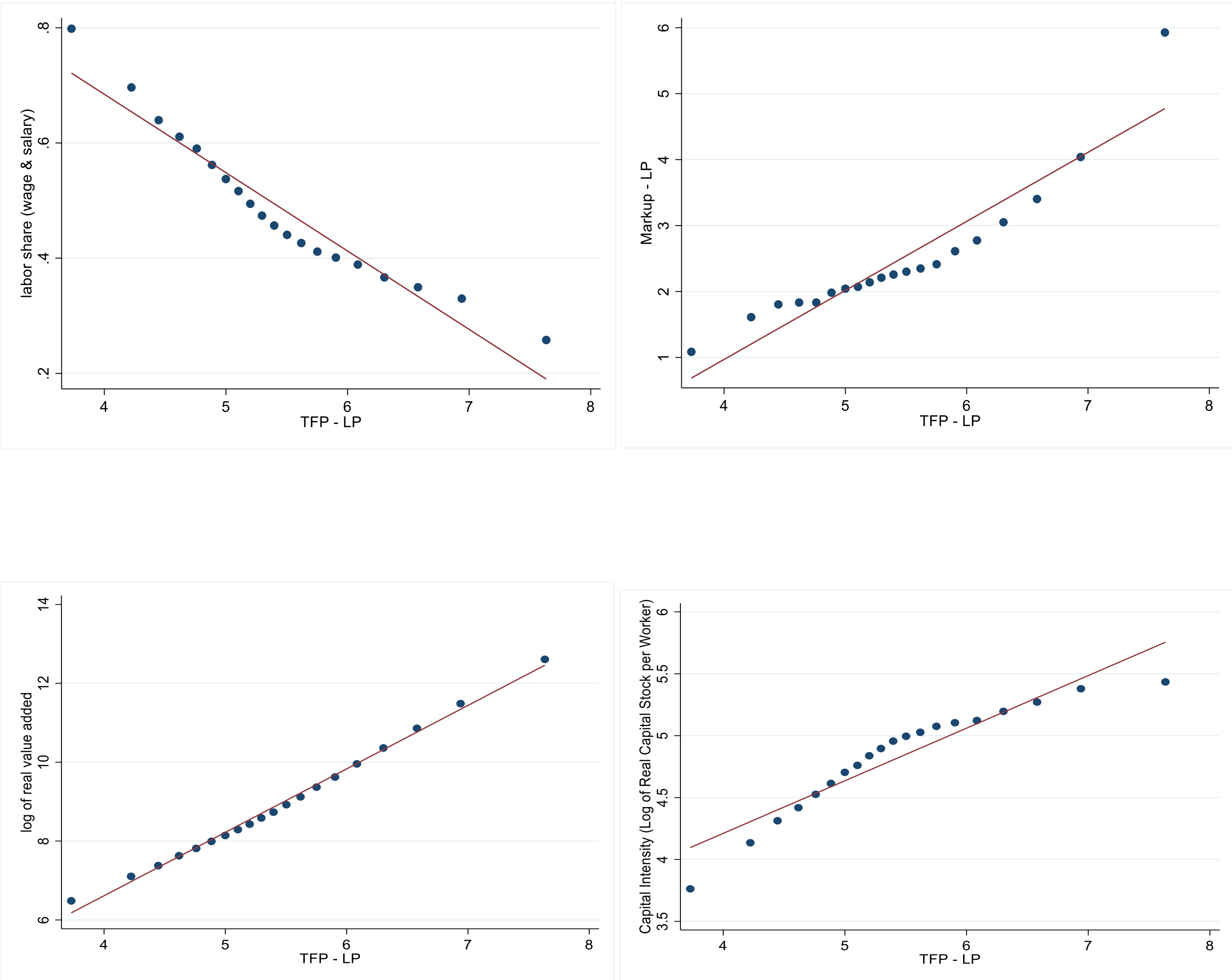
	(1) Total aggregate labor share change	(2) Real wage change	(3) Real productivity change	(4) Employment change	(5) Real value-added change
Decomposition by real wages and productivity					
Change of ERP	0.00265*	-0.00157**	-0.00422***		
	[0.00118]	[0.000581]	[0.00119]		
Decomposition by real wages, employment, and real value-added					

Change of ERP	0.00265*	-0.00157**		0.00117**	-0.00305*
	[0.00118]	[0.000581]		[0.000403]	[0.00135]
N	1948	1948	1948	1948	1948
R^2	0.116	0.094	0.162	0.133	0.149

Note: All estimations are based on TWFE regression at 4-digit industry categories, where each observation is weighted by its value-added share at the initial observation period. All the models use the same covariates as in the main model. Standard errors are in brackets and clustered at the industry level. The following sign shows significance level: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Source: Authors' estimation.

Figure 7. Relation between firms' productivity, labor share, markup, size, and capital intensity



Note: These are bin scatter plots of firms' performance, controlling for 4-digit ISIC by year dummies. Hence, they reflect relationships within the industry in a particular year. Capital intensity is computed as real capital stock per worker. We use only the imputed data of set 1 to establish relation. The result is almost identical using the other four imputation datasets and is available upon request.

Source: Authors' estimation.

From a micro-level perspective, we confirm that low labor share firms tend to have higher productivity. In addition, the more productive a firm is, the larger the size and the more likely it is to charge larger markup and have higher capital intensity. Higher markup among more productive firms is consistent with the heterogenous firms' model (Melitz and Ottaviano, 2008, Autor et al., 2020). Productive firms are able to operate at low marginal cost while only incompletely passing the efficiency gain to the consumer price due to their market power in the industry. Furthermore, the highly productive firms with low labor share are also associated with greater capital intensity, which is broadly consistent with the prediction in Berkowitz et al. (2015), where capital-intensive firms tend to have a lower labor cost component in their value-added.

The superstar firms model does not explain why a market competitiveness shock triggers reallocation into low labor share firms. This paper finds that the low labor share firms are equipped with higher productivity, size, and markups; hence they have a higher chance of surviving a competitive open trade environment. As the market becomes more competitive due to a more liberalized trade policy, the industry is then increasingly dominated by better-performing firms with more capital intensity, hence lowering aggregate labor share. This is consistent with the selection effect of trade predicted in the heterogeneous firms model (Melitz, 2003, Melitz and Ottaviano, 2008). Thus, DLS could be seen as an inevitable outcome as trade liberalization selects only the more productive ones. We find preliminary support for this notion. It is shown in Figure A3 of Appendix A that the industry reallocates into more productive firms during the 1990s period where strong trade's effect is detected.[21]

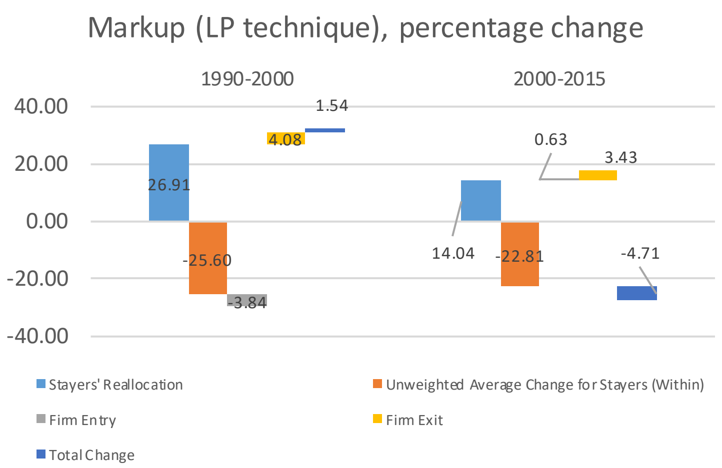
[21] However, this is not to conclude that reallocation into more productive ones is actually induced by trade liberalization. This is beyond the scope of this paper.

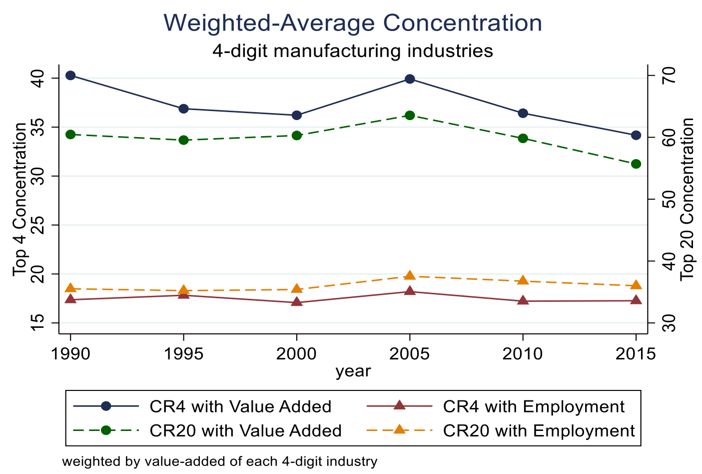
While trade liberalization can trigger superstar firms effect, the superstar firms model cannot fully characterize the labor share dynamics in Indonesia's manufacturing sector.[22] We find that the increase in markup due to reallocation into high-markup superstar firms in the industry is more than compensated by the decrease in average markup within firms (Figure 8). As a result, aggregate markup does not rise in Indonesia's manufacturing sector. In addition, aggregate concentration tends to be steady as well. This stands in contrast to the full superstar firms model that predicts a considerable increase in markup and concentration in the industry due to reallocation into better-performing superstar firms.[23] Therefore, although this paper detects the superstar firms effect which is induced by trade liberalization, it does not find full support for the mechanics outlined in the model.

[22] Autor et al. (2020) assert that reallocation of market share into better performing and high markup firms (called "superstar firms") could raise aggregate markup and overall concentration in the industry, which then consequently decrease the relative income payments to workers. Based on this thinking, the rise of superstar firms in the industry should have been the culprit behind the aggregate labor share fall in many countries.

[23] This story still holds even when using the decomposition method based on the annual changes and different measurement technique for concentration ratio (see Figures A4-A5 in Appendix A). We thus conclude that aggregate markup does not rise in Indonesia's manufacturing sector.

Figure 8. DOPD of (log) markup and aggregate manufacturing sector concentration (percentage)





Note: Aggregate markup is calculated as the weighted average of firms' (log) markup, where each firm's value added in manufacturing sector is used as weight. Decomposition of markup changes is conducted in a five-year interval. It means the analysis uses information from the start and the end of the period while ignoring variations in between. Therefore, survivors, entrants, and exit-ers are all categorized within that five-year interval. All changes are measured in percentage changes. Markup is recovered after estimating production function coefficient using LP technique. The period split accumulates the value derived from the five-year interval decomposition. For example, it means that the value of 1990-2000 represents the sum over 1990-1995 and 1995-2000 periods. The concentration ratio is first calculated at 4-digit ISIC level and then aggregated up to the overall manufacturing sector using each industry's share of value-added.

Source: Authors' estimation.

Instead, the accompanying industry outcomes in Indonesia's manufacturing sector show a resemblance to the prediction made by Melitz and Ottaviano (2008). Their model predicts a strong pro-competitive gain from trade within firms that exceeded the increase of markup due to reallocation into high-markup firms following trade openness. Hence, trade liberalization could be the differentiating factor that prevents markup from rising as per the prediction in the superstar firms model.[24] This adds support to the critical role that trade liberalization plays in explaining the DLS in Indonesia's manufacturing sector.

5.4. Is the effect of trade confounded by other competing factors?

Studies have put forward various explanations for the phenomenon of labour share decline. Apart from trade, the literature has highlighted the important role of technological change, rising product market power of firms, as well as declining bargaining power of workers as potential sources of declining labor share (see Grossman and Oberfield (2022) for more detailed review). Furthermore, other policy reforms may coincide with trade reform during the period of this study. All of these factors can potentially confound the estimation of the effect of trade on labor share if not controlled appropriately. The empirical strategy has added industry and period fixed effects to control any time- and industry-specific policy shocks that may happen during the period of observations. Some relevant industry characteristics have also been controlled to minimize the confounding issues. However, there may be other factors that have not been adequately accounted for, especially the ones that vary across time and industries. We evaluate the extent to which these other factors may confound the effect of trade on labor share in this study by focusing on the two big themes in the labor share's literature: the role of superstar firms (Autor et al., 2020) and capital-biased technological change (Karabarbounis and Neiman, 2014).

As noted, in their superstar firms model, Autor et al. (2020) highlight that the reallocation of market share into high-markup superstar firms could raise aggregate markup and overall

[24] It is important to note that this is not definitive evidence on the impact of trade liberalization on markup.

concentration in the industry. This will, in turn, depress the relative income paid to workers. Empirical studies have supported the superstar firms' hypothesis and discovered that market power and markup are essential factors behind the labour share decline (De Loecker et al., 2020, Autor et al., 2020). Hence, omitting markups and market power from the regression may threaten the validity of the result as it could capture a false positive effect of trade liberalization where it might otherwise be a reflection of changing industry's markups and market power.

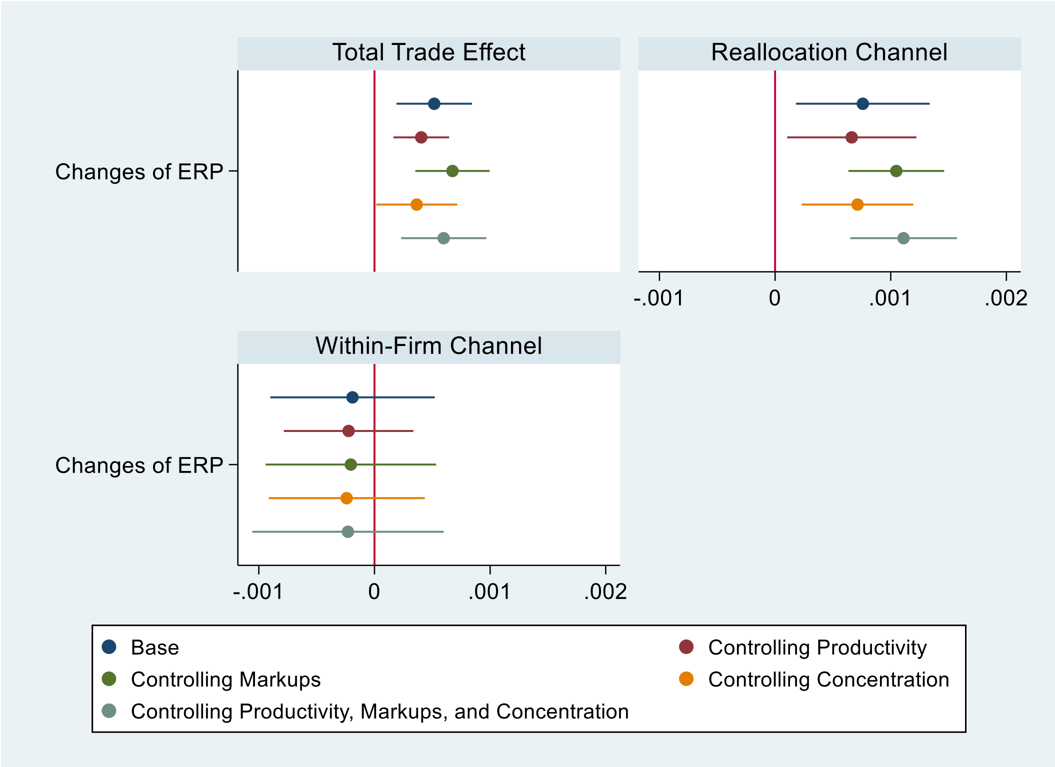
To assess the extent of this problem, we modify the base model by controlling for changes in markups (ΔM_{jt}) and concentration ratio of the four biggest firms in the industry ($\Delta CR4_{jt}$). Consistent with superstar firms' literature, markup and the extent of market power measured by CR4 are all associated negatively with labor share (see Table A10 in Appendix A). However, the effect of trade liberalization on labor share does not go away and only changes a little with the addition of these confounding factors (see Figure 10). This result indicates that trade liberalization matters in explaining the aggregate labor share decline. The confounding issues presented by those characteristics do not distort the result much; hence we argue that the trade effect is not merely a reflection of the movement in markups and concentration.

Considering that aggregate markup is relatively steady, the DLS is likely translated into higher income for capital owners. Under these circumstances, the role of capital-biased technological change could not be ignored. An exogenous change might have occurred in the manufacturing sector resulting in a favorable condition for the greater adoption of capital compared to labor. This shock could stimulate greater capital intensity in the industry, creating downward pressure for aggregate labor share. Such presence of capital-biased technological change makes it difficult to ascertain the role of trade liberalization.

Currently, literature has offered two ways of looking at this issue. On the one hand, capital-biased technological change is seen as a separate force from trade liberalization in influencing aggregate labor share dynamics. Karabarbounis and Neiman (2014) argue that the decrease in the relative price of capital goods induced by the advancement in ICT has stimulated greater use of capital at the expense of labor, thus decreasing the labor share in many countries. However, on the other hand, trade could also induce capital-biased technological progress. Several studies highlight that firms alter their production technology when facing tougher trade competition (Navas-Ruiz and Sala, 2007, Saad, 2017, Impullitti and Licandro, 2017, Lileeva and Trefler, 2010). A greater market access and selection effect of trade rationalize firms' decision to invest in cost-saving technology. As a result, the adoption of new technology, especially the labor-saving one, will likely influence labor share at the aggregate level.

This paper acknowledges the potentially important role of capital-biased technological change in the DLS. Mainly, we consider any changes that result in greater adoption of capital separate from the effect of trade liberalization. We control for the industry's capital intensity ($\Delta KINT_{jt}$) in the main regression to check if the effect of trade liberalization is still relevant. As expected, the regression result shows that greater capital intensity is inversely related to labor share (see Table C8 in Appendix C). However, the effect of trade liberalization is still present after controlling for capital intensity. Further exercise that controls for a measure of technological progress, namely TFP growth ($\Delta\Phi_{jt}$), also does not meaningfully change or explain away the effect of trade liberalization, even though productivity is inversely related to labor share (see Figure 9 and Table A10 in Appendix A). This result shows that there is a distinct effect of trade liberalization on labor share, which is not a mere reflection of changes in capital intensity and technological progress.

Figure 9. The effect of ERP on labor share, controlling for TFP and markup



Note: All estimations are based on TWFE regression at 4-digit industry categories, where each observation is weighted by its value-added share at the initial observation period. The dot represents the point of estimate, while the line is the 95% confidence interval. The confidence interval that crosses the red reference line indicates a statistically insignificant effect at the five percent level. The corresponding decomposed components for markups and productivity are used as covariates when regressing against the decomposed components of labor share changes as the outcome variable.

Source: Authors' estimation.

Despite mainly treating capital-biased technological change as a separate force, this paper also documents that the increasing capital intensity in the industry is an inherent characteristic following trade liberalization. Since lower labor share firms have higher capital intensity (recall Figure 8), when the industry reallocates towards firms with lower labor share following trade liberalization, it will also gravitate toward capital-intensive firms. As a result, aggregate labor share will decline, and capital intensity will likely rise. However, the role of trade in inducing greater capital intensity, in this paper, is confined within the boundary of the reallocation mechanism. Therefore, nothing can be said about how trade liberalization may induce firms to alter their capital intensity.

Despite the relatively stable result after controlling for potential observed confounders, one can still argue that there is a potential role of unobserved confounders that may contaminate the estimated effect. To assess this threat, we follow Oster's framework (Oster, 2019) and assume that selection on unobservable is as important as that on observable. We also assume that in a hypothetical setting where all unobserved controls can be accounted for, outcome can be explained by treatment and full controls, including both the observed and the unobserved, by around 90 to 100 percent (R^2_{max}), in particular due to the fact that labor share estimation is less likely plagued by measurement error.[25] To assess robustness of our baseline estimate from potential unobserved confounding effects, we produced a set of bounds for the treatment effect. One bound is the effect from baseline estimate where only observed covariates are used, while the other bound is the adjusted treatment effect, where equal proportionality assumption is used and R^2_{max} is set at 90 and 100 percent. If the identified set of bounds does not include zero, this raises the confidence that the estimated baseline effect captures the causal effect of change of ERP on aggregate labor share changes. Table 8 shows the set of bounds produced by using equal proportionality assumption and setting R^2_{max} at given values.

[25] Oster (2019) suggested that practitioners should consider setting R^2max below 100 percent due to the problem of measurement error.

Table 8. Estimated bounds of ERP effect with proportionality assumption

Dependent (labor share)	(1) Effect of change of ERP [Std. error], (R^2)	(2) Set of bounds, δ=1 & R^2_{max}=0.9	(3) Set of bounds, δ=1 & R^2_{max}=1
Total change	0.000517**	[0.000517, 0.00360]#	[0.000517, 0.00403]#
	[0.000164]		
	(0.0680)		
Stayers' reallocation	0.000758*	[0.000288, 0.000758]#	[0.000206, 0.000758]#
	[0.000291]		
	(0.308)		
Unweighted average change for stayers (within)	-0.000190	[-0.000190, 0.000740]	[-0.000190, 0.000901]
	[0.000357]		
	(0.312)		
Firm entry	-0.000181***	[-0.000181, 0.00475]	[-0.000181, 0.00601]
	[0.0000509]		
	(0.0170)		
Firm exit	0.000141**	[-0.0101, 0.000141]	[-0.0141, 0.000141]
	[0.0000416]		
	(0.0129)		

Note: Standard errors in brackets only for column (1). Brackets in column (2) – (3) indicate set of bounds of estimated treatment effects as in Oster (2019). # indicates set of bounds that excludes zero value

+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

Source: Authors' estimation.

The identified sets do not include zero value for the effect of change of ERP on total labor share changes and the reallocation term of labor share changes, while they include zero value for the remaining channels, namely within, entry, and exit. This result implies that our baseline effects are still robust against potential unobserved confounding effects. In particular, the confidence is high in accepting that there is an effect of change of ERP on labor share changes that works via reallocation channel. However, the effect from other channels remains doubtful, since it does not exclude zero value in their set of bounds.

In summary, various robustness checks confirm the effect of trade liberalization on labor share via reallocation channel. Thus, the role of trade liberalization should not be ignored in explaining the labor share decline in the manufacturing sector.

6. Concluding Remarks

Many studies have documented a declining trend in aggregate labor share of income across developing countries, including in their manufacturing sector. This paper finds that trade liberalization measured by reduction in ERP contributes to decline in the labor share in Indonesia's manufacturing sector. Specifically, trade liberalization induces superstar firms' effect whereby market share reallocates towards firms with low labor share. This trade-induced intra-industry reallocation channel outweighs other channels in driving the decline of aggregate labor share in the manufacturing sector. The trade effect remains significant across alternative model specifications and trade policy variables, and despite the inclusion of other competing factors that could potentially drive labor share downward. Our baseline estimate is robust from potential unobserved confounding factors, assuming equal proportionality assumption. The importance of the intra-industry reallocation channel also survives most of the alternative models. We further find that the labor share-reducing effect of trade liberalization is more substantial during the period of 1990 to 2000 and in the low technology sectors. Trade liberalization appears to be driving the decline of labor share across sectors organized within the networked trade and those that are not. The extent of exposure to changes in ERP seems to drive the heterogeneous results.

Despite the labor share-reducing effect, trade liberalization raises overall real wages and labor productivity across industries. However, this seems to have been achieved by using fewer workers and greater capital intensity, hence driving labor share downward. This paper confirms that low labor share firms tend to have higher productivity. Hence, reallocation into low labor share superstar firms following trade liberalization resembles the selection effect of trade along the lines of Melitz's heterogeneous firms' model. In this sense, the DLS trend can be seen as an inevitable outcome as trade selects the more productive firms in the industry. We find indicative support for this notion: reallocation into better-performing firms occurs during the period where strong trade's effect is detected, namely in 1990s. This result complements the superstar firms' model, which does not pinpoint a specific mechanism for reallocation channels.

This paper acknowledges the role of other factors in the decline of aggregate labor share in Indonesia's manufacturing sector. Two alternative explanations are assessed: the role of superstar firms and capital-biased technological change. We extend the main models to account for additional controls, such as aggregate markups, market power, and technological progress which is measured by TFP. We confirm that these variables are associated negatively with labor share as expected from the previous literature. However, it does not explain away the effect of trade liberalization. This points to a suggestion that the effect of trade liberalization does not simply pick up the effect of (omitted) confounding factors. Hence, the role of trade liberalization should not be ignored in an attempt to explain the DLS in Indonesia's manufacturing sector.

Despite the relatively stable estimates, the result of this study needs to be interpreted with caution. The SI dataset may not reflect the true firms' turnover as it could reflect attrition and new participation in the survey. However, the contribution from firms' exit and entry tends to be minor, making it less likely to change the overall story of this paper. It is also worth to note that this study only focuses on the manufacturing sector, not the overall economy. The interpretation is then best limited to the manufacturing sector. Trade could be an important factor in the dynamics of labor share in manufacturing sector, but maybe less so for the agriculture or services sector. In addition, readers might also notice that the decomposed trade effect is rather sensitive to the choice of tariff measures and technical coefficient used in ERP construction. Nevertheless, the overall negative effect of trade liberalization on aggregate labor share is still maintained by using various tariff measures and technical coefficients.

Lastly, the finding in this study does not imply that trade liberalization is harmful to labor. One of the most obvious reasons is that labor share is not an ideal indicator of the absolute welfare of the workers; instead, it only measures the relative income held by workers compared to other actors in the economy. The finding in this paper suggests that workers in the manufacturing sector obtain a smaller share of income over time compared to other actors and that trade liberalization plays its part in exacerbating income inequality across various factors of production.

However, it is doubtful that trade protection will work more favourably for workers' interests. Based on the heterogeneous firms' model, trade protection will likely allow less productive firms with larger labor share to stay in the industry; these firms would not have survived or at least shrunk under a liberalized setting. As a result, the aggregate industry's productivity will be lower in a protective trade environment. In addition, trade liberalization and greater access to the global market have been found to improve manufacturing firms' productivity (Amiti and Konings, 2007, Pane and Patunru, 2021). Therefore, the use of trade protection will potentially come at the expense of industries' and firms' productivity. Less productive industries and firms are unlikely to align with workers' interests as high-paying jobs tend to come from more productive industries and firms (Bernard et al., 2007). This finding implies that a more neutral policy, such as income support or other redistributive fiscal policies, might be more effective in maintaining workers' relative welfare with minimum side effects on the manufacturing sector's performance.

However, the nature of our dataset only allows for assessing the impact of liberalization on labor share across firms and industries, where all labor inputs, both low-skilled and high-skilled ones, are combined. Hence, it cannot answer how liberalization affects workers with different skill sets. It is possible that liberalization works more favourably for workers with certain skills, in particular the higher-skilled ones. Whether liberalization affect workers differently depending on their skill is, thus, a promising avenue for future research, especially in the context of developing economies. Finally, future study should include non-tariff barriers, as their use in trade policy has been increasingly recently, including in Indonesia.

References

Abeberese, A. B., Barnwal, P., Chaurey, R. & Mukherjee, P. 2021. Democracy and Firm Productivity: Evidence from Indonesia. *The Review of Economics and Statistics*, 1-30.

Ackerberg, D. A., Caves, K. & Frazer, G. 2015. Identification Properties of Recent Production Function Estimators. *Econometrica,* 83, 2411-2451.

Ahsan, R. N. & Mitra, D. 2014. Trade liberalization and labor's slice of the pie: Evidence from Indian firms. *Journal of Development Economics,* 108, 1-16.

Alatas, V. & Cameron, L. A. 2008. The Impact of Minimum Wages on Employment in a Low-Income Country: A Quasi-Natural Experiment in Indonesia. *ILR Review,* 61, 201-223.

Amiti, M. & Konings, J. 2007. Trade liberalization, intermediate inputs, and productivity: Evidence from Indonesia. *American Economic Review,* 97, 1611-1638.

Athukorala, P. 2006. Trade Policy Reforms and the Structure of Protection in Vietnam. *The World Economy,* 29, 161-187.

Athukorala, P. 2014. Global production sharing and trade patterns in East Asia. *In:* Kaur, Inderjit, and Nirvikar Singh (Eds), *The Oxford Handbook of the Economics of the Pacific Rim* , Oxford: Oxford University Press, 333-361.

Atkinson, A. B. 2009. Factor shares: the principal problem of political economy? *Oxford Review of Economic Policy,* 25, 3-16.

Autor, D., Dorn, D., Katz, L. F., Patterson, C. & Van Reenen, J. 2020. The Fall of the Labor Share and the Rise of Superstar Firms*. *The Quarterly Journal of Economics,* 135, 645-709.

Autor, D. H. 2018. Trade and labor markets: Lessons from China's rise. *IZA World of Labor*.

Autor, D. H., Dorn, D. & Hanson, G. H. 2013. The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review.*

Autor, D. H., Dorn, D. & Hanson, G. H. 2016. The China Shock: Learning from Labor-Market Adjustment to Large Changes in Trade. *Annual Review of Economics,* 8**,** 205-240.

Balassa, B. A. 1971. *The structure of protection in developing countries,* Maryland, The Johns Hopkins Press.

Berkowitz, D., Ma, H. & Nishioka, S. 2015. Declining labour shares and heterogeneous firms. Working Paper 552. Stanford Centre for International Development, Stanford, US.

Bernard, A. B., Jensen, J. B., Redding, S. J. & Schott, P. K. 2007. Firms in International Trade. *The Journal of Economic Perspectives,* 21**,** 3-130.

Blalock, G. & Gertler, P. J. 2004. Learning from exporting revisited in a less developed setting. *Journal of Development Economics,* 75**,** 397-416.

Böckerman, P. & Maliranta, M. 2012. Globalization, creative destruction, and labour share change: evidence on the determinants and mechanisms from longitudinal plant-level data. *Oxford Economic Papers,* 64**,** 259-280.

Brandt, L., Van Biesebroeck, J., Wang, L. & Zhang, Y. 2017. WTO Accession and Performance of Chinese Manufacturing Firms. *American Economic Review,* 107**,** 2784-2820.

Brownstone, D. & Valletta, R. G. 1996. Modeling Earnings Measurement Error: A Multiple Imputation Approach. *The Review of Economics and Statistics,* 78**,** 705.

Corden, W. M. 1966. The Structure of a Tariff System and the Effective Protective Rate. *Journal of Political Economy,* 74**,** 221-237.

Dao, M. C., Das, M. & Koczan, Z. 2019. Why is labour receiving a smaller share of global income? *Economic Policy,* 34**,** 723-759.

De Loecker, J., Eeckhout, J. & Unger, G. 2020. The Rise of Market Power and the Macroeconomic Implications*. *The Quarterly Journal of Economics,* 135**,** 561-644.

De Loecker, J. & Warzynski, F. 2012. Markups and Firm-Level Export Status. *American Economic Review,* 102, 2437-2471.

Dix-Carneiro, R. & Kovak, B. K. 2019. Margins of labor market adjustment to trade. *Journal of International Economics.*

Eddings, W. & Marchenko, Y. 2012. Diagnostics for multiple imputation in Stata. *Stata Journal,* 12, 353-367.

Elsby, M. W., Hobijn, B. & Şahin, A. 2013. The decline of the US labor share. *Brookings Papers on Economic Activity,* 2013, 1-63.

Enders, C. K. 2010. Applied missing data analysis. New York: Guilford Press.

Feenstra, R. C. & Hanson, G. H. 2003. Global Production Sharing and Rising Inequality: A Survey of Trade and Wages. *In:* CHOI, E. K. & HARRIGAN, J. (eds.). Oxford, UK: Blackwell Publishing Ltd.

Fernandes, A. M., Kee, H. L. & Winkler, D. 2022. Determinants of Global Value Chain Participation: Cross-Country Evidence. *The World Bank Economic Review,* 36, 329-360.

Gandhi, A., Navarro, S. & Rivers, D. A. 2020. On the Identification of Gross Output Production Functions. *Journal of Political Economy,* 128, 2973-3016.

Goldberg, P. K. & Pavcnik, N. 2007. Distributional Effects of Globalization in Developing Countries. *Journal of Economic Literature,* 45, 39-82.

Gradzewicz, M. & Mućk, J. 2019. *Globalization and the fall of markups,* Narodowy Bank Polski.

Greenaway, D. & Milner, C. 2003. Effective Protection, Policy Appraisal and Trade Policy Reform. *The World Economy,* 26, 441-456.

Grether, J.-M. & Tissot-Daguette, B. 2021. Multiple imputation techniques: An application to Swiss value-added data. IRENE Institute of Economic Research.

Grossman, G. M. & Helpman, E. 1994. Protection for Sale. *The American Economic Review,* 84**,** 833-850.

Grossman, G. M. & Oberfield, E. 2022. The Elusive Explanation for the Declining Labor Share. *Annual Review of Economics,* 14**,** 93-124.

Hollenstein, H. & Woerter, M. 2008. Inter- and intra-firm diffusion of technology: The example of E-commerce: An analysis based on Swiss firm-level data. *Research Policy,* 37**,** 545-564.

Hopenhayn, H., Neira, J. & Singhania, R. 2018. From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share. *National Bureau of Economic Research Working Paper Series,* No. 25382.

Impullitti, G. & Licandro, O. 2017. Trade, Firm Selection and Innovation: The Competition Channel. *The Economic Journal,* 128**,** 189-229.

Kamal, F., Lovely, M. E. & Mitra, D. 2019. Trade liberalisation and labour shares in China. *The World Economy,* 42**,** 3588-3618.

Karabarbounis, L. & Neiman, B. 2014. The Global Decline of The Labor Share. *The Quarterly Journal of Economics,* 129**,** 61-104.

Kis-Katos, K. & Sparrow, R. 2015. Poverty, labor markets and trade liberalization in Indonesia. *Journal of Development Economics.*

Kofman, P. 2003. Using Multiple Imputation in the Analysis of Incomplete Observations in Finance. *Journal of Financial Econometrics,* 1**,** 216-249.

Krugman, P. R., Obstfeld, M. & Melitz, M. J. 2018. *International economics: theory & policy,* Harlow, Essex, Pearson.

Leblebicioğlu, A. & Weinberger, A. 2021. Openness and factor shares: Is globalization always bad for labor? *Journal of International Economics,* 128**,** 103406.

Levinsohn, J. & Petrin, A. 2003. Estimating Production Functions Using Inputs to Control for Unobservables. *The Review of Economic Studies,* 70**,** 317-341.

Lileeva, A. & Trefler, D. 2010. Improved Access to Foreign Markets Raises Plant-level Productivity…For Some Plants. *The Quarterly Journal of Economics,* 125**,** 1051-1099.

Manjón, M. & Mañez, J. 2016. Production Function Estimation in Stata Using the Ackerberg–Caves–Frazer Method. *The Stata Journal: Promoting communications on statistics and Stata,* 16**,** 900-916.

Melitz, M. J. 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica,* 71**,** 1695-1725.

Melitz, M. J. & Ottaviano, G. I. P. 2008. Market Size, Trade, and Productivity. *Review of Economic Studies,* 75**,** 295-316.

Melitz, M. J. & Polanec, S. 2015. Dynamic Olley-Pakes productivity decomposition with entry and exit. *The RAND Journal of Economics,* 46**,** 362-375.

Montfaucon, A.F., Cali, M., Agnimaruto, B., Silbering, J.M., Hasmand, A.A., Lakatos, C. & Pasha, M. 2023. *Building a Dataset for Non-tariff Measures and Its Usage: The Case of Indonesia and Applicability for Other Countries*. World Bank: Washington, DC.

Navas-Ruiz, A. & Sala, D. 2007. Technology Adoption and the Selection Effect of Trade. *EUI ECO Working Papers,* 58.

Olley, G. S. & Pakes, A. 1996. The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica,* 64**,** 1263-1297.

Oster, E. 2019. Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics,* 37**,** 187-204.

Pane, D. D. & Patunru, A. A. 2021. Does Export Experience Improve Firms' Productivity? Evidence from Indonesia. *The Journal of Development Studies***,** 1-21.

Panon, L. 2022. Labor share, foreign demand and superstar exporters. *Journal of International Economics,* 139, 1–27.

Pavcnik, N. 2017. The Impact of Trade on Inequality in Developing Countries. *National Bureau of Economic Research Working Paper Series,* No. 23878.

Perugini, C., Vecchi, M. & Venturini, F. 2017. Globalisation and the decline of the labour share: A microeconomic perspective. *Economic Systems,* 41**,** 524-536.

Petreski, M. 2021. Has globalization shrunk manufacturing labor share in transition economies? *Journal of Comparative Economics,* 49**,** 201-211.

Petrin, A. & Levinsohn, J. 2012. Measuring aggregate productivity growth using plant-level data. *The RAND Journal of Economics,* 43**,** 705-725.

Putra, C. T. 2021. *Essays on firm diversification strategies: Evidence from Indonesian medium and large manufacturing plants.*

Rovigatti, G. & Mollisi, V. 2018. Theory and practice of total-factor productivity estimation: The control function approach using Stata. *Stata Journal,* 18**,** 618-662.

Saad, A. F. 2017. Trade and technology adoption. *The Journal of International Trade & Economic Development,* 26**,** 1-24.

Schafer, J. L. & Graham, J. W. 2002. Missing data: our view of the state of the art. *Psychological methods,* 7**,** 147-177.

Topalova, P. 2010. Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India. *American Economic Journal: Applied Economics.*

Vadila, Y. & Resosudarmo, B. P. 2020. Tariff reform and income inequality in Indonesia. *Regional Science Policy & Practice.* John Wiley & Sons, Ltd.

White, T. K., Reiter, J. P. & Petrin, A. 2018. Imputation in U.S. Manufacturing Data and Its Implications for Productivity Dispersion. *The Review of Economics and Statistics,* 100**,** 502-509.

Wing, C., Simon, K. & Bello-Gomez, R. A. 2018. Designing Difference in Difference Studies: Best Practices for Public Health Policy Research. *Annual review of public health.* United States: Annual Reviews.

Xu, X., Li, D. D. & Zhao, M. 2018. “Made in China” matters: Integration of the global labor market and the global labor share decline. *China Economic Review,* 52**,** 16-29.

Appendix A. Supporting tables and figures

Table A1. ISIC revision 2 industry definition

4-digit ISIC rev 2	Description	4-digit ISIC rev 2	Description
3111	Slaughtering, preparing and preserving meat	3523	Manufacture of soap and cleaning preparations, perfumes, cosmetics and other toilet preparations
3112	Manufacture of dairy products	3529	Manufacture of chemical products not elsewhere classified
3113	Canning and preserving of fruits and vegetables	3540	Manufacture of miscellaneous products of petroleum and coal
3114	Canning, preserving and processing of fish, crustaces and similar foods	3551	Tyre and tube industries
3115	Manufacture of vegetable and animal oils and fats	3559	Manufacture of rubber products not elsewhere classified
3116	Grain mill products	3560	Manufacture of plastic products not elsewhere classified
3117	Manufacture of bakery products	3610	Manufacture of pottery, china and earthenware
3118	Sugar factories and refineries	3620	Manufacture of glass and glass products
3119	Manufacture of cocoa, chocolate and sugar confectionery	3691	Manufacture of structural clay products
3121	Manufacture of food products not elsewhere classified	3692	Manufacture of cement, lime and plaster

3122	Manufacture of prepared animal feeds	3699	Manufacture of non-metallic mineral products not elsewhere classified
3131	Distilling, rectifying and blending spirits	3710	Iron and steel basic industries
3132	Wine industries	3720	Non-ferrous metal basic industries
3133	Malt liquors and malt	3811	Manufacture of cutlery, hand tools and general hardware
3134	Soft drinks and carbonated waters industries	3812	Manufacture of furniture and fixtures primarily of metal
3140	Tobacco manufactures	3813	Manufacture of structural metal products
3211	Spinning, weaving and finishing textiles	3819	Manufacture of fabricated metal products except machinery and equipment not elsewhere classified
3212	Manufacture of made-up textile goods except wearing apparel	3821	Manufacture of engines and turbines
3213	Knitting mills	3822	Manufacture of agricultural machinery and equipment
3214	Manufacture of carpets and rugs	3823	Manufacture of metal and wood working machinery
3215	Cordage, rope and twine industries	3824	Manufacture of special industrial machinery and equipment except metal and wood working machinery
3219	Manufacture of textiles not elsewhere classified	3825	Manufacture of office, computing and accounting machinery
3220	Manufacture of wearing apparel, except footwear	3829	Machinery and equipment except electrical not elsewhere classified
3231	Tanneries and leather finishing	3831	Manufacture of electrical industrial machinery and apparatus
3233	Manufacture of products of leather and leather substitutes, except footwear and wearing apparel	3832	Manufacture of radio, television and communication equipment and apparatus
3240	Manufacture of footwear, except vulcanized or moulded rubber or plastic footwear	3833	Manufacture of electrical appliances and housewares
3311	Sawmills, planing and other wood mills	3839	Manufacture of electrical apparatus and supplies not elsewhere classified
3312	Manufacture of wooden and cane containers and small cane ware	3841	Ship building and repairing
3319	Manufacture of wood and cork products not elsewhere classified	3842	Manufacture of railroad equipment
3320	Manufacture of furniture and fixtures, except primarily of metal	3843	Manufacture of motor vehicles

3411	Manufacture of pulp, paper and paperboard	3844	Manufacture of motorcycles and bicycles
3412	Manufacture of containers and boxes of paper and paperboard	3845	Manufacture of aircraft
3419	Manufacture of pulp, paper and paperboard articles not elsewhere classified	3851	Manufacture of professional and scientific, and measuring and controlling equipment not elsewhere classified
3420	Printing, publishing and allied industries	3852	Manufacture of photographic and optical goods
3511	Manufacture of basic industrial chemicals except fertilizers	3853	Manufacture of watches and clocks
3512	Manufacture of fertilizers and pesticides	3901	Manufacture of jewellery and related articles
3513	Manufacture of synthetic resins, plastic materials and man-made fibres except glass	3902	Manufacture of musical instruments
3521	Manufacture of paints, varnishes and laquers	3903	Manufacture of sporting and athletic goods
3522	Manufacture of drugs and medicines	3909	Manufacturing industries not elsewhere classified

Source: UN Stats

Table A2. Description of variables used in the analysis

Variables	Description
Value added (Rp million, constant 2000)	Output value subtracted by all expenses other than labor cost, land rent, tax, interest payment and gifts
Raw material (total) expenses (Rp million, constant 2000)	Value of all raw materials used in the production process, both imported and domestically sourced
Fuel expenses (Rp million, constant 2000)	Value of all fuel and lubricants used in the production process including gasoline, diesel, kerosene, coal, gas, lubricants, and other fuels
Electricity expenses (Rp million, constant 2000)	Value of all electricity purchased from PLN and non-PLN
Auxiliary expenses (Rp million, constant 2000)	Value of all other expenses such as rent, tax, industrial services, interest payment, royalty, and others
Total number of workers: All type	Number of all type of workers
Number of workers: non-production (skilled)	Number of non-production workers
Foreign ownership (%)	Percentage of capital owned by foreign entities
Export status (1 = exporter, 0 = non-exporter)	Does the firm export a product?

Capital stock, original & imputed (Rp million, constant 2000)	Estimated value of fixed assets at the end of the year, including land, building, machinery & equipment, vehicles, and others (combined estimation of five imputed datasets using Rubin's rule)
Labor share: wage or salary relative to value added	Value of Wage or salary payment to workers relative to value added
Labor share: total compensation relative to value added	Value of all compensation (including overtime, bonus, pension and other social security contribution, and accident allowance)
Markups - LP	Markups estimated by production function approach that use LP technique
Markups - ACF Cobb - Douglas	Markups estimated by production function approach that use ACF Cobb-Douglas technique
Markups - Simple accounting	Markups estimated by simple accounting technique: value of goods produced relative to total variable cost: workers' and materials' expenses
TFP – LP	TFP estimated by LP technique
TFP - ACF Cobb - Douglas	TFP estimated by ACF Cobb-Douglas technique
Capacity utilization	Percentage of Realized Production relative to Installed Capacity
Effective Rate of Protection (ERP), input weight: average 2000-2015	ERP calculated by Corden's formula that use fixed input coefficient, averaging 2000-2015
Nominal Rate of Protection (NRP): output tariff including ad-valorem estimation	Output tariff that include also ad-valorem equivalent of specific tariffs
Input tariff, input expenses to value added weight: average 2000-2015	Input tariff of a given (output) industry, using fixed input coefficient averaging 2000-2015

Source: SI Questionnaires and various sources

Table A3. Statistical support for the exogeneity assumption of ERP

	Dep: ERP deviation		Dep: change in ERP deviation	
	(1)	(2)	(3)	(4)
TFP (log) (Φ_{jt})	6.708	19.97		
	[8.653]	[21.54]		
Share of exporters ($EXPORT_{jt}$)	51.43	-3.479		
	[52.08]	[20.21]		
Share of skilled workers ($SKILL_{jt}$)	181.6	9.062		
	[193.3]	[44.23]		
Capital per worker (log) ($KINT_{jt}$)	-8.388	3.835		
	[9.592]	[4.180]		

Average firm's age ($AAGE_{jt}$)	-207.4	-2.276		
	[218.8]	[21.13]		
Average firm's capacity utilization	0.302	-0.0719		
($UTIL_{jt}$)	[0.280]	[0.219]		
Auxiliary expenditure share to value	5.020	1.058		
added ($OIEXPSH_{jt}$)	[7.175]	[3.704]		
Average firm's FDI share (DFO_{jt})	30.83	-37.20		
	[50.74]	[40.15]		
Initial share of exporters ($EXPORT_j{}^{1990}$)		-20.74		10.13
		[36.08]		[18.64]
Initial share of skilled workers		-124.8		53.48
($SKILL_j{}^{1990}$)		[135.4]		[53.20]
Initial capital intensity (log) ($KINT_j{}^{1990}$)		-2.658		0.908
		[5.856]		[2.939]
Initial average firm's capacity		-0.327		0.318
Utilization ($UTIL_j{}^{1990}$)		[0.666]		[0.442]
Initial auxiliary expenditure share to		-10.21		9.154
value added ($OIEXPSH_j{}^{1990}$)		[16.94]		[10.11]
Initial average firm's FDI share		-34.34		23.73
($DFO_j{}^{1990}$)		[42.83]		[27.38]
TFP growth ($\Delta\Phi_{jt}$)			-0.0149	-0.0264
			[0.0144]	[0.0247]
Change in share of exporters			0.0518	0.0623
($\Delta EXPORT_{jt}$)			[0.0693]	[0.0793]
Change in share of skilled workers			-0.0683	-0.0119
($\Delta SKILL_{jt}$)			[0.127]	[0.0997]
Capital per worker growth ($\Delta KINT_{jt}$)			-0.290	-0.404
			[0.597]	[0.669]
Average firm's age growth ($\Delta AAGE_{jt}$)			0.109	-0.0313
			[0.0903]	[0.0723]
Change in capacity utilization ($\Delta UTIL_{jt}$)			-0.274	-0.259
			[0.289]	[0.275]
Change in auxiliary expenditure share to			-0.0268	-0.0217
value added ($\Delta OIEXPSH_{jt}$)			[0.0271]	[0.0238]
Change in average firm's FDI share			-0.00434	-0.0108
(ΔDFO_{jt})			[0.0822]	[0.0978]
Observations	2027	2027	1948	1948
R^2	0.017	0.015	0.013	0.016

Note: ERP deviation is defined as ERP minus output tariff, while the change in ERP deviation reflects the change of that deviation over time. All estimations are conducted at 4-digit ISIC industry level without any weight. Standard errors are in brackets and clustered at the industry level. The following sign shows significance level: + p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001.

Source: Authors' estimation

Table A4. Does treatment variable of interest correlate with observables and industry performance?

	(1) Changes of ERP	(2) Changes of output tariff
TFP growth ($\Delta\Phi_{jt}$)	-0.0309	-0.0159^{+}
	[0.0200]	[0.00867]
Change in share of exporters ($\Delta EXPORT_{jt}$)	0.0751	0.0233
	[0.0874]	[0.0225]
Change in share of skilled workers ($\Delta SKILL_{jt}$)	-0.0711	-0.00277
	[0.174]	[0.0642]
Capital per worker growth ($\Delta KINT_{jt}$)	-0.178	0.112
	[0.798]	[0.298]
Average firm's age growth ($\Delta AAGE_{jt}$)	0.225^{+}	0.117*
	[0.119]	[0.0521]
Change in capacity utilization ($\Delta UTIL_{jt}$)	-0.376	-0.102
	[0.353]	[0.0809]
Change in auxiliary expenditure share to value added ($\Delta OIEXPSH_{jt}$)	-0.0399	-0.0131
	[0.0333]	[0.00812]
Change in average firm's FDI share (ΔDFO_{jt})	0.0101	0.0145
	[0.103]	[0.0253]
Observations	1948	1948
R^2	0.013	0.022

Note: all estimations are based on TWFE regression at 4-digit industry categories without any weight. Standard errors are in brackets and clustered at the industry level. The following sign shows significance level: + p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001.

Source: Authors' estimation.

Table A5. The statistical difference between ERP and output tariff, paired t-test

	Observations	Mean 1: ERP	Mean 2: output tariff	a two-sided p-value of the two-sample t-test
Level	2028	32.367	18.611	0.0792
Change	1950	-8.127	-1.757	0.0895

Source: Authors' estimation.

Table A6. Correlation between treatment variable, covariates, and TFP

	(1) Change in ERP	(2) Change in TFP	(3) Change in exporter share	(4) Change in skilled worker share	(5) Change in capital intensity	(6) Change in average age of firms	(7) Change in capacity utilization	(8) Change in auxiliary expenditure share to value added	(9) Change in FDI-firm share
(1)	1								
(2)	-0.0118	1							
(3)	-0.00196	0.0176	1						
(4)	-0.00191	-0.0164	-0.0788***	1					
(5)	-0.00553	0.0453*	-0.0457*	0.00228	1				
(6)	-0.0695**	0.0798***	0.0414	0.0627**	0.0384	1			
(7)	-0.0179	-0.0112	0.00322	-0.0284	0.0539*	-0.0451*	1		
(8)	-0.00644	0.0176	-0.0680**	0.0948***	0.0585**	-0.0157	-0.0637**	1	
(9)	-0.00254	0.0898***	0.0721**	-0.161***	-0.0388	0.0312	-0.0641**	-0.145***	1

Note: The following sign shows significance level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Source: Authors' estimation

Table A7. Weighted-average of DOPD of aggregate labor share change at 4-digit ISIC Rev 2

	Stayers' reallocation (1)	Unweighted average change for stayers (within) (2)	Firm entry (3)	Firm exit (4)	Total change (5)
A. Initial period (year 1990) value-added share as weight					
1990-1995	-3.65	4.73	-0.84	-0.08	0.15
1995-2000	0.82	0.05	-0.75	-0.04	0.08
2000-2005	1.97	-0.40	0.01	0.20	1.78
2005-2010	3.64	-7.98	-0.06	-0.13	-4.53
2010-2015	3.31	-4.76	-0.49	0.20	-1.74
1990-2000 (cumulative)	-2.83	4.78	-1.60	-0.13	0.23
2000-2015 (cumulative)	8.92	-13.13	-0.55	0.27	-4.49
1990-2015 (cumulative)	6.09	-8.35	-2.14	0.14	-4.25
B. Annual value-added share as weight					
1990-1995	-5.05	4.29	-0.96	-0.04	-1.75
1995-2000	-1.08	-0.59	-0.83	0.19	-2.30
2000-2005	1.44	-0.51	-0.09	0.23	1.07
2005-2010	4.44	-7.69	-0.36	-0.17	-3.78
2010-2015	2.23	-4.19	-0.28	0.05	-2.18
1990-2000 (cumulative)	-6.13	3.70	-1.78	0.16	-4.05
2000-2015 (cumulative)	8.12	-12.39	-0.73	0.11	-4.89
1990-2015 (cumulative)	1.99	-8.68	-2.52	0.26	-8.94

Note: Aggregate labor share is calculated as the weighted average of firms' labor share, where each firm's value added in manufacturing sector is used as weight. Decomposition of labor share changes is conducted in a five-year interval, meaning that the analysis uses information from the start and the end of the period while ignoring variations in between. Therefore, survivors, entrants, and exit-ers are all categorized within that five-year interval. The decomposition is applied for each industry at a 4-digit level of ISIC rev-2 and then aggregated using the industry's value-added. All change is measured in percentage point changes unless specified otherwise. Cumulatively, the reallocation channel plays a significant role during 1990-2000, while the role of the within-firm channel gets more substantial during 2000-2015 in explaining the decline in aggregate labor share.

Source: Authors' estimation.

Table A8. The full result of the effect of ERP on aggregate labor share

	(1) Total change	(2) Stayers' reallocation	(3) Unweighted average change for stayers (within)	(4) Firm entry	(5) Firm exit
Changes of ERP (ΔERP_{jt})	0.000517**	0.000758*	-0.000190	-0.000181***	0.000141**
	[0.000164]	[0.000291]	[0.000357]	[0.0000509]	[0.0000416]
Change in share of exporters ($\Delta EXPORT_{jt}$)	0.0125	-0.0394	0.0868	-0.0271	-0.0129
	[0.0789]	[0.0871]	[0.0946]	[0.0280]	[0.0251]
Change in share of skilled workers ($\Delta SKILL_{jt}$)	0.0545	0.129	-0.0470	-0.0278	0.000565
	[0.111]	[0.138]	[0.161]	[0.0260]	[0.0197]
Capital per worker growth ($\Delta KINT_{jt}$)	-0.860+	3.817+	-4.256+	-0.110	-0.305*
	[0.454]	[2.187]	[2.350]	[0.141]	[0.139]
Average firm's age growth ($\Delta AAGE_{jt}$)	-0.115	0.102	-0.258	0.0202*	0.00804
	[0.0819]	[0.122]	[0.156]	[0.0100]	[0.0185]
Change in capacity utilization ($\Delta UTIL_{jt}$)	0.0314	-0.0682	0.109	-0.00275	-0.00382
	[0.0744]	[0.0945]	[0.0917]	[0.0221]	[0.0297]
Change in auxiliary expenditure share to value added ($\Delta OIEXPSH_{jt}$)	0.117***	0.0840*	0.0349	0.00121	0.000288
	[0.0312]	[0.0372]	[0.0490]	[0.00748]	[0.00663]
Change in average firm's FDI share (ΔDFO_{jt})	0.0275	-0.270+	0.248	-0.0462	0.0641
	[0.0632]	[0.141]	[0.194]	[0.0320]	[0.0816]
Observations	1948	1948	1948	1879	1865
R^2	0.068	0.308	0.312	0.017	0.013

Note: The table shows the estimated coefficient of change of ERP on change of aggregate labor share as well as its components. All estimations are based on TWFE regression at 4-digit industry categories, where each observation is weighted by its value-added share at the initial observation period. Standard errors are in brackets and clustered at the industry level. The following sign shows significance level: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Source: Authors' estimation

Table A9. Robustness check: by components

	(1) Total change	(2) Stayers' reallocation	(3) Unweighted average change for stayers (within)	(4) Firm entry	(5) Firm exit
Baseline	0.000517** [0.000164]	0.000758* [0.000291]	-0.000190 [0.000357]	-0.000181*** [0.0000509]	0.000141** [0.0000416]
With industry-specific time trend	0.000850*** [0.000231]	0.00136*** [0.000380]	-0.000319 [0.000493]	-0.000329*** [0.0000614]	0.000173*** [0.0000489]
With year-industry (2 digit) FE	0.000596*** [0.000163]	0.000950*** [0.000195]	-0.000375+ [0.000213]	0.0000280 [0.000187]	0.00000896 [0.000105]
Total compensation	0.000997*** [0.000277]	0.00115** [0.000347]	-0.0000428 [0.000481]	-0.000312*** [0.0000814]	0.000230*** [0.0000643]
OLS	0.000505*** [0.000139]	0.000893*** [0.000255]	-0.000424 [0.000308]	-0.000156*** [0.0000442]	0.000192*** [0.0000366]
OLS with initial controls	0.000520*** [0.000138]	0.000878*** [0.000249]	-0.000411 [0.000306]	-0.000121* [0.0000464]	0.000178*** [0.0000435]
ERP with technical coefficient of 2006	0.00113* [0.000511]	0.00116 [0.000743]	-0.00000360 [0.000987]	-0.000276* [0.000119]	0.000274** [0.0000975]
Original tariff data	0.0605** [0.0207]	-0.0418 [0.0491]	0.0946+ [0.0502]	0.00142 [0.00641]	0.00550 [0.00640]
Ad-valorem tariff data only	0.0343+ [0.0183]	-0.0479 [0.0452]	0.0772 [0.0476]	0.00373 [0.00480]	0.00261 [0.00403]

Note: The table shows the estimated coefficient of change of ERP on change of aggregate labor share as well as its components. All estimations are based on TWFE regression at 4-digit industry categories, where each observation is weighted by its value-added share at the initial observation period. Standard errors are in brackets and clustered at the industry level. The following sign shows significance level: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Source: Authors' estimation.

Table A10. The Effect of ERP on aggregate labor share with additional controls

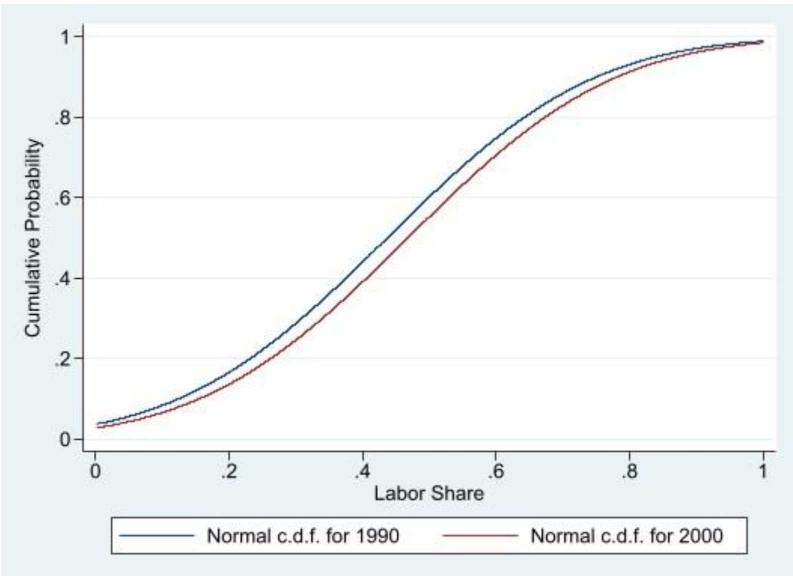
	(1) Controlling for productivity	(2) Controlling for markups	(3) Controlling for market power	(4) Controlling for productivity, markups, and market power
Changes of ERP (ΔERP_{jt})	0.000405**	0.000675***	0.000366*	0.000599**
	[0.000121]	[0.000161]	[0.000176]	[0.000185]
Change in share of exporters ($\Delta EXPORT_{jt}$)	0.0133	-0.00415	0.0237	0.00403
	[0.0731]	[0.0770]	[0.0764]	[0.0776]
Change in share of skilled workers ($\Delta SKILL_{jt}$)	0.0528	0.0542	0.0778	0.0751
	[0.114]	[0.108]	[0.110]	[0.107]
Capital per worker growth ($\Delta KINT_{jt}$)	-0.746	-0.375	-0.810+	-0.338
	[0.485]	[0.563]	[0.434]	[0.543]
Average firm's age growth ($\Delta AAGE_{jt}$)	-0.119	-0.0954	-0.110	-0.0881
	[0.0845]	[0.0872]	[0.0802]	[0.0846]
Change in capacity utilization ($\Delta UTIL_{jt}$)	0.0666	0.0624	0.0323	0.0520
	[0.0804]	[0.0777]	[0.0721]	[0.0796]
Change in auxiliary expenditure share to value added ($\Delta OIEXPSH_{jt}$)	0.115***	0.132***	0.0940**	0.115***
	[0.0296]	[0.0292]	[0.0331]	[0.0320]
Change in average firm's FDI share (ΔDFO_{jt})	0.0229	0.0350	-0.0127	0.00242
	[0.0552]	[0.0711]	[0.0682]	[0.0795]
Total change in the TFP ($\Delta\Phi_{jt}$)	-0.0594**			0.0229
	[0.0212]			[0.0140]
Total change in the (log) markup ($\Delta \mathrm{M}_{jt}$)		-0.0651***		-0.0701***
		[0.0168]		[0.0155]

Change in share of market power ($\Delta CR4_{jt}$)			-0.140*** [0.0285]	-0.122*** [0.0304]
Observations	1948	1948	1948	1948
R^2	0.095	0.155	0.089	0.169

Note: The table shows the estimated coefficient of change of ERP on change of aggregate labor share as well as its components. All estimations are based on TWFE regression at 4-digit industry categories, where each observation is weighted by its value-added share at the initial observation period. Standard errors are in brackets and clustered at the industry level. The following sign shows significance level: + p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001.

Source: Authors' estimation.

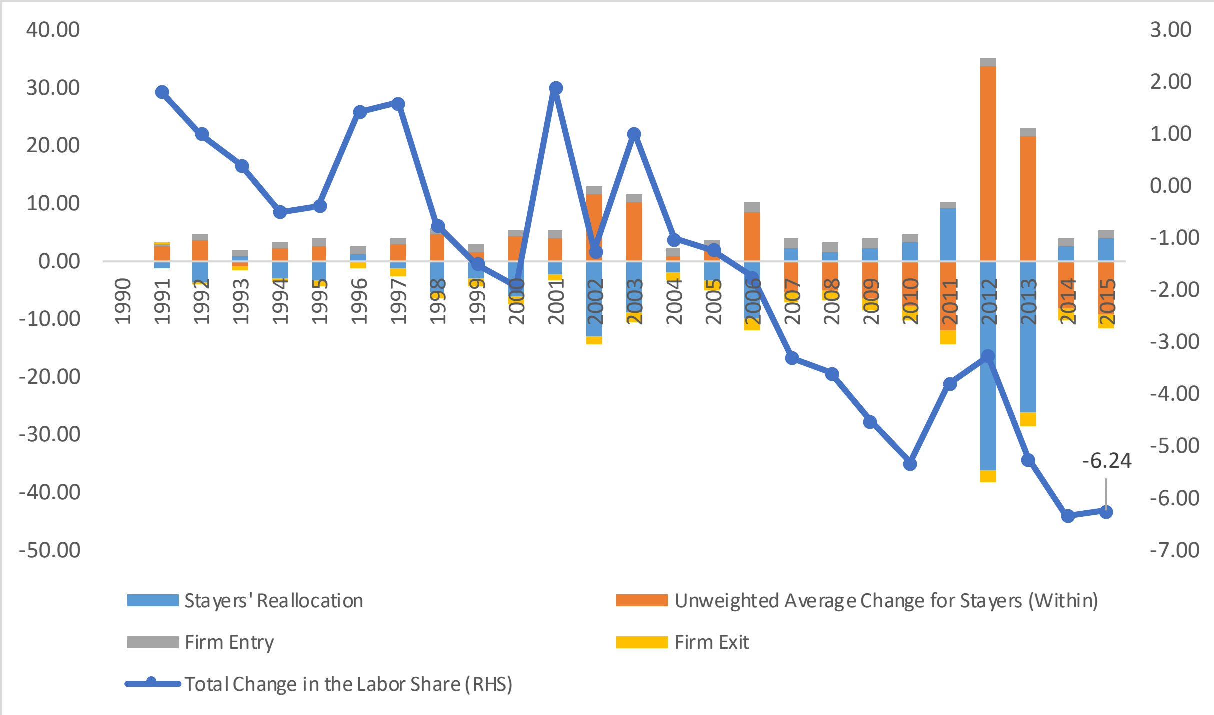
Figure A1. Firms' cumulative distribution of labor share



Note: The observations with labor share of more than 100% are excluded. The total sample used is 96% of the original one.

Source: Author's estimation

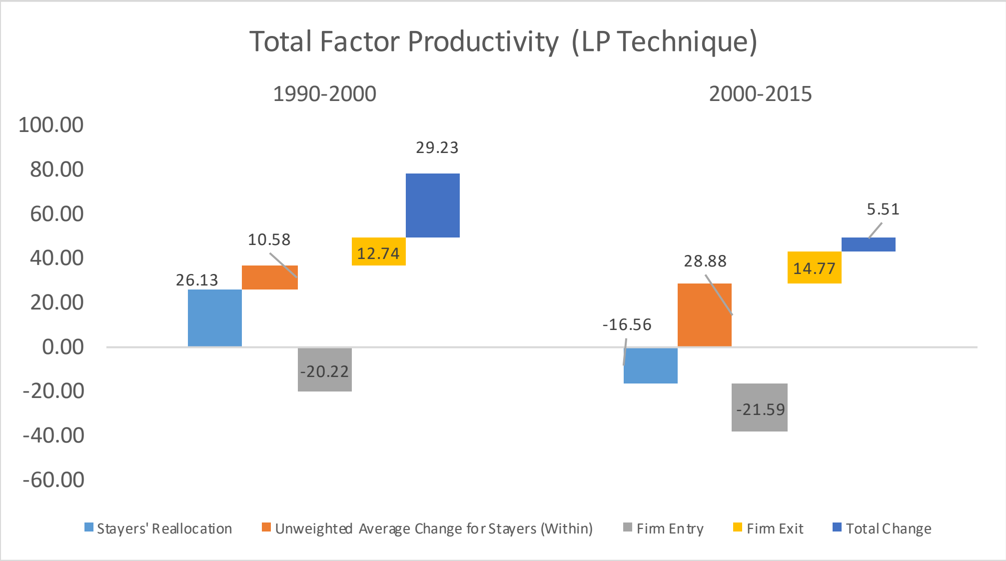
Figure A2. Annual DOPD of labor share change in manufacturing sector (in percentage point), cumulative from 1990



Note: aggregate labor share is calculated as the weighted average of firms' labor share, where each firm's value-added in the manufacturing sector is used as weight. Decomposition of labor share changes is conducted on an annual basis. All changes are measured in percentage point changes.

Source: Authors' estimation

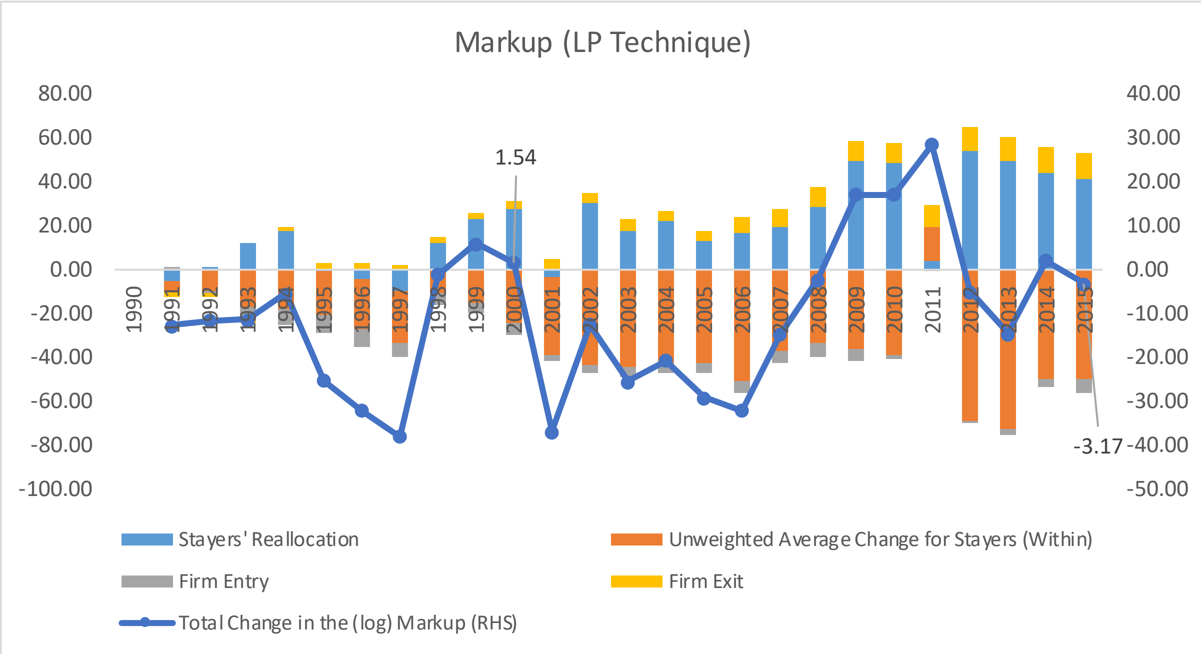
Figure A3. Aggregate TFP decomposition (in percentage change)



Note: Aggregate TFP is calculated as the weighted average of firms' TFP (in log term), where each firm's value added in manufacturing sector is used as weight. Decomposition of TFP changes is conducted in a five-year interval. It means the analysis uses information from the start and the end of the period while ignoring variations in between. Therefore, survivors, entrants, and exit-ers are all categorized within that five-year interval. TFP is recovered after estimating production function coefficient using LP technique. The period split accumulates the value derived from the five-year interval decomposition. For example, it means that the value of 1990-2000 represents the sum over 1990-1995 and 1995-2000 periods.

Source: Authors' estimation.

Figure A4. Annual decomposition of (log) markup changes (% change), cumulative from 1990



Note: Aggregate markup is calculated as the weighted average of firms' (log) markup, where each firm's value-added in the manufacturing sector is used as weight. Decomposition of markup changes is conducted on an annual basis. Markup is recovered after estimating production function coefficient using LP technique. All changes are measured in percentage changes.

Source: Authors' estimation.

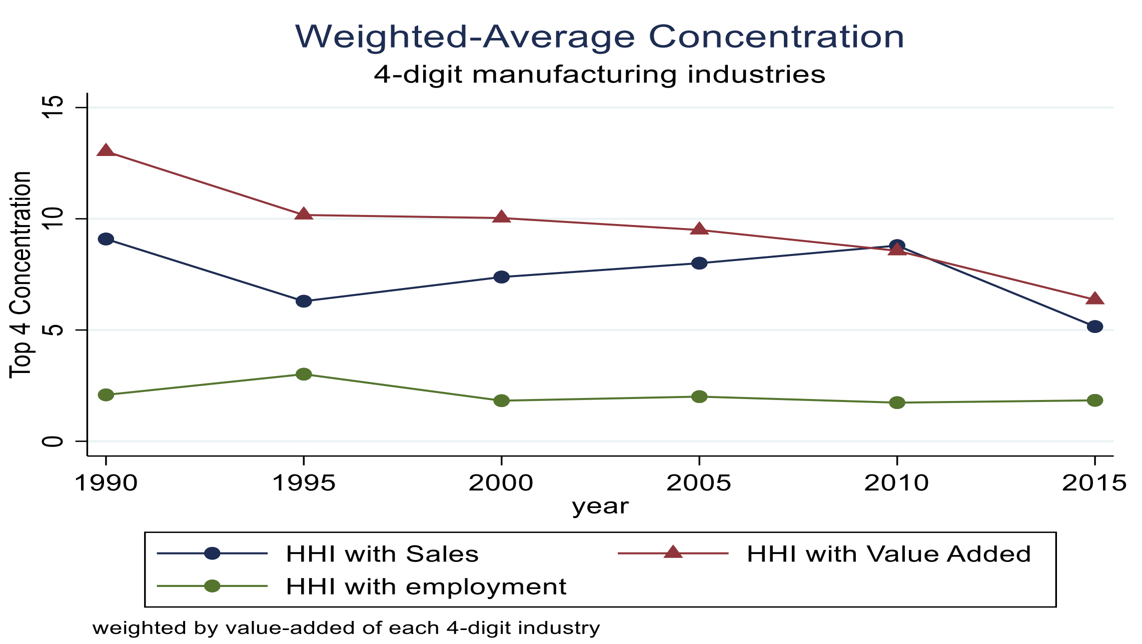


Figure A5. The Herfindahl-Hirschman Index (HHI), all manufacturing industries

Note: The HHI is first calculated at 4-digit ISIC level and then aggregated up to the overall manufacturing sector using each industry's share of value-added.

Source: Authors' estimation

Appendix B. Imputation of capital stock data

As noted in the main text, the original manufacturing dataset suffers from missing capital stock data. However, these problems are not uncommon among micro-level survey data like the SI database. Studies that use survey-based microdata, such as household income and firm data, have faced similar challenges. The dominant approach in the economics literature to deal with that problem is the deletion method, where analysis only focuses on the complete observations (Kofman, 2003). Only a few studies acknowledge and further treat the data reporting problems. Those that decided to do so use some variants of imputation method. Previous studies that use the SI database implement a simple imputation method based on the interpolation of the observed data (Blalock and Gertler, 2004, Amiti and Konings, 2007, Pane and Patunru, 2021). Other studies employ the mean imputation technique, where the missing data point is imputed by its average industry value or the fitted value of univariate regression. White et al. (2018)

highlighted that almost half of the observations for the key production function variables in the US Census of Manufactures data are imputed using mean value. Some highly cited studies in economics have used the US Census manufacturing data, including Autor et al. (2020).

This paper uses the multiple imputation method instead of the simple one. The rationale is as follows. First, the pattern of the missing cases for capital stock data fits the missing at random (MAR) assumption. After treating the likely measurement error of negative and zero capital stock data as missing cases[26], it appears that the missingness in capital stock data is associated with several observables. This result implies that the missing capital stock data is not fully random and that the value of completely observed data can likely predict the value of the missing cases (Kofman, 2003). This missingness pattern in capital stock data that is consistent with MAR assumption then gives way to the multiple imputation method (Enders, 2010). Second, removing the problematic capital stock data (deletion strategy) will drop almost 40 percent of the firm-level observations. This strategy will result in a severe bias for the intra-industry dynamics decomposition that this paper attempts to obtain. In particular, the contribution of within-firm, reallocation, and firms' turnover in aggregate industry performance will change dramatically if too many observations are dropped within an industry. In addition, the missing data literature also advises against deletion strategy when the missing cases are not fully random (Enders, 2010). However, leaving the reporting issues of capital stock data untreated worsens the bias as it results in a negative capital stock coefficient in the production function, which is highly unrealistic.

Many studies show that the multiple imputation method trumps other methods in treating missing data problems (see the review in Enders (2010) and Schafer and Graham (2002)), especially when the deletion strategy needs to be avoided. One major reason is that multiple imputation technique can produce multiple plausible replacement values, hence acknowledging uncertainty in dealing with missing data. Any imputation method that produces only a single imputed value lacks data variability and consequently underestimates standard errors, increasing the risk of type 1 error (Enders, 2010). Multiple imputation strategy has found application across studies that uses microdata. This includes estimating firm-level total factor productivity

[26] Producing anything without capital would not be possible particularly when firms are operating in the formal sector and are of medium and large establishments.

(TFP) (White et al., 2018), consistent earnings equation under measurement errors (Brownstone and Valletta, 1996), as well as other firm-level analyses (Kofman, 2003, Hollenstein and Woerter, 2008, Grether and Tissot-Daguette, 2021). We acknowledge that there is no single best route in dealing with missing data. The multiple imputation method, however, is the most appropriate technique to be used for this paper, given that the deletion strategy needs to be minimized so that a more accurate picture of intra-industry decomposition can be obtained.

We create five sets of imputed capital stock data where each of which has unique estimates. The process is as follows. Firstly, zero and negative capital stock are treated as missing data since producing anything without capital would not be possible particularly when firms are operating in the formal sector and are medium and large establishments.[27] Next, we apply multiple imputation which consists of three phases: imputation, analysis, and pooling phase. In the imputation phase, a set of linear regression equations based on several observed variables is produced to predict the missing cases. We predict the real capital stock in logarithmic natural term as it is closer to a normal distribution than the level term; hence works better with the linear regression-based model.

The imputing equation is:

$$cap_{it}^{*} = [\widehat{\beta_0} + \widehat{\beta_1} va_{it} + \widehat{\beta_2} ltl_{it} + \widehat{\beta_3} mat_{it} + \widehat{\beta_4} fuel_{it} + \widehat{\beta_5} elec_{it} + \widehat{\beta_6} oiexp_{it} + \widehat{\beta_7} dfo_{it} + \widehat{\beta_8} export_{it}] + z_{it} \quad \text{(B1)}$$

where cap_{it}^{*} is the imputed real capital stock for firm i in period t, z_{it} is residual term drawn from a normal distribution in a Monte Carlo simulation, while the term in bracket comprises the regression coefficient used to generate predicted capital stock. Therefore, the imputed values are inclusive of the stochastic component from the residual term, not just the predicted one. All variables are in log terms except for own_{it} and $export_{it}$ which are dummy variables. The nominal values are deflated using an appropriate deflator.

[27] The zero capital data case would be more believable if the dataset cover small and micro enterprises, where traditional production system are more prevalent.

We use several imputing variables that could plausibly be associated with the level of capital stock during firm's production process, such as real value-added (va_{it}), number of workers (ltl_{it}), raw materials (mat_{it}), fuel ($fuel_{it}$), electricity ($elec_{it}$), and other input expenses ($oiexp_{it}$), as well as share of foreign ownership (dfo_{it}) and export status ($export_{it}$). Although it may seem problematic to use outcome variables such as value-added to impute the missing case, the goal of the imputation is to preserve important associations across variables for the observed data, not establishing causality. Furthermore, the inclusion of error terms in the imputed value could minimize any bias (Enders, 2010). Since the association between capital stock and other observed variables could vary across industries, the imputation is conducted separately for each 4-digit ISIC code.

There are some key features of the multiple imputation technique. Firstly, the missing cases are assumed to be missing at random (MAR). This assumption means that the missing data is independent of the value of the capital stock itself but could be related to other completely observed variables. It is distinct from the missing completely at random (MCAR) assumption as the MCAR implies independence from both its own value and other observables. MAR assumption allows for the use of multiple imputation technique as the value of completely observed variables can be used to predict the missing cases. Table B1 supports this assumption since the missingness pattern is related to some completely observed variables or covariates. Secondly, the imputed data in this technique includes the residual term. This adds greater variability to the multiple-imputed data compared to other basic imputation technique, such as mean imputation.

The completely observed variables in equation B1 are used to predict the missing cases. After the first imputation, a new set of residual term is randomly drawn from a Monte Carlo simulation and added to the previously imputed data to make alternative estimates of missing data. The new imputed values then necessitate a different set of residual terms and coefficients, which will be used to make alternative estimates of imputed data. The whole process reiterates until five sets of imputed data are produced (see the more detailed procedure in Enders (2010)).

Table B1. Mean difference of other observed variables by missingness status in capital stock data

Other observed variables	Mean value: Observed case of the incomplete data	Mean value: Missing case of the incomplete data	two-sided p-value of the two-sample t-test
Value-added (real, log)	8.709	9.347	0.0000
Total labor (log)	4.164	4.269	0.0000
Raw material	8.397	8.687	0.0000
Fuel expenses	4.849	4.976	0.0000
Electricity expenses	3.639	4.268	0.0000
Other expenses	6.058	5.920	0.0000
Foreign ownership	5.814	7.047	0.0000
Exporting firms	0.152	0.092	0.0000
Observations	363,360	211,628	

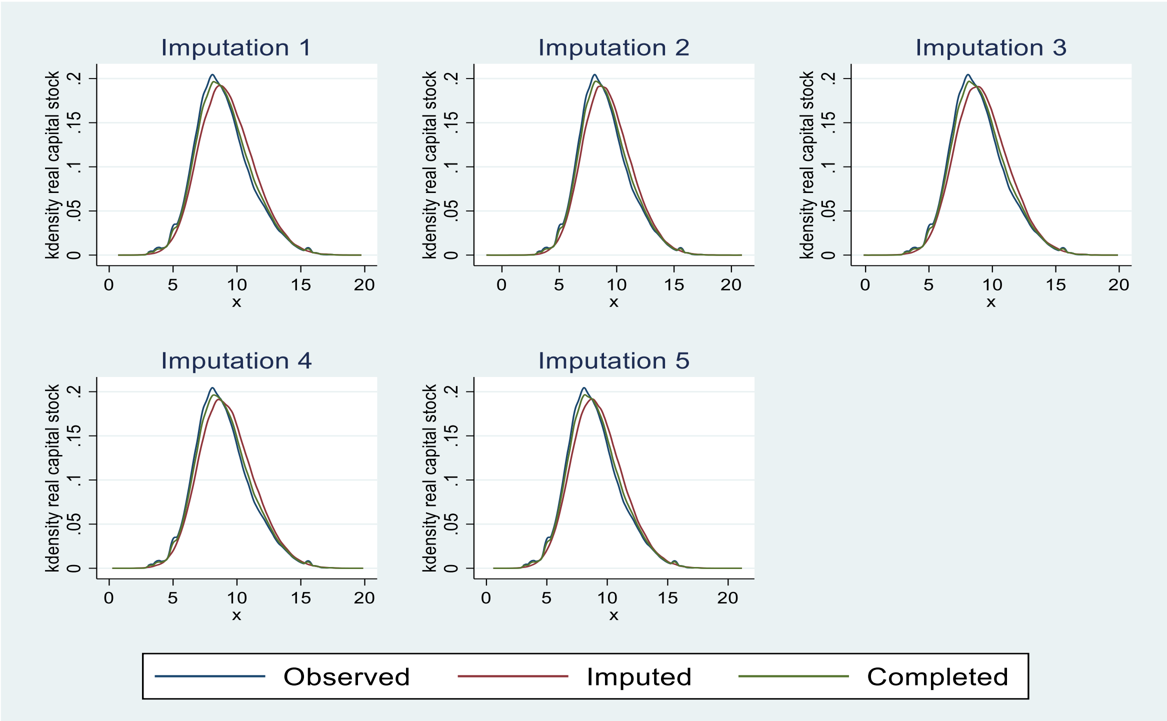
Note: The two-sample *t*-test is applied for the pool of all firms across different industries and periods. The findings are qualitatively similar when the *t*-test is applied for each period. Results are available upon request.

Source: Authors' estimation.

The five complete sets of capital stock data are then used for subsequent analysis. This includes recovering TFP and markup after production function estimation. To obtain the input coefficient, production function estimation is conducted separately for each of the complete datasets. Finally, parameter estimates and standard errors are combined into a single result following Rubin's rule as described in Enders (2010).[28] Diagnostics based on Eddings and Marchenko (2012) shows that there is a high degree of overlap between observed and imputed values, indicating a good fit of the imputation model to produce the complete datasets on capital stock (Figure B1). As a result, the total value of real capital stock data from these sets shows an upward trend over time, consistent with the increasing number of firms covered in the SI over the period (Figures B2 and B3).

[28] We use the default mi command provided in Stata.

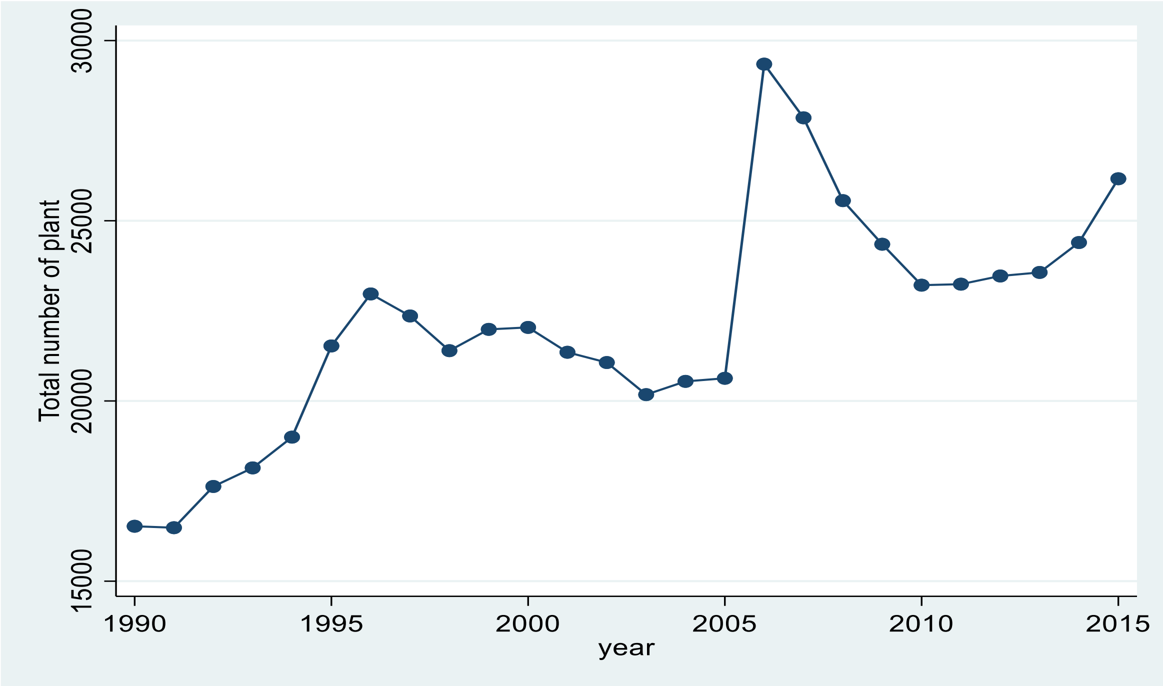
Figure 1. Diagnostics for multiple imputation of capital stock data



Note: Observed data represents the original data before being filled in by multiple imputation method, meanwhile, completed data consists of both original and the imputed data.

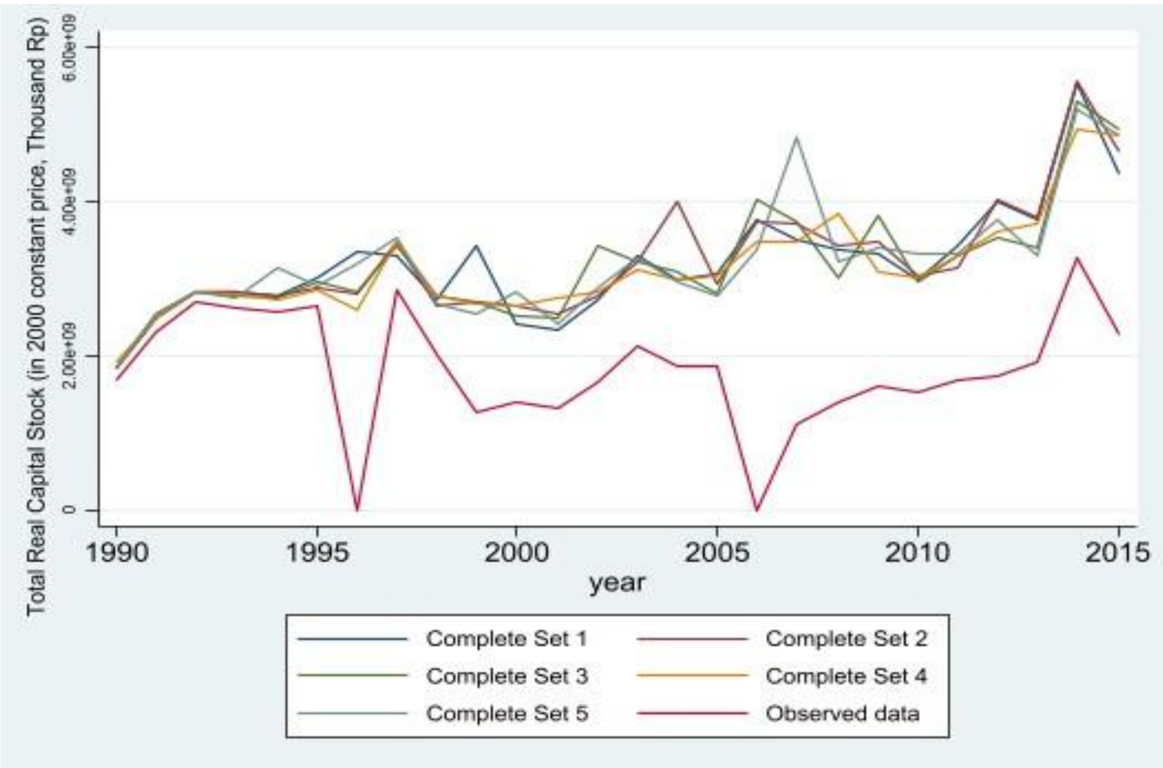
Source: Authors' estimation.

Figure B2. Total number of firms in SI dataset, 1990-2015



Source: Authors' estimation.

Figure B3. Real capital stock trend, imputed and observed, 1990–2015



Note: Complete set 1 – 5 include the original as well as imputed data

Source: Authors' estimation.

Appendix C. Estimating production function, TFP, and markup

This study follows Pane and Patunru (2021), among others, in using value-added as an outcome in the production function estimation.[29] The main estimating equation assumes Cobb-Douglas functional form and is given by:

$$va_{it} = \beta_0 + \beta_l ltl_{it} + \beta_k cap^*_{it} + \varepsilon_{it} \qquad \text{(C1)},$$

where va_{it} denotes firm i's value added at time t, while ltl_{it} and cap^*_{it} are number of workers and (imputed) real capital stock respectively—all in log terms. Estimation of input coefficient in a single equation as in Equation C1 suffers from a well-known endogeneity bias. The problem arises because firm's input choice at time t may not be determined exogenously but may be influenced by productivity shock ω_{it} which is contained in ε_{it} and is only observed by firms—not by researchers. In this case, ε_{it} can be thought of as a composite error consisting of unobserved productivity and a pure error ϵ_{it} in additive manner ($\varepsilon_{it} = \omega_{it} + \epsilon_{it}$).

The most widely used strategy in the literature is the control function approach where production function estimation is conducted in two different stages. Olley and Pakes (1996) (OP estimator) used firm's investment as a proxy for the unobserved firm's productivity, while Levinsohn and Petrin (2003) (LP estimator) resort to intermediate goods to be the proxy as the investment data is usually sparse in the firm-level database. In the OP and LP method, labor input coefficient is estimated in the first stage along with non-parametric representation of

[29] Literature uses both the gross and value-added production function in estimating input coefficient. The reason we prefer value-added one is twofold. First, the value-added production function, rather than the gross output one, provides more flexibility to be used with the more recent production function estimator, especially the Ackerberg-Caves-Frazer (ACF) estimator (see Gandhi et al. (2020) and Ackerberg et al. (2015) on the issue of using gross output production function with ACF estimator). Second, the value-added production function estimates have a closer linkage to overall welfare as aggregate value-added mimics the aggregate demand function (Petrin and Levinsohn, 2012, Melitz and Polanec, 2015). Consistent use of value-added in the labor share definition and production function, therefore, allows for evaluating how welfare is distributed across production factors over time.

unobserved productivity [30], while estimation of capital input coefficient is conducted at the second stage to avoid collinearity issues.

However, recent studies show that estimating the labor input coefficient along with non-parametric representation of unobserved productivity may suffer from bias. This issue emerges because labor input may be colinearly related to the non-parametric control function in the first stage. Ackerberg et al. (2015) (ACF) highlighted that labor input might not be a free variable and constitute a deterministic function along the lines of investment and material demand function in OP and LP, respectively. In this case, the OP and LP method suffers from functional dependence problem. There is not enough variation in the labor input variable to identify the labor input coefficient in the first stage. According to the ACF method, labor input may also depend on capital stock and unobserved productivity shock: $ltl_{it} = ltl_{it}(cap_{it}, \omega_{it})$, just as how material input mat_{it} demand function behave in LP method: $mat_{it} = mat_{it}(cap_{it}, \omega_{it})$. Therefore, labor input should constitute part of the non-parametric function to control unobserved productivity: $\omega_{it} = \omega_{it}(cap_{it}, mat_{it}, ltl_{it})$, hence all of the input coefficient is estimated at the second stage (see De Loecker and Warzynski (2012) and Manjón and Mañez (2016) for more detail explanation on ACF procedures).

Considering the recent development and practicality, we use the LP estimator in conjunction with the ACF method instead of the OP estimator. However, this paper will mainly use estimates based on LP as a benchmark since the ACF method still suffers from convergence and stability issues, despite the functional dependence problem it addresses (Manjón and Mañez, 2016). Table C1 indicates the parameter estimates of the production function from Fixed Effect, LP, and ACF methods.

Table C1. Comparison of parameter estimates from various methods

	(1) Fixed effect	(2) Levinsohn-Petrin	(3) ACF Cobb-Douglas
Labor (l)	0.780***	0.478***	0.531***
	[0.00289]	[0.00183]	[0.000648]

[30] Non-parametric function of unobserved productivity is often approached using high-degree polynomial function in practice.

Capital stock (k)	0.128^{***}	0.170^{***}	0.218^{***}
	[0.00141]	[0.00151]	[0.00188]
Number of observations	574988	574988	574988
Number of groups	67187	67187	67187

Note: Production function estimation uses the complete set of data that includes imputed data using multiple imputation technique. The estimates are then pooled using Rubin's rule. Production function based on LP and ACF technique is executed by prodest command provided in STATA. For ACF technique, the value of 0.5 is used as initial point for both coefficients. Standard errors are in brackets. The following sign shows significance level: + p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001.

Source: Authors' estimation.

Table C2 compare the estimates of input coefficients using three different capital stock data: imputed sets, original data (including missing, zero, and negative capital stock), and the strictly positive ones. Capital coefficient produces a negative result when the original data is used, although not statistically significant. Meanwhile, by including only positive capital stock data in production function estimation, although we obtained a positive capital coefficient, the number of observations dropped significantly by almost 40 percent of the original samples. In contrast, the imputed capital stock datasets produce a realistic result with a positive capital input coefficient and a relatively consistent one with the literature without sacrificing the number of observations. For this reason, we opt for using the imputed datasets throughout the analysis.

Table C2. Comparison of input coefficient using various capital stock datasets: LP and ACF methods

	Multiple imputation		All data (incl. k=0)		Positive capital stock only (k>0)	
	(1) LP	(2) ACF	(3) LP	(4) ACF	(5) LP	(6) ACF
Labor (l)	0.478^{***}	0.531^{***}	0.571^{***}	0.317^{*}	0.522^{***}	0.472^{***}
	[0.00334]	[0.000935]	[0.00274]	[0.129]	[0.00227]	[0.00604]
Capital stock (k)	0.170^{***}	0.218^{***}	-0.000711^{**}	-0.0134^{***}	0.169^{***}	0.195^{***}
	[0.00149]	[0.00239]	[0.000268]	[0.00259]	[0.000882]	[0.00463]
Number of Observations	574988	574988	552374	552374	363360	363360

Number of Group	67187	67187	62089	62089	52921	52921

Note: Production function estimation that uses the complete dataset from multiple imputation techniques (column (1) and (2)) is pooled following Rubin's rule. The zero capital data is added by value "1" before taking log transformation to allow for regression in column (3) and (4). Production function based on LP and ACF technique is executed by prodest command provided in Stata. For ACF technique, the value of 0.5 is used as initial point for both coefficients. Standard errors are in brackets. The following sign shows significance level: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Source: Authors' estimation.

This study uses the coefficients from the LP method to construct TFP and markup variables. The log of TFP can be derived directly following production function estimation. Meanwhile, markups need to be further calculated using the input coefficient and the firm's labor share. We follow production function literature to construct markup as in De Loecker and Warzynski (2012), De Loecker et al. (2020), and Autor et al. (2020) that define markups as a ratio of price over marginal cost. Assuming firms choose input based on cost minimization procedures, markups (μ_{it}) can be defined as the wedge between labor's output elasticity and labor expenditure share in revenue (labor share) given by the following formula: $\mu_{it} = \frac{\beta_l}{s_{it}}$.[31] The formula implies that in a setting where a firm has market power, firms choose labor input below its optimal productivity level, driving a wedge between the two and pushing up markup. However, we further follow De Loecker and Warzynski (2012) and Gradzewicz and Mućk (2019) in correcting the standard markup formula with the residual term (ϵ_{it}) from production function estimation to avoid bias due to improper deflator. The corrected markup ($\widetilde{\mu_{it}}$) formula

[31] Output elasticity of labor input varies across firms and time (β_{it}^{l}) in the translog specification of ACF estimator depending on firm's level of worker and capital stock in that particular period (see De Loecker and Warzynski (2012) for the derivation).

then become: $\widetilde{\mu_{it}} = \frac{\mu_{it}}{\epsilon_{it}}$. For completeness, we present the markup estimation based on the production function approach and the one based on a simple accounting formula.[32]

[32] The simple markup formula is defined as total production value over variable cost covering expenses for materials and worker's compensation