# Trade Liberalization and Markup Dispersion: Evidence from China's WTO Accession

Yi Lu and Linhui Yu*

First Draft: May 2013
Revised: March 2015

**Abstract**

In this paper, we empirically investigate whether trade liberalization affects markup dispersion, a potential source of resource misallocation. The identification uses China's WTO accession at the end of 2001. We show that trade liberalization reduces markup dispersion within a narrowly defined industry. We also examine both price and cost responses to trade liberalization, as well as heterogeneous effects across firms and across locations. Our study contributes to the literature by identifying another potential channel through which free trade benefits a nation.

Resource misallocation is common, especially in developing countries (e.g., Banaerjee and Duflo (2005)), and helps explain substantial differences in productivity across countries (e.g., Hsieh and Klenow (2009), Bartelsman, Haltiwanger, and Scarpetta (2013)). While recent studies have exploited imperfections in the input markets (such as capital markets) to identify the source of misallocation (e.g., Restucccia and Rogerson (2008), Midrigan and Xu (2014)), distortions in the product market also play an important role in generating allocative inefficiency. Robinson (1934) show that first-best efficiency is achieved when markups are the same across products. In a world with markup dispersion, industries/firms with higher markups employ resources at less than optimal levels, while those with lower markups produce more than optimal, resulting in efficiency losses (e.g., Lerner (1934), Opp, Parlour, and Walden (2014)).

*Lu: Department of Economics, National University of Singapore, 1 Arts Link, Singapore, 117570 (e-mail: justinly6@gmail.com); Yu: School of Economics, Zhejiang University, 38 Zheda Road, Hangzhou, China 310027 (e-mail: yulh@zju.edu.cn). We would like to thank Wen-Tai Hsu, Lianming Zhu, anonymous referees, and participants in conferences and workshops for their insightful comments and suggestions, which have significantly improved the paper. Lu acknowledges the support of a research grant from the National University of Singapore (the Ministry of Education AcRF Tier 1 FY2014-FRC2-001). Yu acknowledges the support of a research grant from the National Natural Science Foundation of China (project code 71402162).

Instead of identifying sources of misallocation, we investigate how to reduce misallocation degree, and, more specifically, examine the effect of trade liberalization on markup dispersion. With substantial reductions in trade costs and advancements in telecommunications and logistics, globalization has become a dominant feature of the world, and has significantly changed world production in the past decades. By intensifying market competition, trade may affect the distribution of firm markups through two channels: adjustment of markups by surviving firms (the intensive margin) and entry and exit (the extensive margin). Depending on the distributional assumption of firm productivity draw, theoretical predications are mixed. Using the Pareto distribution, Bernard, Eaton, Jensen, and Kortum (2003) and Arkolakis et al. (2012) find that intensive and extensive margins of trade liberalization on markup dispersion cancel each other out in the equilibrium. However, de Blas and Russ (2012) and Holmes, Hsu, and Lee (2014) find that the departure from the Pareto distribution can cause the distribution of firm markups to be responsive to trade costs.

This paper contributes to the above literature on three grounds. First, instead of assuming distributional functions and deriving theoretical results, we use real data to check whether trade liberalization affects markup dispersion. Understanding whether markup dispersion changes or not in response to trade liberalization is an important step in investigating the pro-competitive role of trade—that is, change in trade costs may affect resource reallocation across firms through changes in markup distribution. Second, we explore one of the most important trade episodes in the 2000s—China's accession into the WTO—and use the most comprehensive firm-level data in China from 1998 to 2005. The liberalization degree upon WTO accession in China is found to be large (for a detailed description of China's WTO accession, see Lardy (2002)), and its effects on the competitiveness of Chinese firms and welfare gains are found to be significant. For example, Brandt et al. (2012) find that a 10 percent reduction in tariff protection leads to a permanent 6 percent increase in industry-level productivity. And di Giovanni, Levchenko, and Zhang (2012) calculate overall welfare gains of 3.72 percent compared to autarky. Third, we use newly compiled data to more accurately estimate production function and calculate firm markups (for details on empirical data, see Section 4.1). Specifically, we are able to measure output in physical terms (which avoids the omitted output price bias), adopt a flexible specification of production function (i.e., translog), and use a control function approach developed by De Loecker et al. (2014) to address the issue of omitted firm-specific input prices (for detailed discussion of our production function estimation and comparison to other approaches, see Appendix B). The markup calculation is based on the methodology developed by De Loecker and Warzynski (2013), which relies on the intuition that the output elasticity of a variable production factor equals its expenditure share in total revenue only when price equals marginal cost and in a world with imperfect competition, markup is the wedge between input revenue share and output elasticity of this input (see Section 3.3 for details of the markup estimation).

Our identification is essentially a difference-in-differences (DID) estimation, that is, we compare markup dispersion in industries experiencing greater tariff reduction upon WTO accession (the treatment group) to that in industries experiencing less tariff reduction (the control group) before and after 2001, the time of WTO accession. We find that trade liberalization significantly reduces the dispersion of firm markups. Results hold for different measures of markup dispersion, inclusion of many industry-varying characteristics, and finer definition of industries.

The validity of our DID estimation hinges on two assumptions: (1) the treatment group would have followed the same trend as the control group in the case without WTO accession, and (2) no other policy reform differentially targeted our treatment and control groups at the time of WTO accession. As checks on the first identifying assumption, we first show that treatment and control groups followed the same trends in markup dispersion until the WTO accession, then began to diverge right after the accession. Second, we carefully investigate what caused tariffs to differ across industries before WTO accession (or the pretreatment differences between treatment and control groups), then control for potential differential trends in markup dispersion after WTO accession generated by significant pretreatment differences, a approach similar to the one used by Gentzkow (2006). Third, we check and rule out the possibility that firms may have changed their behavior and, hence, markup dispersion in anticipation of the WTO accession.

As checks on the second identifying assumption, we control for two important ongoing policy reforms in the early 2000s, i.e., the state-owned-enterprises (SOE) reform and the relaxation of FDI regulations. We also control for changes in intermediate input tariffs and the effect of accessing foreign markets. As further robustness checks, we use two placebo tests (i.e., the sample from the pre-WTO period and the sample of processing traders), control for cross-product-within-industry tariff variations, and control for multi-industry issues (see Sections 5.3-5.4 for details).

To gain further insights about how trade liberalization changes markup dispersion, we first verify that imports increase more in product categories experiencing greater tariff reduction, thereby establishing the market competition linkage. We then investigate markup responses at different quantiles along the distribution, and find that trade liberalization increases markups at the lower quantiles but reduces them at higher quantiles, which in turn flattens the markup distribution. Furthermore, we look at price and marginal cost components of markup separately, and find that trade liberalization reduces the dispersion of both prices and marginal costs. Finally, we uncover heterogeneous effects across firms (i.e., surviving firms versus entries/exiters; SOEs versus non-SOEs) and across locations (i.e., coastal versus inland cities).

Our paper is related to several strands of literature, as well as the studies on resource misallocation mentioned above. The first strand reflects the recent renaissance in gains from trade, due to the availability of micro-level data and development of new theories. In a recent influential paper, Arkolakis, Costinot, and Rodríguez-Clare (2012) show that benefits from free trade can be pinned down by two parameters: share of expenditure on domestic goods and an elasticity of imports with respect to variable trade costs; these results apply to a variety of trade models with or without firm heterogeneity. However, the constant markup assumption used by Arkolakis, Costinot, and Rodríguez-Clare (2012) excludes any pro-competitive effects of trade. In contrast, using a variable markup setup, Edmon, Midrigan, and Xu (2014) quantify the gains from trade using data from Taiwan. Our work departs from Edmon, Midrigan, and Xu (2014)'s in that, first, they use a structural estimation approach, while ours is a reduced-form estimation; second, they conduct a counterfactual analysis of trade liberalization, whereas we study a real incident of trade liberalization.

Our study is also related to literature on the relation between trade and average markups, such as work by Levinsohn (1993), Harrison (1994), Krishna and Mitra (1998), Konings, Van Cayseele, and Warzynksi (2001), Chen, Imbs, and Scott (2009), and De Loecker et al. (2014). However, there are significant differences in focus between these studies and ours (i.e., markup level vs. markup dispersion), and hence the welfare implication. If the reduction in markup levels comes through productivity improvement, this constitutes a productive efficiency gain channel from trade. However, as industries/firms are potentially affected differentially by trade, allocative efficiency may be improved or worsened. Lipsey and Lancaster (1956-1957) make the point that rendering already competitive sectors even more competitive reduces overall welfare. Recently, Holmes, Hsu, and Lee (2014) develop a formula to decompose overall welfare gains from trade into improvements in productive efficiency and allocative efficiency, and Hsu, Lu, and Wu (2014) further show that the latter can be a significant component of welfare analysis of trade in the case of China's WTO accession.

In the context of China, Brandt et al. (2012) investigate how China's WTO accession affects productivity growth at both firm and industry levels, as well as outcomes such as industry average price deflators and industry average markups. While both their and our papers consider beneficial effects of WTO accession on the domestic economy in China and use the same data, there are several important differences. First, the two papers investigate different channels for gains from trade. Brandt et al. (2012)'s focus on productivity gains (as well as industry-average markups) confirms traditional welfare gains from trade through productive efficiency improvement as identified by Arkolakis, Costinot, and Rodríguez-Clare (2012), whereas we investigate another important channel of gains from trade, that is, the change in markup dispersion, which in turn affects resource allocation within an industry. Second, the two papers use different methods of production function estimation, which is a crucial step in the construction of firm productivity and markups. Brandt et al. (2012) estimate a revenue-based and Cobb-Douglas production function, which results in (1) the same output elasticities of inputs across firms in the same sector and (2) potential biases from omitted firm-specific output and input prices (see De Loecker and Goldberg, 2014, for more discussion of these issues). To address these concerns, we estimate a quantity-based and translog production function with a control function for omitted firm-specific input prices (for comparison of different production function estimations, see Appendix B).[1]

# I. Background: China's WTO Accession

*The Process.*—In July 1986, China notified GATT (WTO's predecessor) that it would like to resume its status as a GATT contracting party; this lasted for 15 years. Between 1987 and 1992, as China debated whether to continue the market reform or move back to the planned economy system, the return to GATT was suspended. Momentum resumed after Deng Xiaoping's southern tour speech in 1992, and in July 1995, China officially applied to join the WTO.

A pivotal aspect of China's WTO accession process involved bilateral negotiations between China and WTO members. New Zealand was the first country to sign a bilateral agreement (in August 1997) with China regarding WTO accession. However, negotiations between China and the U.S. took 25 rounds and four years before an agreement was reached in November 1999. After that, China signed agreements with 19 countries within six months, including Canada in November 1999 and the European Union in May 2000. In September 2001, China signed an agreement with Mexico, at which point negotiations with all WTO member countries were complete. Finally, on November 10, 2001, WTO's Ministerial Conference approved by consensus the text of the agreement for China's entry into the WTO.

*Tariff Reduction.*—As a condition to joining the WTO, China carried out a large and widespread tariff reduction between 1992 and 1997. Specifically, in 1992, China's (unweighted) average tariff was as high as 42.9%. Shortly after the GATT Uruguay round of negotiations, China substantially reduced tariffs, i.e., the average tariff dropped from 35% in 1994 to around 17% in 1997. There was little change in tariffs after 1997, however, until China joined the WTO at the end of 2001.

In early 2002, China started to fulfill its tariff reduction responsibilities as a WTO member. According to the WTO accession agreement, China was required to complete tariff reductions by 2004 (with a few exceptions to be completed by 2010); average tariffs for agriculture and manufacturing goods would be reduced to 15% and 8.9%, respectively.

[1] Specifically, we find that markups from our estimation and those from Brandt et al. (2012)'s method are negatively correlated (i.e., the correlation is $-0.1379$; see Appendix Table 6).

Figure 1 plots China's (unweighted) average tariffs during the period 1996-2007. Tariff rates dropped substantially in 1996, followed by a relatively stable period from 1997 to 2001 and another round of gradual reduction in 2002 before reaching a steady state in 2005. Unweighted average tariffs dropped from 15.3% in 2001 to 12.3% 2004, while weighted average tariffs declined from 9.1% to 6.4%.

[Insert Figure 1 here]

Interestingly, tariff reduction upon WTO accession exhibited great heterogeneity across products. As shown in Figure 1, the ratio of tariffs at the 25th percentile to those at the 75th percentile also dropped sharply in 2002 and then stabilized after 2005. In Figure 2, we further plot the relation between tariffs in 2001 (the year just before WTO accession) and changes in tariffs between 2001-2005 across 3-digit industries (the unit used in the main regression analysis).[2] Clearly, there is a strong, positive correlation, implying that industries with higher tariffs before WTO accession experienced greater tariff reduction after WTO accession. Presumably, China had to reduce tariffs to WTO-determined levels, which are quite uniform across products, whereas China's pre-WTO tariffs varied widely across products.

[Insert Figure 2 here]

# II. Empirical Strategy

## A. Specification

To identify the impact of trade liberalization on markup dispersion, we explore the fact that after China joined WTO, industries that had previously been more protected (i.e., industries with higher tariffs in 2001) experienced greater tariff reduction under the WTO agreement and therefore higher degrees of liberalization, whereas previously more open industries (i.e., industries with lower tariffs in 2001) witnessed small changes in tariffs and therefore less liberalization. These differential degrees of liberalization and the timing of tariff reductions (i.e., 2002) allow us to conduct a DID estimation—that is, to compare the change in markup dispersion in previously more protected industries (the treatment group) before and after 2001 to the corresponding change in those previously more open industries (the control group) during the same period (see, for example, Guadalupe and Wulf (2010), for a similar approach).

[2] A similar pattern is seen at the HS-6 product level (see Appendix Figure 1).

The specification for our DID estimation is

$$y_{it} = \alpha_i + \beta Tariff_{i2001} \cdot Post02_t + \mathbf{X}'_{it}\boldsymbol{\gamma} + \lambda_t + \varepsilon_{it}, \tag{1}$$

where $i$ and $t$ represent industry and year, respectively; $y_{it}$ is the measure of markup dispersion in industry $i$ at year $t$ (see Section 4.3 for details); $Tariff_{i2001}$ is the tariff rate of industry $i$ in 2001;[3] $Post02_t$ denotes a post-WTO period, taking a value of 1 if it is year 2002 and onwards, and 0 otherwise; $\alpha_i$ is the industry fixed effect, controlling for all time-invariant differences across industries; $\lambda_t$ is the year fixed effect, controlling for all yearly shocks common to industries, such as business cycles; and $\varepsilon_{it}$ is the error term. To cope with the potential heteroskedasticity and serial autocorrelation, we cluster standard errors at the industry level (see Bertrand, Duflo, and Mullainathan, 2004).

To isolate the effect of trade liberalization, we control for several time-varying industry characteristics ($\mathbf{X}_{it}$) that may affect markup dispersion, such as industrial agglomeration degree (measured by the Ellison-Glaeser [EG] index, with a higher value indicating a higher degree of geographic concentration; see Ellison and Glaeser, 1997, for the development of the measurement) and entry barriers (proxied by the average fixed asset [in log] and the number of firms).

In the main specification, we define an industry at the three-digit Chinese Industrial Classification (CIC) level—presumably there are relatively more observations within such defined industries and therefore smaller measurement errors of our outcome variable. However, to address concerns regarding any potential aggregation bias, we conduct a robustness check at the four-digit CIC level, which is the finest definition in our data.

Note that we use the interaction of tariffs in 2001 ($Tariff_{i2001}$) and the post-WTO indicator ($Post02_t$) as our regressor of interest, instead of yearly tariffs ($Tariff_{it}$). One motivation is that the schedule of tariff reduction upon WTO accession in China was released in 2002, and hence the phase-out process is expected and could be exploited by the producers. As explained by Liu and Trefler (2011), use of the interaction between $Tariff_{i2001}$ and $Post02_t$ can capture both real and expected effects of trade liberalization. Nonetheless, using yearly tariffs ($Tariff_{it}$) produces similar results (see Appendix Table 1, Column 3), albeit marginally insignificant.

## B. Identifying Assumption and Checks

The identifying assumption associated with our DID estimation specification (1) is that conditional on a list of controls ($\alpha_i, \mathbf{X}_{it}, \lambda_t$), our regressor of interest, $Tariff_{i2001} \cdot Post02_t$,

[3] Using average tariffs over 1997-2001 or tariffs in 1997 generates similar results (see Appendix Table 1, Columns 1-2); presumably, tariffs did not change much between 1997 and 2001.

is uncorrelated with the error term, $\varepsilon_{it}$, i.e.,[4]

$$E\left[\varepsilon_{it}|Tariff_{i2001} \cdot Post02_t, \alpha_i, \mathbf{X}_{it}, \lambda_t\right] = E\left[\varepsilon_{it}|\alpha_i, \mathbf{X}_{it}, \lambda_t\right]. \quad (2)$$

In other words, markup dispersion in the treatment group would have followed the same trend as that in the control group if there had been no trade liberalization in 2002.

Concerns may exist, however, about the satisfaction of our identifying assumption—specifically, the timing of the WTO accession, the nonrandom selection of tariffs in 2001, and other simultaneous policy reforms. First, one might be concerned that approval of China's WTO accession at the end of 2001 was expected, and therefore firms could adjust their behavior even before tariff reductions took effect in 2002. However, China's WTO accession process was lengthy, taking about 15 years to complete, and approval required the consensus of all WTO member countries. Although China achieved important breakthroughs by signing agreements with the U.S. in 1999 and the EU in 2000, many issues remained unsolved until mid-2001. Hence, the timing of China's WTO accession was quite uncertain before 2001. Nonetheless, as a robustness check, we include an additional control in the DID regression, $Tariff_{i2001} \times One\ Year\ Before\ WTO\ Accession_t$, to examine whether firms changed their behavior—and therefore markup dispersion—in anticipation of WTO accession the following year.

Second, while the use of tariffs in 2001 is less susceptible to endogeneity concerns, the choice of these tariffs was nonrandom, raising the possibility that treatment and control groups could be systematically different *ex ante*. To alleviate the possibility that some pre-existing differences between treatment and control groups might also differentially affect markup dispersion by these two groups even after WTO accession (and therefore contaminate our DID estimates), we first carefully characterize significant tariff determinants in the pre-WTO period (for details, see Appendix A and Appendix Table 2), and then control flexibly for post-WTO differences in the time path of the outcome variable generated by these pre-existing differences (see Gentzkow (2006) for details on this approach). Specifically, we add interactions between those significant tariff determinants ($\mathbf{Z}_{i2001}$) with our post-WTO indicator ($Post02_t$), i.e., $\mathbf{Z}_{i2001} \cdot Post02_t$, to our DID regression.

Third, if other policy reforms differentially targeted our treatment and control groups around the time of the WTO accession (i.e., the end of 2001), our DID estimates might also capture the effects of these reforms, making it hard to pinpoint the effect of trade liberalization. Two important ongoing reforms in the early 2000s were the SOEs reform

[4] Note that the identification does not require our control variables to be exogeneous, i.e.,

$$E\left[\varepsilon_{it}|\alpha_i, \mathbf{X}_{it}, \lambda_t\right] = 0.$$

In other words, for these control variables, estimated coefficients may not have causal interpretations. See Stock and Watson (2012, p.274) for more discussion of this point.

and the relaxation of FDI regulations (i.e., fewer regulations on wholly owned FDI). To control for any confounding effects from these policy reforms, we include in our DID estimation $SOE\ Share$ (measured by the ratio of the number of SOEs to the number of domestic firms) and $FDI$ (measured by the logarithm of the number of foreign-invested firms).

To further check our identifying assumption, we conduct two placebo tests: one using only the pre-WTO data as by Topalova (2010), and the other using the sample of processing traders. For details, see Section 5.3.

## C. Estimation of Firm Markups

The crucial component in constructing our outcome variable is firm-level markup, defined as the ratio of price to marginal cost. However, firm-level data rarely contain information on product prices, let alone information on marginal costs. To recover firm-level markup, we follow the recent work of De Loecker and Warzynski (2012). Specifically, it is assumed that the production function of firm $i$ at time $t$ is[5]

$$Q_{it} = F_{it}\left(L_{it}, K_{it}, M_{it}, \omega_{it}\right) \tag{3}$$

where $L_{it}$, $K_{it}$, and $M_{it}$ are physical inputs of labor, capital, and intermediate materials, respectively; and $\omega_{it}$ denotes firm-specific productivity. Production function $F\left(.\right)$ is assumed to be continuous and twice-differentiable with respect to all of its arguments.

Consider the following cost-minimization problem faced by firm $i$ at time $t$

$$\min_{\{L_{it}, K_{it}, M_{it}\}} w_{it}L_{it} + r_{it}K_{it} + p_{it}^{m}M_{it} \tag{4}$$

$$s.t.\ F_{it}\left(L_{it}, K_{it}, M_{it}, \omega_{it}\right) \geq \bar{Q}_{it} \tag{5}$$

$$L_{it} \geq I\left[D_{it} = 1\right]\bar{S}_{it} \tag{6}$$

where $w_{it}$, $r_{it}$, and $p_{it}^{m}$ denote the wage rate, rental price for capital, and the price of intermediate inputs, respectively; $D_{it}$ is an indicator of state-owned enterprise, i.e., taking a value of 1 if firm $i$ is an SOE at time $t$ and 0 otherwise; and $I\left[.\right]$ is an indicator function that takes a value of 1 if the statement in the bracket is true and 0 if not.

The constraint equation (6) captures a prominent feature of SOEs, namely that they are often required to hoard redundant labor to meet a minimum level of employment ($\bar{S}_{it}$) so as to help bureaucrats maintain social stability.[6]

[5] Note that the framework is robust to any arbitrary number of inputs. As we mainly observe three inputs (labor, capital, and intermediate materials) in our data, we here focus on a production function with only these three inputs.

[6] For example, during the financial crisis of 2008-2009, Chinese President Hu Jintao announced publicly

Estimation of firm-level markup hinges on the optimal choice of inputs free of any adjustment cost and estimation of output elasticities of inputs. As labor is not freely chosen due to constraint (6), and capital is often considered to be a dynamic input, we focus on the optimal choice of intermediate materials.[7] Specifically, the Lagrangian function associated with optimization problem (4) can be written as

$$\begin{aligned}\mathcal{L}\left(L_{it}, K_{it}, M_{it}, \lambda_{it}, \eta_{it}\right) &= w_{it}L_{it} + r_{it}K_{it} + p_{it}^{m}M_{it} \\ &\quad +\lambda_{it}\left[\bar{Q}_{it} - F_{it}\left(L_{it}, K_{it}, M_{it}, \omega_{it}\right)\right] \\ &\quad +\eta_{it}\left[I\left[D_{it}=1\right]\bar{S}_{it} - L_{it}\right]. \end{aligned} \tag{7}$$

Hence, the first-order-condition for intermediate materials is

$$\frac{\partial \mathcal{L}}{\partial M_{it}} = p_{it}^{m} - \lambda_{it}\frac{\partial F_{it}}{\partial M_{it}} = 0. \tag{8}$$

Rearranging equation (8) and multiplying both sides by $\frac{M_{it}}{Q_{it}}$ leads to

$$\begin{aligned}\frac{\partial F_{it}}{\partial M_{it}}\frac{M_{it}}{Q_{it}} &= \frac{1}{\lambda_{it}}\frac{p_{it}^{m}M_{it}}{Q_{it}} \\ &= \frac{P_{it}}{\lambda_{it}}\frac{p_{it}^{m}M_{it}}{P_{it}Q_{it}}, \end{aligned} \tag{9}$$

where $P_{it}$ is the price of the final good.

Note that $\lambda_{it} = \frac{\partial \mathcal{L}}{\partial Q_{it}} = c_{it}$ represents the marginal cost of production at a given level of output. We define the markup $\mu_{it}$ as the ratio of price to marginal cost, i.e., $\mu_{it} \equiv \frac{P_{it}}{\lambda_{it}}$. Therefore, equation (9) leads to our estimation expression of firm-level markup[8]

$$\mu_{it} = \theta_{it}^{m}\left(\alpha_{it}^{m}\right)^{-1}, \tag{10}$$

where $\theta_{it}^{m} \equiv \frac{\partial F_{it}}{\partial M_{it}}\frac{M_{it}}{Q_{it}}$ is the output elasticity of intermediate materials and $\alpha_{it}^{m} \equiv \frac{p_{it}^{m}M_{it}}{P_{it}Q_{it}}$ is

that SOEs could not lay off their employees and should instead try to expand employment.

[7] We admit that cost-minimization with respect to material inputs is at best an approximation for characterizing SOE behavior. It is likely that SOEs would use more materials than necessary in production because of their lack of incentives to minimize costs. Nonetheless, compared with the problem of overemployment of labor, the overuse of material inputs is less of a concern in the literature. Du et al. (2014) use the DID method to examine changes in labor employment and material inputs after restructuring, and find that labor employment exhibits a significant decline after privatization, but materials show no significant changes. These findings suggest that SOEs truly suffered from redundant employment problems before privatization (a prominent symptom of SOEs around the world) but they did not have a serious problem with the overuse of material inputs. Given that material inputs had been adjusted relatively freely even in SOEs due to their much smaller adjustment costs than those for labor, we employ materials to recover firm-level markup.

[8] Note that this expression holds under any form of competition. In particular, De Loecker and Warzynski (2012) discuss alternative settings of market competition, including Cournot competition, Bertrand competition, and monopolistic competition, which lead to a similar estimation expression for firm-level markup.

the share of expenditure on intermediate materials in total revenue.

As information on expenditure on intermediate materials and total sales is available in the data, $\alpha_{it}^{m}$ can be readily calculated. However, the output elasticity of intermediate materials $\theta_{it}^{m}$ requires the estimation of production function. There is extensive literature on the estimation of production function, which focuses on how to control for the unobserved productivity shock (see Ackerberg, Benkard, Berry, and Pakes, 2007, for a review). Solutions range from instrumental variable estimation to GMM estimation and the control function approach pioneered by Olley and Pakes (1996). We adopt the control function approach developed by Ackerberg, Caves, and Frazier (2006), which consists of a two-step estimation.

In Appendix B.1, we lay out details of our procedure in estimating the production function. Specifically, we use a translog specification of production function, i.e.,

$$\begin{aligned} q_{it} &= \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \beta_{ll} l_{it}^2 + \beta_{kk} k_{it}^2 + \beta_{mm} m_{it}^2 \\ &\quad + \beta_{lk} l_{it} k_{it} + \beta_{km} k_{it} m_{it} + \beta_{lm} l_{it} m_{it} \\ &\quad + \beta_{lkm} l_{it} k_{it} m_{it} + \omega_{it} + \varepsilon_{it}, \end{aligned} \tag{11}$$

where the lowercase letters represent the logarithm of the uppercase letters; $\omega_{it}$ is firm-specific productivity; and $\varepsilon_{it}$ is an i.i.d. error term.

We estimate the translog production function (11) separately for each two-digit industry. Once $\hat{\boldsymbol{\beta}} = \left(\hat{\beta}_l, \hat{\beta}_k, \hat{\beta}_m, \hat{\beta}_{ll}, \hat{\beta}_{kk}, \hat{\beta}_{mm}, \hat{\beta}_{lk}, \hat{\beta}_{km}, \hat{\beta}_{lm}, \hat{\beta}_{lkm}\right)$ is obtained, we can calculate the output elasticity of materials as $\hat{\theta}_{it}^{m} = \hat{\beta}_m + 2\hat{\beta}_{mm} m_{it} + \hat{\beta}_{lm} l_{it} + \hat{\beta}_{km} k_{it} + \hat{\beta}_{lmk} l_{it} k_{it}$, then firm markup using equation (10).

A few practical details are worth noting. First, to estimate equation (11), we use a merged dataset that contains the information on output ($q_{it}$) in physical terms and therefore avoid the omitted output price bias in the production function estimation pointed out by Klette and Griliches (1996).

Second, the estimation of equation (11) also requires three inputs ($l_{it}$, $k_{it}$, $m_{it}$) measured in physical quantity terms. Our dataset has information on employment, which allows us to measure labor input $l_{it}$ in physical terms. However, capital $k_{it}$ and material $m_{it}$ inputs are only available in value terms; specifically, we use the net value of fixed assets as a measure of $k_{it}$ and the total value of intermediate materials as a measure of $m_{it}$. To back out the physical quantity of $k_{it}$ and $m_{it}$, we deflate these values with the price indices provided by Brandt, Van Biesebroeck, and Zhang (2012). But this practice may result in estimation biases due to the omitted firm-specific input prices (see De Loecker and Goldberg, 2014, for a detailed discussion). To correct this omitted input price bias, we use a control function approach developed by De Loecker et al. (2014). Specifically, the omitted firm-specific input prices are assumed to be a reduced-form function of output prices, market shares, and exporter status,[9] and these factors are also interacted with the deflated inputs to construct a flexible control function.

Third, we focus on a group of single-product producers to avoid the potential bias caused by the multi-product producer issue. After obtaining $\hat{\boldsymbol{\beta}} = \left(\hat{\beta}_l, \hat{\beta}_k, \hat{\beta}_m, \hat{\beta}_{ll}, \hat{\beta}_{kk}, \hat{\beta}_{mm}, \hat{\beta}_{lkm}\right)$ and assuming that multi-product firms use the same technology as single-product firms in the same industry, we are able to calculate the firm-product-level markups and then average across products to get firm-level markups.

Fourth, in estimating the production function, we also control for demand and supply shocks by including output prices, 5-digit product dummies, city dummies, product market shares, exporter status, input tariffs at the industry level, and output tariffs at the industry level. For discussion of the importance of controlling for demand and supply shocks in the production function estimation, see De Loecker (2011).

# III. Data and Variables

## A. Data

The main dataset used in this study comes from the *Annual Survey of Industrial Firms* (ASIF), conducted by the National Bureau of Statistics of China for the period 1998-2005. This is the most comprehensive and representative firm-level dataset in China, and surveyed firms contribute the majority of China's industrial value-added. The dataset is used to calculate matrices in the national income account (e.g., GDP) and major statistics published in China Statistical Yearbooks. This dataset has also proved to be reasonably accurate and reliable due to the strict double-checking procedures used in data collection (Cai and Liu, 2009). Accordingly, it has been widely used by economic researchers in recent years, e.g., Lu, Lu, and Tao (2010), Brandt, Van Biesebroeck and Zhang (2012), Brandt et al. (2012).

One drawback of this dataset is that it covers all SOEs, but for non-SOEs, only those with annual sales of five million RMB (Chinese currency) or more are surveyed. Hence, it is possible that both the overall degree of markup dispersion and the effect of trade liberalization on markup dispersion are underestimated, as this is a relatively more homogeneous sample due to data truncation.

The number of firms ranges from 140,000+ in 1998 to 244,000+ in 2005. These firms are distributed among 29 two-digit (or 164 three-digit, 464 four-digit) manufacturing industries,[10] and across China's 31 provinces (including four municipalities), 344 cities, and

[9] De Loecker et al. (2014) also include region dummies and product dummies as the determinants of firm-specific input prices. However, this may increase the parameters of interest to more than 5,000, which is beyond our computational capacity in a sample with 314,421 observations.

[10] Later, we exclude Tobacco industry from our analysis as (1) there are few observations, and (2) this is a monopoly industry, protected by the government.

2,829 counties. The dataset provides detailed firm information, including industry affiliation, location, and all operations and performance items from the accounting statements such as output, intermediate materials, employment, and book value and net value of fixed assets, which are of interest to us.

During the sample period, there were several changes in China's administrative boundaries and, consequently, in the county or city codes in our data set. For example, new counties were established, while existing counties were combined into larger ones or even elevated to cities. Using the 1999 National Standards (promulgated at the end of 1998 and called GB/T 2260-1999) as the benchmark codes, we convert the regional codes of all the firms to these benchmark codes to achieve consistency in the regional codes throughout the sample period. Meanwhile, a new classification system for industry codes (GB/T 4754-2002) was adopted in 2003 to replace the old classification system (GB/T 4754-1994), which had been used from 1995 to 2002. To achieve consistency in the industry codes for the whole sample period (1998-2005), we convert the industry codes in the 1998-2002 data to the new classification system.

The dataset of Chinese tariffs is downloaded from the WTO website. Specifically, we use the *Tariff Download Facility* to obtain standardized tariff statistics. Tariff data provide, for each product defined at HS-6 digit level, detailed information on the number of tariff lines; average, minimum and maximum ad valorem tariff duties; etc. Tariff data are available for 1996, 1997 and 2001-latest. As tariff information on the WTO website is missing for 1998-2000, we supplement the missing tariff data from the World Integrated Trade Solution website maintained by the World Bank. Meanwhile, as different HS codes were used before and after 2002, we match the 1996 HS codes (also used for 1997-2001 tariffs) to the 2002 HS codes (used for 2001-2006 tariffs) using the standard HS concordance table. There are a total of 5,036 HS-6 products from manufacturing industries in our tariff data.

As our outcome variable can be only calculated at the industry level, we need to aggregate tariffs from the HS-product level to the industry level. To this end, we first match the HS classification to CIC using the concordance table from the National Bureau of Statistics of China.[11] Then, for each industry and each year, we calculate the simple average tariff. However, one could be concerned that such aggregation might conceal substantial variations in tariff reduction across products within an industry, which, in turn, could underestimate the effect of trade liberalization. To address this concern, in a robustness check, we add the interaction between our regressor of interest ($Tariff_{i2001} \times Post02_t$) with the number of products within a 3-digit industry, to check whether industries with more HS-6 products (and therefore potentially more tariff variations) behave differently from those with fewer products.

A crucial step in obtaining firm markup involves the estimation of production function,

[11] We thank Yifan Zhang for sharing this concordance table.

which requires the observation of firm-level output in physical terms. As this information is missing in the ASIF data, we use product-level data from the National Bureau of Statistics of China for the period 2000-2006, which contains information on each product (defined at the 5-digit product level) produced by the firm, and in particular, output quantity. As the product-level data and the ASIF data share the same firm identity, we can easily match the two.

## B. Output Elasticities and Firm Markups

For each two-digit manufacturing industry, we report in Appendix Table 3 estimated output elasticities of labor, capital, and materials at different quantiles (i.e., p5, p25, p50, p75, p95). It is found that materials play a dominant role in China's manufacturing production, while the role of capital is limited. However, there are some abnormalities in the estimated output elasticities of labor for a few industries. Specifically, output elasticities of labor are mostly negative in the Metal Products, General Purpose Machinery, and Electrical Machinery and Equipment industries. To ensure that our results are not driven by these three industries, we conduct a robustness check by excluding these industries.

To compare our estimation of production function (i.e., quantity-based and translog production function with adjustment for input prices, denoted as $Q-TL-IP$) to alternative approaches used in the literature, we list the output elasticities of three alternative estimations (i.e., quantity-based and translog production function, denoted as $Q-TL$; revenue-based and Cobb-Douglas production function, denoted as $R-CD$; and revenue-based and translog production function, denoted as $R-TL$; all without adjustment for input prices) in Appendix Table 4. For translog production function, we use median output elasticities in the comparison. It is found that these four production function estimations have different values and distribution of output elasticities across industries. Many output elasticities from the quantity-based and translog production function estimation without adjustment for input prices are negative; De Loecker and Goldberg (2014) argue that this is mainly due to the omitted input price bias in the production function estimation. This problem is partly avoided in the revenue-based production function estimation, similar to findings by De Loecker and Goldberg (2014).

Figure 3 displays mean markups for each two-digit manufacturing industry for 1998-2005, and the mean values as well as values of different quantiles (i.e., p5, p25, p50, p75, p95) are reported in Appendix Table 5. Average markups across industries range from 0.825 to 1.372, and most firms have markups above 1. Labor-intensive industries have low markups; for example, the four industries with average markups lower than 1 are Garment, Food Ware and Caps (0.822); Manufacture of Foods (0.825); Timber, Manufacture of Wood, Bamboo, Rattan, Palm and Straw Products (0.888); and Artwork and Other Manufacturing (0.892). And industries with large markups are Chemical

Fibers (1.372); Paper and Paper Products (1.363); Plastics (1.320); and Printing and Reproduction of Recording Media (1.305).

[Insert Figure 3 here]

We report correlations among markups from the four different production function estimations in Appendix Table 6. It is found that markups from the production function estimation used in our analysis ($Q-TL-IP$) are positively correlated with those from $Q-TL$ and $R-TL$, but are negatively correlated with those from $R-CD$. Meanwhile, correlations range from $-0.09$ to 0.31, suggesting that different production function estimations result in different estimated firm markups.

As a sanity check on our estimated firm markups, we report the correlation between mean markups and competition degree across two-digit manufacturing industries over the sample period. Specifically, we use the Herfindahl index (HHI) to characterize industry competition degree, and find a correlation of 0.2057 (with statistical significance at 1%) between these two variables. As a lower value of HHI means fiercer competition, this result indicates that markups are lower in more competitive industries, which is consistent with our intuition.

## C. Markup Dispersion

A widely used measure of dispersion in the literature is the Gini index, with the value ranging from 0 (perfect equality) to 1 (perfect inequality). While the Gini coefficient has many desirable properties (e.g., mean independence, population size independence, symmetry, and Pigou-Dalton transfer sensitivity), it suffers from the problems of decomposability and statistical testability (Cowell, 1995). As a result, a number of entropy measures have been developed to overcome these problems and reap the benefits of the Gini index. The most widely used entropy measure is the Theil index[12]—specifically,

$$Theil_{it} = \frac{1}{n_{it}} \sum_{f=1}^{n_{it}} \frac{y_{fit}}{\bar{y}_{it}} \log(\frac{y_{fit}}{\bar{y}_{it}}), \tag{12}$$

where $y_{fit}$ is the markup ratio for firm $f$ located in industry $i$ at year $t$; $\bar{y}_{it}$ is the average markup value in industry $i$ at year $t$; and $n_{it}$ is the number of firms in industry $i$ at year $t$.[13]

Given the superiority of the Theil index over the Gini index, we use the former as the

[12] We have experimented with another commonly used entropy measure, the Mean Log Deviation (i.e., $MLD_{it} = \frac{1}{n_{it}} \sum_{f=1}^{n_{it}} \log(\frac{\bar{y}_{it}}{y_{fit}})$), and find similar results (see Appendix Table 1, Column 4).

[13] To alleviate the concern that outliers may drive the degree of dispersion, we exclude the top and bottom 2.5% of markups in constructing the dispersion measures.

main measure of markup dispersion and the latter for the robustness check. Meanwhile, we have also experimented with two other dispersion measures in the robustness checks. One is the coefficient of variation (CV), defined as the ratio of the standard deviation to the mean (i.e., $CV_{it} = \frac{\sqrt{V_{it}}}{\bar{y}_{it}}$, where $V_{it}$ is the standard deviation of firm markups in industry $i$ at year $t$), and the other is the relative mean deviation (RMD), defined as the average absolute distance of each unit from the mean and expressed as a proportion of the mean (i.e., $RMD_{it} = \frac{1}{n_{it}} \sum_{f=1}^{n_{it}} \left| \frac{y_{fit}}{\bar{y}_{it}} - 1 \right|$).

Average values for these four measures at the two-digit industry level over the period 1998-2005 are reported in Appendix Table 7. Industries having the largest degree of markup dispersion are Chemical Fibers industry; and Manufacture of Foods industry. Industries with the smallest Theil values are Textiles industry; Paper and Paper Products industry; Articles for Culture, Education and Sport Activities industry; Smelting and Pressing of Ferrous Metals industry; and Metal Products industry.

In Appendix Figure 2, we further show the relation between markup dispersion (i.e., Theil index) and the pre-period mean markup level at the three-digit industry level. While there is a modest negative relation between these two variables, the overall correlation is quite noisy, indicating that they are capturing different underlying factors.

Note that in calculating markup dispersion, we implicitly assume that firms only produce in one industry (i.e., the one that the firm reports in the data). However, it could be possible that a firm produces goods in multiple industries, but we only observe one due to data limitations. This could cause two potential estimation issues: First, our outcome variable could be measured with errors; second, our estimation might ignore the effect of trade liberalization from other industries in which firms have production but do not report in the data. To check whether our estimates are biased due to a multiple-industries issue, we first conduct a robustness check at the 2-digit industry level, in which the incidence of this is less severe. Moreover, we use the firm-product merged data to determine whether a firm produces goods in different 3-digit industries (the classification level used in our main analysis). In a robustness check, we focus on a subsample of firms producing all goods within only one three-digit industry.

# IV. Empirical Findings

## A. Graphical Results

To illustrate our identification strategy, we plot, in Figure 4, time trends of markup dispersion (measured by the Theil index) for high-tariff industries (i.e., industries with tariffs above the sample median in 2001, or our treatment group) and low-tariff industries (i.e., industries with tariffs below the sample median in 2001, or our control group), conditional on industry dummies.

[Insert Figure 4 here]

It is clear that in the pre-WTO period, the two groups have quite similar trends. This parallel pretreatment trend in markup dispersion between treatment and control groups alleviates the concern that our treatment and control groups are *ex ante* incomparable, which lends support to the satisfaction of our DID identifying assumption.

Meanwhile, there is visible divergence in trends of markup dispersion after 2002, when China started to reduce tariffs upon WTO accession. The consistency in timing between the divergence in markup dispersion and WTO accession suggests that trade liberalization reduces dispersion of firm markups.

## B. Main Results

Regression results for the DID specification (1) are reported in Table 1. We start with a simple DID specification that includes only industry and year fixed effects in Column 1. Our regressor of interest, $Tariff_{i2001} \cdot Post02_t$, is statistically significant and negative, suggesting that markup dispersion decreased more after 2002 in industries with higher tariffs in 2001 than in industries with lower tariffs in 2001. Given that industries with higher tariffs in 2001 experienced greater tariff reduction after 2002, these results imply that trade liberalization reduces markup dispersion.

[Insert Table 1 here]

In Column 2, we add some time-varying industry characteristics that may correlate with both our outcome variable (markup dispersion) and our regressor of interest (trade liberalization). Specifically, we use the EG index to measure industrial concentration degree, which may affect firm markups on the one hand and respond to trade liberalization on the other hand (e.g., Hanson (1998)). The mean value of fixed assets and the number of firms in each industry are used to capture the degree of entry barriers, which may be affected by trade liberalization and also affect the distribution of firm markups. Evidently, our results are found to be robust to these additional controls.

One could be concerned that tariffs in 2001 were not randomly determined, and hence our treatment and control groups could be systematically different *ex ante*, which may spuriously generate the negative effect of trade liberalization on markup dispersion. However, as displayed in Figure 4, markup dispersion degrees in high-2001-tariff industries and in low-2001-tariff industries have similar time trends in the pre-WTO period and start to diverge upon WTO accession, implying that our treatment and control groups are largely comparable. To further alleviate the concern that the nonrandom determination of tariffs in 2001 could bias our estimates, we conduct a robustness check following Gentzkow (2006). Specifically, in Appendix A, we first identify what are the important determinants of tariffs in 2001. As shown in Appendix Table 2, three determinants are found to be robustly statistically significant: (1) output share of SOEs has a positive effect, consistent with the story that politically connected SOEs are protected by the government; (2) average wage per worker has a positive effect; and (3) export intensity has a positive effect, implying an export-promotion industrial policy. Meanwhile, conditional on these potential tariff determinants, we find that China's industrial tariff structure is not affected by markup dispersion, thereby relieving the concern about reverse causality.

We then add interactions between these significant tariff determinants with the post-WTO indicator to control for flexible time trends in markup dispersion generated by these significant tariff determinants. As shown in Column 3 of Table 1, the coefficient of our regressor of interest remains negative and statistically significant, and magnitude also barely changes.

## C. Checks on the Identifying Assumption

In this subsection, we report results of a battery of robustness checks on the identifying assumption of our aforementioned DID estimation.

*Expectation Effect.*—In Column 1 of Table 2, we add to the regression an additional control, $Tariff_{i2001} \times One\ Year\ Before\ WTO\ Accession_t$, to check whether firms changed their behavior (and thereby markup dispersion changed) in anticipation of the coming WTO accession, which may in turn make our treatment and control groups *ex ante* noncomparable and bias our estimates. The coefficient of $Tariff_{i2001} \times One\ Year\ Before\ WTO\ Accession_t$ is found to be statistically insignificant, suggesting little expectation effect. Moreover, the coefficient of our regressor of interest remains negative and statistically significant.

[Insert Table 2 here]

*Control for Other Policy Reforms.*—To control for the two ongoing policy reforms in the early 2000s (SOEs reform and the relaxation of some FDI regulations), we add two control variables (i.e., the share of SOEs among domestic firms, and the number of foreign-invested firms) in Column 2 of Table 2. Our main findings remain robust to these additional controls.

In addition, WTO accession is multilateral and multidimensional; that is, China's trading partners may also reduce their tariffs on Chinese imports. To fix the idea that the change in markup dispersion comes from the increase in domestic competition degree generated by tariff reduction, we additionally include total exports (to control for access to foreign markets) and input tariffs at the industry level (to control for the use of foreign intermediate inputs). Regression results are reported in Column 3 of Table 2. Clearly, our main findings remain robust to these additional controls, lending support to the argument regarding import competition.

*Placebo Test I: Pre-WTO Period.*—As a placebo test, we follow Topalova (2010) in looking at the effect of tariffs on markup dispersion in the pre-WTO period (i.e., 1998-2001). The premise is that because tariffs did not change much during this period,[14] we should not expect any significant effects; otherwise, that could indicate the existence of some underlying confounding factors.[15] As shown in Column 4 of Table 2, we indeed find tariffs have almost zero effect on markup dispersion in the pre-WTO period.

*Placebo Test II: A Sample of Processing Traders.*—A unique feature of the Chinese trade regime is that some firms are allowed to import materials free of tariffs but required to export their entire output—the so-called "processing trade regime." This policy was meant not only to protect a fragile domestic economy from foreign competition but also to open the economy when the Chinese government adopted its "reform and opening" policy in 1978. Given that processing traders are relatively immune from the liberalization caused by WTO accession, the estimation using the sample of processing traders should show insignificant liberalization effect. Regression results are reported in Column 5 of Table 2. As expected, we find the coefficient of $Tariff_{i2001} \times Post02_t$ is highly insignificant and small in magnitude.

## D. Other Robustness Checks

In this subsection, we present another series of robustness checks on other econometric concerns. Regression results are reported in Table 3.

[Insert Table 3 here]

*Alternative Measures of Markup Dispersion.*—In Columns 1-3, we experiment with three alternative measures of markup dispersion: Gini index, CV, and RMD. We find that $Tariff_{i2001} \times Post02_t$ has consistently negative and statistically significant coefficients, implying that our aforementioned results are not driven by any specific dispersion measure.

*Finer Industry Definition.*—Thus far, our analysis has been based on the three-digit CIC industry level. To alleviate concerns about any aggregation bias, we conduct a

[14] The correlation between tariffs in 1997 and in 2001 is 0.95.

[15] Formally, assume $\varepsilon_{it} = \delta\omega_{it} + \tilde{\varepsilon}_{it}$ such that $cov\left[Tariff_{i2001} \times Post02_t, \omega_{it}|\mathbf{W}_{it}\right] \neq 0$, and $cov\left[Tariff_{i2001} \times Post02_t, \tilde{\varepsilon}_{it}|\mathbf{W}_{it}\right] = 0$, where $\mathbf{W}_{it}$ summate all the other controls. In other words, the identification problem comes from the omitted variable $\omega_{it}$. Hence, $\hat{\beta} = \beta + \delta\theta$, where $\theta \equiv \frac{cov[Tariff_{i2001} \times Post02_t, \omega_{it}|\mathbf{W}_{it}]}{var[Tariff_{i2001} \times Post02_t|\mathbf{W}_{it}]}$, and $\hat{\beta} \neq \beta$ if $\delta\theta \neq 0$. We now replace $Tariff_{i2001} \times Post02_t$ with $Tariff_{it}$, and estimate the equation for the pre-WTO period. Given that $Tariff_{it}$ barely changed in this period, its effect ($\beta$) is close to zero. Meanwhile, if our equation (1) is well specified such that $\delta = 0$, the estimator of $Tariff_{it}$ shall then be zero.

robustness check at the four-digit CIC industry level (note that a trade-off is that there are fewer observations within each industry-year cell and, therefore, potential measurement errors in the dispersion variable). Regression results are reported in Column 4. Clearly, our aforementioned results are robust to this finer industry definition.

*Check on Cross-Product, Within-Industry Tariff Variations.*—As noted in Section 4.1, one drawback of our data is that tariff information is at the HS-6 product level, while our markup dispersion calculation is at the three-digit CIC industry level. Hence, mapping from the HS-6 product to the three-digit CIC industry level might conceal variations in tariff reduction across different HS-6 products but within the same three-digit industry, which could lead to underestimation of our trade liberalization effect. As a check on this issue, we add an interaction between our regressor of interest ($Tariff_{i2001} \times Post02_t$) and the number of products within a three-digit industry. As shown in Column 5, the triple interaction term is not statistically significant, implying that industries with more HS-6 products (and therefore potentially more variations in tariffs within the industry) does not behave differently from those with fewer products.

*Checks on the Multi-Industry Issue.*—One could be concerned that as firms produce multiple products spanning different three-digit industries, our aforementioned DID estimation could miss the liberalization effect from other related three-digit industries. To check this, we first investigate the effect at the two-digit industry level, where the multi-product issue is less severe. As shown in Column 6, we still find a negative, albeit imprecisely estimated, effect of trade liberalization on markup dispersion. Meanwhile, in Column 7, we focus on a subsample of firms that produce in only one three-digit industry, and continue to find a negative and statistically significant effect of trade liberalization.

*A Sample of Non-Exporters.*—Our data include many exporters, and hence their markups could also reflect the conditions of foreign markets. To check whether our results are driven by changes in foreign markets, we focus on the sample of non-exporters. Regression results are reported in Column 8. Clearly, our findings remain robust to the sample of non-exporters, alleviating concerns about any complications due to foreign markets.

*Two Periods Estimation.*—One concern with the DID estimation is how to accurately calculate standard errors and, in turn, statistical inference. Thus far, we have followed the suggestion by Bertrand, Duflo, and Mullainathan (2004) to cluster standard errors at the industry level. As a robustness check, we use another approach suggested by Bertrand, Duflo, and Mullainathan (2004), which is to collapse the panel structure into two periods, one before and the other after WTO accession, then use the White-robust standard errors. Meanwhile, this exercise allows us to compare the long-run average effect of trade liberalization on markup dispersion. Regression results are reported in Column 9, and show similar results.

*Exclusion of Industries with Abnormal Estimated Output Elasticities.*—As shown in

Appendix Table 3, three two-digit industries have negative estimated output elasticities of labor. To address the concern that our results might be driven by these industries, we exclude them and repeat our analysis. As shown in Column 10, our results are robust to the exclusion of these industries, suggesting that our findings are not driven by industries with abnormal estimated output elasticities.

## E. Discussion

We have established that trade liberalization reduces the dispersion of firm markups within a narrowly defined industry, which is an important step for the pro-competitive role of trade. In this subsection, we provide further evidence to support and understand this allocative efficiency channel of trade liberalization.

We first summarize the domestic market structures for Chinese two-digit industries before and after WTO accession as well as their changes during this period. Specifically, we use HHI to characterize overall competition degree and EG index to capture the spatial structure. Appendix Figure 3 reports the average HHI before the WTO accession and change in the HHI after the WTO accession, and Appendix Table 8 contains detailed summary statistics. Most Chinese manufacturing industries were already quite competitive before WTO accession; for example, the average HHI between 1998 and 2001 was 0.0171, and 13 out of 28 two-digit industries had HHI values below 0.01. A majority of manufacturing industries (24 out of 28) experienced decreases in HHI or increases in competition degree after WTO accession, and some industries even saw their HHI values drop by more than 30%.

Appendix Figure 4 displays average EG indices before WTO accession and their changes after accession. Manufacturing industries in China were quite dispersed across the space, with an average EG index of 0.0118 before WTO accession (numbers in the U.S. were 0.039 in 1972, 0.039 in 1977, 0.038 in 1982, 0.036 in 1987 and 0.034 in 1992; see Dumais, Ellison, and Glaeser (2002)). WTO accession largely increased the geographic concentration of Chinese manufacturing industries: EG index values for all but one industry increased after WTO accession. As the geographic concentration is found to increase local competition and productivity (for a review, see Melo, Graham, and Noland, 2009), trade liberalization also intensifies domestic market competition through the geographic location of production.

We further investigate whether imports increase in response to tariff reduction, which is direct evidence of the competition effect of trade. With both import and tariff information available at the HS-6 product level, we investigate import response to trade liberalization at the product level. However, there are many HS-6 product categories with zero import values, which creates a potential estimation bias (i.e., the sample selection issue). To correct for this zero-trade-value issue, we use the Poisson pseudo maximum likelihood estimation by Silva and Tenreyro (2006). Specifically, we regress the level of imports on our regressor of interest (i.e., $Tariff_{p2001} \times Post02_t$, where $Tariff_{p2001}$ is the tariff of product $p$ in 2001), along with a set of product and year dummies, using the Poisson estimation. Regression results are reported in Column 1 of Table 4. We find that imports increase in product categories experiencing more tariff reduction, corroborating our import-competition argument.

[Insert Table 4 here]

To further understand how trade liberalization changes markup dispersion, we look at the response of markup distribution at the different quantiles, specifically, p5, p25, p50, p75, p95 and the mean level. Regression results are summarized in Columns 2-7 of Table 4. Trade liberalization increases markups at the lower quantiles but reduces markups at higher quantiles, thereby flattening markup distribution. We also find an insignificant effect of trade liberalization on the mean markup level, which is different from Brandt et al. (2012)'s (i.e., they find a positive and significant effect of tariff reduction on mean markups).[16] If entry and exit mostly occurs at the lower end, these results suggest that firm selection induced by trade liberalization improves markups at lower quantiles. If large firms exist mostly at the higher end, these results also imply that competition from trade liberalization negatively affects larger firms, consistent with the findings by di Giovanni and Levchenko (2013).

Note that the markup measure contains both price and cost information, and hence the effect of trade liberalization on the dispersion of firm markups can operate through price changes, cost changes, or both. To further understand the underlying mechanisms, we conduct two analyses, each having its own pros and cons due to data limitations. First, we use the ASIF data, which has a large coverage of firms but no information on product prices. Instead, we calculate productivity for each firm and each year based on the estimation of production functions, and use firm productivity as a proxy for firms' marginal costs. We then use the dispersion of firm productivity to investigate the cost-change channel of the liberalization effect on the dispersion of firm markups. Meanwhile,

[16] There are two main differences between these two studies—(1) markup estimation: we use quantity-based and translog production function with adjustment for firm-specific input prices, while Brandt et al. (2012) use revenue-based and Cobb-Douglas production function without adjustment for input prices; (2) regression specification: we use $Tariff_{i2001} \times Post02_t$ as the regressor of interest along with a set of controls and estimate using the fixed effect approach, whereas Brandt et al. (2012) use $Tariff_{it}$ as the regressor of interest with year dummies and estimate in the first-difference approach. To further understand which drives the different findings, we conduct two experiments. First, we use mean markups from Brandt et al. (2012)'s production function estimation in our regression specification, and also find a negative but insignficant effect (i.e., the coefficient is $-0.042$ with a standard error of 0.076). Second, we use our estimated mean markups in Brandt et al. (2012)'s regession specification, and also find a negative but insignificant effect (i.e., the coefficient for one-year change is $-0.057$ with a standard error of 0.048). These results suggest that two differences both play an important role in generating different findings between ours and Brandt et al. (2012)'s.

we control for productivity dispersion in the regression of markup dispersion on trade liberalization to partially isolate the price-change channel of the liberalization effect. Regression results are reported in Columns 1-2 of Table 5. Trade liberalization significantly reduces the dispersion of firm productivity, suggesting the response of costs to trade liberalization. Meanwhile, we continue to find a significant effect of trade liberalization on markup dispersion, after controlling for the dispersion of firm productivity, suggesting that prices are also responsive to trade liberalization.

[Insert Table 5 here]

Second, we use the sample of single-product firms in the merged product-ASIF data, which contains information on output quantity and revenue, and therefore enables us to calculate product price. With estimated firm markup, we are then able to back out marginal cost for each firm and each year (similar to an approach used by De Loecker, Goldberg, Khandelwal, and Pavcnik, 2014). However, a drawback of this analysis is that we are only able to do it for a particular group of firms—that is, single-product firms—for the period 2000-2005, and this raises external validity issues. Regression results using price dispersion and marginal cost dispersion as the outcome variables are reported in Columns 3-4 of Table 5, respectively. We find that trade liberalization has both negative and statistically significant effects on these two outcomes. Combined, these two exercises suggest that both price and cost channels work for the liberalization effect on the dispersion of firm markups.

## F. Heterogeneous Effects

Our aforementioned analyses estimate the average effect of trade liberalization on the dispersion of firm markups across Chinese manufacturing industries. In this subsection, we investigate the heterogeneous effects of trade liberalization on the dispersion of firm markups across firms and regions, to further shed light on how markup dispersion is affected by trade liberalization.

First, as shown in Arkolakis et al. (2012), surviving firms and new entries/exiters respond differently to changes in trade costs. In their model setup, changes in these two groups exactly cancel each other out, generating the unresponsiveness of the dispersion of firm markups to trade costs. Following this argument, we divide firms into two groups: surviving firms (i.e., firms present in our data both before and after WTO accession) and new entries/exiters (i.e., firms that exited or entered our data after WTO accession). Note that our data are truncated; that is, for non-SOEs, only those with annual sales of five million RMB or more are surveyed. This could mean that after WTO accession, SOEs newly entered or exited markets, or non-SOEs shrank annual sales to less than or increased annual sales to more than five million RMB. Regression results using these two groups are reported in Columns 1-2 of Table 6. There is a negative and significant effect of trade liberalization on dispersion of markups among new entries and exiters, but an insignificant effect among surviving firms. These results suggest that much of the liberalization effect on markup dispersion stems from different markups among new entries and exiters.

[Insert Table 6 here]

Second, Holmes, Hsu, and Lee (2014) show that there is a diminishing effect of trade liberalization on allocative efficiency or markup dispersion. This pattern has been confirmed by Edmon, Midrigan, and Xu (2014) and Hsu, Lu, and Wu (2014) in analyses using Taiwanese and Chinese data, respectively. Specifically, the diminishing effect implies that when the market is already competitive—and therefore there is low dispersion of firm markups before trade liberalization—the liberalization effect on markup dispersion is smaller than in a case with a more monopolized pre-liberalization setting. The reasoning is that the markup has a lower bound at 1, and when there is less competition—and therefore more dispersion—there is more room for competition to decrease markup dispersion. Following this argument, we conduct two exercises. In the first, we divide firms into SOEs and non-SOEs.[17] In China, SOEs enjoy various governmental protections, e.g., restrictions on market entry and privileged access to subsidized credit (for anecdotal evidence, see Li, Liu, and Wang (2012)), whereas non-SOEs face market discrimination and huge competitive pressure. As a result, SOEs largely encounter fewer challenges than non-SOEs (Du et al. (2014)). Regression results using subsamples of SOEs and non-SOEs are reported in Columns 3-4 of Table 6, respectively. The effect of trade liberalization on markup dispersion is larger for SOEs than non-SOEs.

In our second exercise, we divide firms based on location—specifically, coastal versus inland cities. When China opened its borders to overseas investors in 1978, access to domestic markets was restricted in the coastal regions, through the establishment of a series of special economic zones. In addition, due to better infrastructures and geographic features, markets in coastal regions have remained more open and more competitive than those in inland regions in the past decades. Regression results using subsamples of coastal and inland regions are reported in Columns 5-6 of Table 6, respectively. The effect in inland regions is bigger than that in coastal regions. Combined, these two analyses imply that distribution of firm markups becomes relatively less dispersed in response to trade liberalization than when competition was fiercer before liberalization.

[17] Classification of SOEs follows the one used by Hsieh and Song (2013). Specifically, a firm is classified as an SOE if it satisfies one of two conditions: (a) the registered capital held directly by the state exceeds 50 percent, or (b) ASIF data identify the state as the controlling shareholder of the firm.

# V. Conclusion

Resource misallocation has recently been the focus of attempts to understand why there are substantial differences in productivity across countries. In this paper, we look at one important source of resource misallocation—product market distortions and specifically markup dispersion—and investigate whether trade liberalization can reduce markup dispersion.

For empirical estimation, we first apply the methodology developed by De Loecker and Warzynski (2012) to Chinese firm-level data to recover firm markups, then use China's WTO accession as an identification strategy. Our results indicate that the distribution of firm markups becomes flattened after trade liberalization. This finding is robust to a battery of checks on the identifying assumption and other econometric concerns.

Our study also contributes to recent literature on gains from trade. While these studies focus on productive efficiency gains from trade, we study another potential channel—change in markup dispersion—through which free trade can benefit a nation. However, calculation of overall gains from trade (and through different channels, including the change in markup dispersion) requires a structural approach (e.g., Edmond, Midrigan, and Xu (2014), Hsu, Lu, and Wu (2014)), which is beyond the scope of this study.

# References


[1] Ackerberg, Daniel, C. Lanier Benkard, Steven Berry, and Ariel Pakes. 2007. "Econometric Tools for Analyzing Market Outcomes" in *Handbook of Econometrics*, Volume 6A, ed. James J. Heckman and Edward E. Leamer, 4171-276. Amsterdam and Boston: Elsevier, North-Holland

[2] Ackerberg, Daniel, Kevin Caves, and Garth Frazer. 2006. "Structural Identification of Production Functions", mimeo, UCLA

[3] Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012. "New Trade Models, Same Old Gains?", *American Economic Review* 102, 94-130

[4] Arkolakis, Costas, Arnaud Costinot, Dave Donaldson, and Andrés Rodríguez-Clare. 2012. "The Elusive Pro-Competitive Effects of Trade", working paper

[5] Banerjee, Abhijit V., and Esther Duflo. 2005. "Growth Theory through the Lens of Development Economics", In *Handbook of Economic Growth*, edited by Philippe Aghion and Steven Durlauf, Amsterdam: Elsevier Science

[6] Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. 2013. "Cross-Country Differences in Productivity: The Role of Allocation and Selection", *American Economic Review* 103, 305-334

[7] Bernard, Andrew, Jonathan Eaton, J. Bradford Jensen, and Samuel S. Kortum. 2003. "Plants and Productivity in International Trade", *American Economic Review* 93, 1268-1290

[8] Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?", *Quarterly Journal of Economics* 19, 249-275

[9] Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang. 2012. "Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing", *Journal of Development Economics* 97, 339-351

[10] Brandt, Loren, Johannes Van Biesebroeck, Luhang Wang, and Yifan Zhang. 2012. "WTO Accession and Performance of Chinese Manufacturing Firms", working paper

[11] Cai, Hongbin and Qiao Liu. 2009. "Competition and Corporate Tax Avoidance: Evidence from Chinese Industrial Firms", *Economic Journal* 119, 764-795

[12] Chen, Natalie, Jean Imbs, and Andrew Scott. 2009 "The Dynamics of Trade and Competition", *Journal of International Economics* 77, 50–62

[13] Cowell, Frank A. 1995. *Measuring Inequality* (2nd edition), Prentice-Hall/Harvester-Wheatsheaf, Hemel Hempstead

[14] de Blas, Beatriz and Katheryn Russ. 2012. "Understanding Markups in the Open Economy", working paper

[15] De Loecker, J. 2011. "Product Differentiation, Multi-Product Firms and Estimating the Impact of Trade Liberalization on Productivity", *Econometrica* 79, 1407–51

[16] De Loecker, Jan and Pinelopi K. Goldberg. 2014. "Firm Performance in a Global Market", *Annual Review of Economics* 6, 201-227

[17] De Loecker, Jan, Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik. 2014. "Prices, Markups and Trade Reform", working paper

[18] De Loecker, Jan and Frederic Warzynski. 2012. "Markups and Firm-level Export Status", *American Economic Review* 102, 2437-2471

[19] di Giovanni, Julian and Andrei A. Levchenko. 2013. "Firm Entry, Trade, and Welfare in Zipf's World", *Journal of International Economics* 89, 283-296

[20] di Giovanni, Julian, Andrei A. Levchenko, and Jing Zhang. 2014. "The Global Welfare Impact of China: Trade Integration and Technological Change", *American Economic Journal: Macroeconomics* 6, 153-183

[21] Du, Julan, Yi Lu, Zhigang Tao, and Linhui Yu. 2014. "The Nature of Chinese-Style Capitalism: the Twin Costs of State Ownership in China", working paper

[22] Dumais, Guy, Glenn Ellison, and Edward L. Glaeser. 2002. "Geographic Concentration As A Dynamic Process", *Review of Economics and Statistics* 84, 193-204

[23] Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2014. "Competition, Markups, and the Gains from International Trade", working paper

[24] Ellison, Glenn and Edward L. Glaeser. 1997. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach", *Journal of Political Economy* 105, 889-927

[25] Gentzkow, Matthew. 2006. "Television and Voter Turnout", *Quarterly Journal of Economics* 121, 931-972

[26] Guadalupe, Maria, and Julie Wulf. 2010. "The Flattening Firm and Product Market Competition: The Effect of Trade Liberalization on Corporate Hierarchies", *American Economic Journal: Applied Economics* 2, 105-127

[27] Hanson, Gordon H. 1998. "Regional Adjustment to Trade Liberalization", *Regional Science and Urban Economics* 28, 419-444

[28] Harrison, Ann E. 1994. "Productivity, Imperfect Competition and Trade Reform. Theory and Evidence", *Journal of International Economics* 36, 53–73

[29] Holmes, Thomas J., Wen-tai Hsu, and Sanghoon Lee. 2014. "Allocative Efficiency, Mark-ups, and Welfare Gains from Trade", *Journal of International Economics* 94, 195-206

[30] Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India", *Quarterly Journal of Economics* 124, 1403-1448

[31] Hsieh, Chang-Tai and Zheng Song. 2013. "Grasp the Large, Let Go of the Small: The Transformation of the State Sector in China", working paper

[32] Hsu, Wen-tai, Yi Lu, and Guiying Wu. 2014. "An Anatomy of Welfare Effects of Trade Liberalization: The Case of China's Entry to WTO", working paper

[33] Klette, T.J. and Z. Griliches. 1996. "The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous", *Journal of Applied Econometrics* 11, 343-361

[34] Konings, Joep, Patrick Van Cayseele, and Frederic Warzynski. 2001. "The Dynamics of Industrial Mark-Ups in Two Small Open Economies: Does National Competition Policy Matter?", *International Journal of Industrial Organization* 19, 841–859

[35] Krishna, Pravin and Devashish Mitra. 1998. "Trade Liberalization, Market Discipline and Productivity Growth: New Evidence from India", *Journal of Development Economics* 56, 447–462

[36] Lardy, Nicholas R. 2002. *Integrating China into the Global Economy*, Washington, U.S.: Brookings Institution Press

[37] Lerner, A. P., 1934, "The Concept of Monopoly and the Measurement of Monopoly Power", *Review of Economic Studies* 1, 157-175

[38] Levinsohn, James. 1993. "Testing the Imports-as-Market-Discipline Hypothesis", *Journal of International Economics* 35, 1–22

[39] Li, Xi, Xuewen Liu, and Yong Wang. 2012. "A Model of China's State Capitalism", working paper

[40] Lipsey, R. G. and Kelvin Lancaster. 1956-1957. "The General Theory of Second Best", *Review of Economic Studies* 24, 11-32

[41] Liu, Runjuan and Daniel Trefler. 2011. "A Sorted Tale of Globalization: White Collar Jobs and the Rise of Service Offshoring", NBER working paper no. 17559

[42] Lu, Jiangyong, Yi Lu, and Zhigang Tao. 2010. "Exporting Behavior of Foreign Affiliates:Theory and Evidence", *Journal of International Economics* 81, 197-205

[43] Melo, Patricia C., Daniel J. Graham, and Robert B. Noland. 2009. "A Meta-Analysis of Estimates of Urban Agglomeration Economies", *Regional Science and Urban Economics* 39, 332-342

[44] Midrigan, Virgiliu and Daniel Yi Xu. 2014. "Finance and Misallocation: Evidence from Plant-level Data", *American Economic Review* 104, 422-458

[45] Olley, G. Steven and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry", *Econometrica* 64, 1263-1298

[46] Opp, Marcus M., Christine A. Parlour, and Johan Walden. 2014. "Markup Cycles, Dynamic Misallocation, and Amplification", working paper

[47] Restuccia, Diego and Richard Rogerson. 2008. "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments", *Review of Economic Dynamics* 11, 707-720

[48] Robinson, Joan. 1934. *The Economics of Imperfect Competition*, Macmillan

[49] Silva, J.M.C. Santos and Silvana Tenreyro. 2006. "The Log of Gravity", *Review of Economics and Statistics* 88, 641-658

[50] Stock, James H. and Mark W. Watson. 2012. *Introduction to Econometrics*, 3rd Edition, Essex, England: Pearson Education Limited

[51] Topalova, Petia. 2010. "Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India", *American Economic Journal: Applied Economics* 2, 1-41

# Appendix

## A. Tariff Determinants

Our identification uses tariff variations across industries in the pre-WTO period, and hence variations in tariff reductions after WTO accession. However, it must be recognized that China's tariff structure before its WTO accession was not randomly determined. Therefore, an understanding of how the pre-WTO tariffs were determined is important to pinpoint potential biases in our DID estimation (i.e., the comparability between our treatment and control groups) and to attribute the change in markup dispersion to trade liberalization.

There are many reasons why the government imposes different tariffs in different industries. According to the political economy literature (e.g., Grossman and Helpman (1994)), industries with more political power are more capable in lobbying and influencing governments for more protection. In the case of China, SOEs are known to conduct businesses under the auspices of governments, and in some circumstances, are cash cows for local governments. Meanwhile, employment is always at the top list of the government's agenda, as it is related to social stability. For example, during the financial crisis in 2008-2009, President Hu Jintao announced publicly that SOEs could not lay off their employees and should instead try to expand labor employment. Thus, to capture such political considerations, we use four variables: output share of SOEs, output share of other domestic firms, total employment (in log), and employment growth rate in past years.

Another important set of tariff determinants is economic factors. For example, governments may protect infant industries to allow enough time for development. Meanwhile, as China is largely a labor-abundant and technologically underdeveloped country, it is expected that government may protect labor-intensive and/or technologically backward industries. To characterize these economic considerations, we use four variables: average wage per worker (in log), capital-labor ratio, value-added ratio, and industry age.

The choice of tariff structure could also reflect the government's industrial policies; for example, import substitution versus export promotion. To capture such industrial policies, we use export intensity (measured as the ratio of total exports to total output). Finally, it is important to check whether the tariff structure is intended to preserve the distribution of firm markups or reverse causality.

Regression results are reported in Appendix Table 2, in which industry-level tariffs in 2001 are regressed on the aforementioned potential determinants, with level variables being measured in 2001 and growth variables being measured in the period 1998-2001. Three variables are found to be robustly statistically significant: (1) output share of SOEs is found to have a positive effect, consistent with the story of the protection of politically connected SOEs; (2) average wage per worker is found to have a positive effect; and (3) export intensity is found to have a positive effect, implying an export-promotion industrial policy.

Moreover, Columns 4-7 show that none of the four alternative measures of markup dispersion is statistically significant and the $t$-statistics are very small. These results suggest that conditional on potential tariff determinants, China's industrial tariff structure is not reversely affected by markup dispersion.

# B. Production Function Estimation

In this appendix, we provide the details of how we estimate the production function (11) and compare our estimation with other methods used in the literature.

## B.1. Quantity-Based Production Function Estimation with Adjustment for Input Prices

We rewrite production function (3) as

$$q_{it} = f_{it}(\mathbf{x}_{it}; \boldsymbol{\beta}) + \omega_{it} + \varepsilon_{it}, \tag{13}$$

where $\mathbf{x}_{it}$ is the vector of (log) physical inputs, specifically, $l_{it}$, $k_{it}$, $m_{it}$; $\boldsymbol{\beta}$ is the vector of production function coefficients to be estimated; $\omega_{it}$ is firm-specific productivity; and $\varepsilon_{it}$ is an i.i.d. error term.

A practical issue in estimating equation (13) is that both output ($q_{it}$) and three inputs ($l_{it}$, $k_{it}$, $m_{it}$) shall be in physical quantity terms. To this end, we use the merged product-ASIF data, which provide the physical quantity of output $q_{it}$. Meanwhile, the ASIF data have information on employment, which allows us to measure labor input $l_{it}$ in physical quantity. However, capital $k_{it}$ and material $m_{it}$ inputs are only available in value terms; specifically, we use the net value of fixed assets as a measure of $k_{it}$ and the total value of intermediate materials as a measure of $m_{it}$. To back out the physical quantity of $k_{it}$ and $m_{it}$, we deflate these values with the price indices provided by Brandt, Van Biesebroeck and Zhang (2012). In other words, the true estimation specification of equation (13) is

$$q_{it} = f_{it}(\tilde{\mathbf{x}}_{it}; \boldsymbol{\beta}) + B(\mathbf{w}_{it}, \tilde{\mathbf{x}}_{it}, \boldsymbol{\beta}) + \omega_{it} + \varepsilon_{it}, \tag{14}$$

where $\tilde{\mathbf{x}}_{it}$ is the vector of (log) deflated inputs; and $\mathbf{w}_{it}$ is the vector of firm-specific input prices. Hence, consistent estimation of $\boldsymbol{\beta}$ requires the proper control for unobserved firm productivity $\omega_{it}$ and the omitted firm-specific input prices $B(\mathbf{w}_{it}, \tilde{\mathbf{x}}_{it}, \boldsymbol{\beta})$.

To proxy $\omega_{it}$, Levinsohn and Petrin (2003) assume that

$$m_{it} = m_t \left( k_{it}, \omega_{it}, \mathbf{Z}_{it} \right),$$

where $\mathbf{Z}_{it}$ is a vector of controls including output price ($p_{it}$), 5-digit product dummies, city dummies, product market share ($ms_{it}$), exporter status ($e_{it}$), input tariff, and output tariff. Given the monotonicity of $m_t(.)$, we can have

$$\omega_{it} = h_t\left(m_{it}, k_{it}, \mathbf{Z}_{it}\right).$$

To control for omitted firm-specific input prices, we follow De Loecker, Goldberg, Khandelwal, and Pavcnik (2014) by assuming that firm-specific input prices $\mathbf{w}_{it}$ are a function of output price, market share, and exporter status, i.e.,

$$\mathbf{w}_{it} = w_t\left(p_{it}, ms_{it}, e_{it}\right).$$

Then, the control function $B(\mathbf{w}_{it}, \tilde{\mathbf{x}}_{it}, \boldsymbol{\beta})$ can be written as

$$B(\mathbf{w}_{it}, \tilde{\mathbf{x}}_{it}, \boldsymbol{\beta}) = B\left((p_{it}, ms_{it}, e_{it}) \times \tilde{\mathbf{x}}^c_{it}; \boldsymbol{\beta}, \boldsymbol{\delta}\right),$$

where $\tilde{\mathbf{x}}^c_{it} = \{1, \tilde{\mathbf{x}}_{it}\}$; and $\boldsymbol{\delta}$ is an additional parameter vector to be estimated.

In the first stage, we estimate the following equation

$$q_{it} = \phi_{it} + \varepsilon_{it},$$

where

$$\phi_{it} = f_{it}(\tilde{\mathbf{x}}_{it}; \boldsymbol{\beta}) + B\left((p_{it}, ms_{it}, e_{it}) \times \tilde{\mathbf{x}}^c_{it}; \boldsymbol{\beta}, \boldsymbol{\delta}\right) + \omega_{it},$$

and obtain estimates of the expected output ($\hat{\phi}_{it}$) and the error term ($\hat{\varepsilon}_{it}$).

Meanwhile, to recover all the production function coefficients $\boldsymbol{\beta}$ in the second stage, we model that firm productivity follows a first-order Markov movement, i.e.,

$$\omega_{it} = g_t\left(\omega_{it-1}\right) + \xi_{it},$$

where $\xi_{it}$ is an idiosyncratic shock.

From the first stage, the productivity $\omega_{it}\left(\boldsymbol{\beta}, \boldsymbol{\delta}\right)$ can be computed as

$$\omega_{it}\left(\boldsymbol{\beta}, \boldsymbol{\delta}\right) = \hat{\phi}_{it} - f_{it}(\tilde{\mathbf{x}}_{it}; \boldsymbol{\beta}) - B\left((p_{it}, ms_{it}, e_{it}) \times \tilde{\mathbf{x}}^c_{it}; \boldsymbol{\beta}, \boldsymbol{\delta}\right).$$

Then the idiosyncratic shock to productivity given $\boldsymbol{\beta}$, $\xi_{it}\left(\boldsymbol{\beta}, \boldsymbol{\delta}\right)$ can be obtained through a nonparametric regression of $\omega_{it}\left(\boldsymbol{\beta}, \boldsymbol{\delta}\right)$ on $\omega_{it-1}\left(\boldsymbol{\beta}, \boldsymbol{\delta}\right)$. Finally, the moment conditions used to estimate the parameters are

$$E\left(\xi_{it}\left(\boldsymbol{\beta}, \boldsymbol{\delta}\right) \mathbf{Y}_{it}\right) = 0.$$

In constructing the moment conditions, we follow the literature by assuming that capital is determined one period beforehand, and hence its current value is used in the moments. Meanwhile, wage rates and prices of intermediate materials are assumed to vary across firms and be serially correlated, as a result of which lagged labor and lagged materials are used in the moments. Moreover, we follow De Loecker et al. (2014) by using the lagged output prices, lagged market shares, lagged exporter status, lagged input tariffs, lagged output tariffs and their interactions with appropriately lagged inputs to form additional moments to jointly estimate $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$. Finally, we follow De Loecker and Warzynski (2012) by using a translog specification of production function, i.e.,

$$\begin{aligned} f_{it}(\mathbf{x}_{it};\boldsymbol{\beta}) &= \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \beta_{ll} l_{it}^2 + \beta_{kk} k_{it}^2 + \beta_{mm} m_{it}^2 \\ &\quad + \beta_{lk} l_{it} k_{it} + \beta_{km} k_{it} m_{it} + \beta_{lm} l_{it} m_{it} + \beta_{lkm} l_{it} k_{it} m_{it}. \end{aligned}$$

Under the translog output production function, the output elasticity of material is calculated as $\hat{\theta}_{it}^m = \hat{\beta}_m + 2\hat{\beta}_{mm} m_{it} + \hat{\beta}_{lm} l_{it} + \hat{\beta}_{km} k_{it} + \hat{\beta}_{lmk} l_{it} k_{it}$.

### B.2. Alternative Production Function Estimations

In this subsection, we discuss three alternative approaches of production function estimation and compare them with our method.

**Alternative I: Quantity-based and translog production function without adjustment for input prices.** This approach still uses output in physical quantity terms, but does not control for omitted input prices in the estimation of production function. In other words, the estimation specification of production function is

$$q_{it} = f_{it}(\tilde{\mathbf{x}}_{it};\boldsymbol{\beta}) + \omega_{it} + \varepsilon_{it}, \tag{15}$$

and

$$\begin{aligned} f_{it}(\tilde{\mathbf{x}}_{it};\boldsymbol{\beta}) &= \beta_l \tilde{l}_{it} + \beta_k \tilde{k}_{it} + \beta_m \tilde{m}_{it} + \beta_{ll} \tilde{l}_{it}^2 + \beta_{kk} \tilde{k}_{it}^2 + \beta_{mm} \tilde{m}_{it}^2 \\ &\quad + \beta_{lk} \tilde{l}_{it} \tilde{k}_{it} + \beta_{km} \tilde{k}_{it} \tilde{m}_{it} + \beta_{lm} \tilde{l}_{it} \tilde{m}_{it} + \beta_{lkm} \tilde{l}_{it} \tilde{k}_{it} \tilde{m}_{it}. \end{aligned}$$

Other procedures are similar to those used in production function estimation in this paper.

**Alternative II: Revenue-based and translog production function without adjustment for input prices.** This approach is similar to the approach used in Alternative I, except that revenue output (instead of quantity output) is used. Specifically, the estimation specification of production function is

$$\tilde{q}_{it} = f_{it}(\tilde{\mathbf{x}}_{it};\boldsymbol{\beta}) + \omega_{it} + \varepsilon_{it}, \tag{16}$$

where $\tilde{q}_{it}$ is (log) deflated output.

**Alternative III: Revenue-based and Cobb-Douglas production function without adjustment for input prices.** This approach is similar to the approach used in Alternative II, except that a Cobb-Douglas (instead of translog) production function is assumed. Specifically, the estimation specification of production function is

$$\tilde{q}_{it} = g_{it}(\tilde{\mathbf{x}}_{it}; \boldsymbol{\beta}) + \omega_{it} + \varepsilon_{it}, \tag{17}$$

and

$$g_{it}(\tilde{\mathbf{x}}_{it}; \boldsymbol{\beta}) = \beta_l \tilde{l}_{it} + \beta_k \tilde{k}_{it} + \beta_m \tilde{m}_{it}.$$

Estimated output elasticities of inputs for these production function estimations are reported in Appendix Table 4, and correlations among markups from these estimations are provided in Appendix Table 6.

Figure 1: Tariffs Evolution During 1996-2007

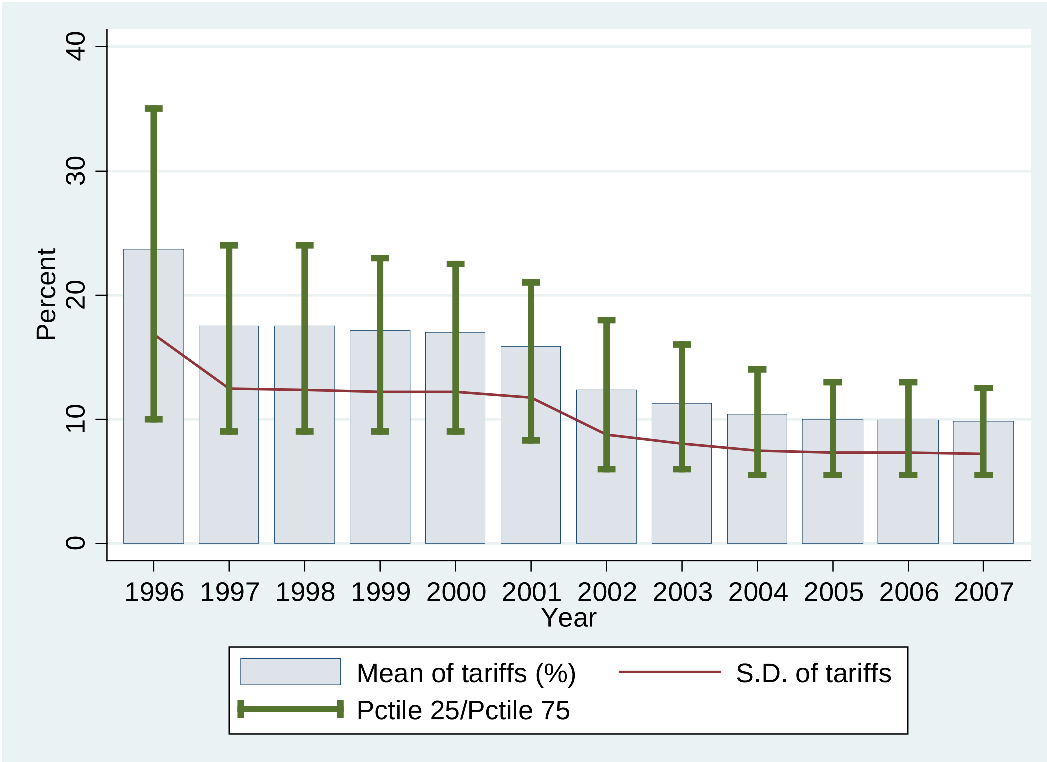

Figure 2: The Correlation Between Tariffs in 2001 and Tariff Changes During 2001-2005 (Three-digit CIC Industries)

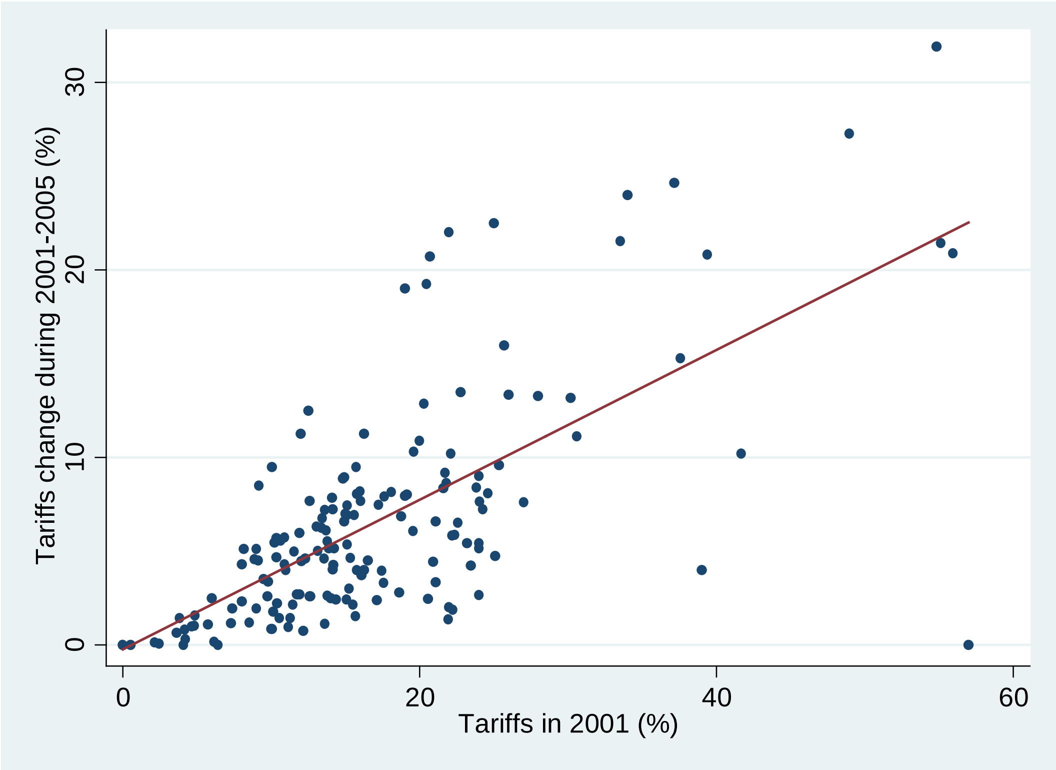

Figure 3: Estimated Markups for Two-digit CIC Manufacturing Industries

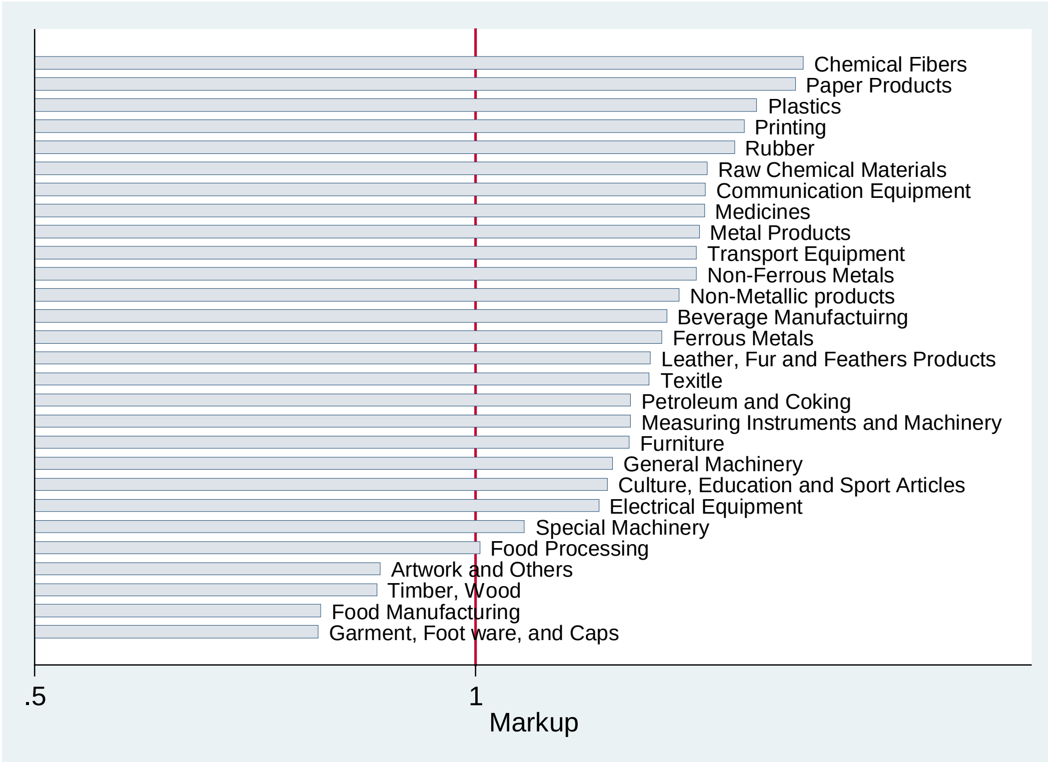

Figure 4: Changes of Markup Dispersion for High vs Low Tariff Industries

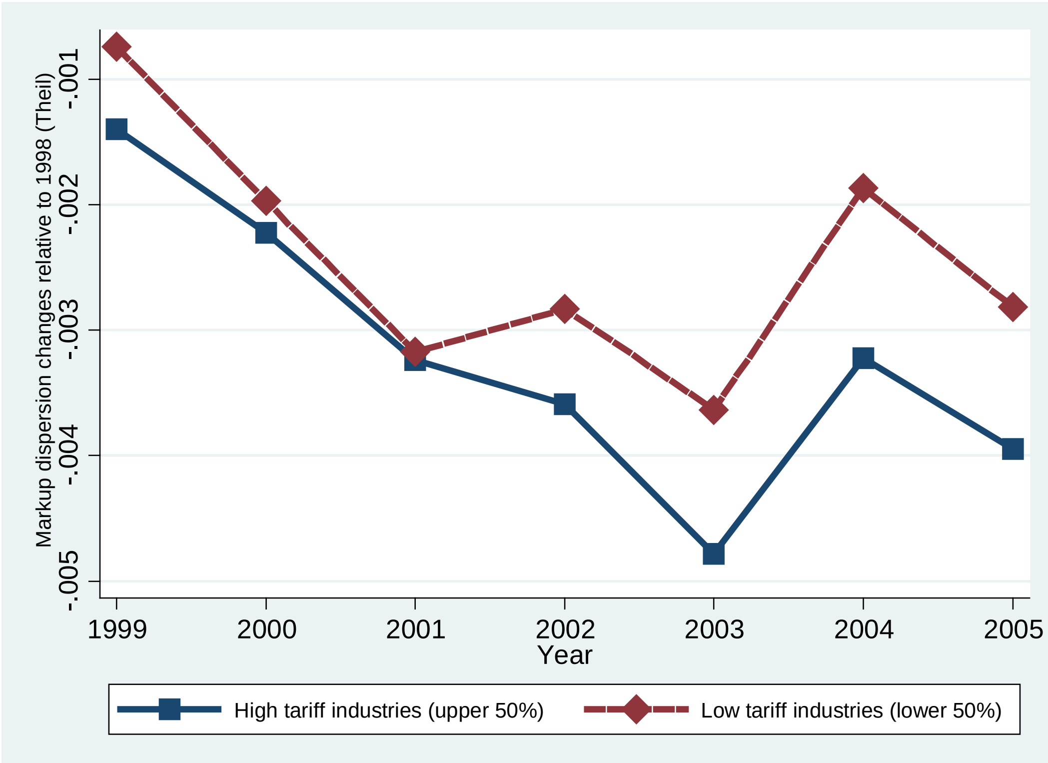

Table 1: Main results

| Dependent variable: Theil dispersion of markups (in log) | (1) | (2) | (3) |
|---|---|---|---|
| $Tariff_{2001}$*$Post_{2002}$ | -0.322*** | -0.307*** | -0.313*** |
| | (0.104) | (0.103) | (0.101) |
| Agglomeration (EG-index) | | -0.635 | -0.900 |
| | | (0.647) | (0.676) |
| Average fixed assets (log) | | 0.019 | 0.017 |
| | | (0.039) | (0.040) |
| Number of firms (log) | | 0.023 | -0.008 |
| | | (0.029) | (0.028) |
| Output share of $SOEs_{2001}$*$Post_{2002}$ | | | -0.148** |
| | | | (0.073) |
| Average wage per $worker_{2001}$*$Post_{2002}$ | | | 0.001 |
| | | | (0.009) |
| Export intensity in $2001_{2001}$*$Post_{2002}$ | | | 0.019 |
| | | | (0.062) |
| Industry fixed effect | X | X | X |
| Year fixed effect | X | X | X |
| Observations | 1,235 | 1,235 | 1,232 |
| R-squared | 0.357 | 0.359 | 0.365 |
| Number of industries | 155 | 155 | 154 |

Note: Standard errors, clustered at 3-digit industry level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2: Checks on the identifying assumptions

| Dependent variable: Theil dispersion of markups (in log) | (1) Next year | (2) Additional controls | (3) Additional controls | (4) Pre-WTO | (5) Processing traders |
|---|---|---|---|---|---|
| $Tariff_{2001}$*$Post_{2002}$ | -0.294** | -0.302*** | -0.300*** | | -0.076 |
| | (0.119) | (0.097) | (0.097) | | (0.460) |
| $Tariff_{2001}$*One Year Before WTO Accession | 0.075 | | | | |
| | (0.128) | | | | |
| FDI (log) | | 0.163** | 0.163** | | |
| | | (0.076) | (0.074) | | |
| SOE share | | -0.637*** | -0.620** | | |
| | | (0.241) | (0.269) | | |
| Input tariff | | | 0.088 | | |
| | | | (0.070) | | |
| Total exports (log) | | | 0.001 | | |
| | | | (0.020) | | |
| Tariff rates | | | | 0.002 | |
| | | | | (0.001) | |
| Industry fixed effect | X | X | X | X | X |
| Year fixed effect | X | X | X | X | X |
| Time-varying industry characteristics | X | X | X | X | X |
| Interactions between Post02 and significant determinants | X | X | X | X | X |
| Observations | 1,232 | 1,232 | 1,232 | 616 | 694 |
| R-squared | 0.365 | 0.375 | 0.377 | 0.343 | 0.101 |
| Number of sic3 | 154 | 154 | 154 | 154 | 97 |

Note: Standard errors, clustered at 3-digit industry level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Other Robustness Checks

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | 3-digit industry | 3-digit industry | 3-digit industry | 4-digit industry | Include # of products |
| Dependent variable: dispersion of markups (in log) | Gini | CV | RMD | Theil | Theil |
| $Tariff_{2001}*Post_{2002}$ | -0.145*** | -0.164*** | -0.152*** | -0.277*** | -0.314*** |
| | (0.052) | (0.053) | (0.056) | (0.080) | (0.107) |
| $Tariff_{2001}*Post_{2002}*Prodnum_{2001}$ | | | | | 0.000 |
| | | | | | (0.001) |
| Controls | X | X | X | X | X |
| Observations | 1,232 | 1,232 | 1,232 | 3,081 | 1,232 |
| R-squared | 0.376 | 0.352 | 0.351 | 0.135 | 0.365 |
| Number of industries | 154 | 154 | 154 | 391 | 154 |

| | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|
| | 2-digit industry | Single industry | Non-exporters | Two periods | Exclude three industries |
| Dependent variable: dispersion of markups (in log) | Theil | Theil | Theil | Theil | Theil |
| $Tariff_{2001}*Post_{2002}$ | -0.232 | -0.268** | -0.255*** | -0.313*** | -0.321*** |
| | (0.198) | (0.128) | (0.106) | (0.101) | (0.103) |
| Controls | X | X | X | X | X |
| Observations | 224 | 1,231 | 1,226 | 308 | 1,040 |
| R-squared | 0.647 | 0.368 | 0.341 | 0.564 | 0.354 |
| Number of industries | 28 | 154 | 154 | 154 | 130 |

Note: Standard errors, clustered at 3-digit industry level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1; Controls include industry fixed effects, year fixed effects, time-varying industry characteristics, and interactions between Post02 and significant determinants.

Table 4: Import effect and markup regressions at different quantiles

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | | P5 | P25 | P50 | P75 | P95 | Mean |
| Dependent variable: | Log imports of HS-6 Products | Markup | Markup | Markup | Markup | Markup | Markup |
| $Tariff_{2001}$*$Post_{2002}$ | 0.021*** | 0.046* | 0.029 | 0.014 | -0.006 | -0.028 | 0.013 |
| | (0.000) | (0.025) | (0.020) | (0.022) | (0.023) | (0.048) | (0.020) |
| Product/Industry fixed effect | X | X | X | X | X | X | X |
| Year fixed effect | X | X | X | X | X | X | X |
| Time-varying industry characteristics | | X | X | X | X | X | X |
| Interactions between Post02 and tariff determinants | | X | X | X | X | X | X |
| Observations | 35,252 | 1,232 | 1,232 | 1,232 | 1,232 | 1,232 | 1,232 |
| R-squared | - | 0.450 | 0.412 | 0.307 | 0.205 | 0.110 | 0.324 |
| Number of products/industries | 5,036 | 154 | 154 | 154 | 154 | 154 | 154 |

Note: Standard errors in parentheses are in column 1, and are clustered at 3-digit industry level in columns 2-7. *** p<0.01, ** p<0.05, * p<0.1;

Table 5: Price and marginal cost effects

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable: Theil dispersion (in log) | Markup | TFP | Price | Marginal cost |
| $Tariff_{2001}$* $Post_{2002}$ | -0.293*** | -1.206*** | -0.949** | -0.961** |
| | (0.104) | (0.435) | (0.398) | (0.405) |
| TFP dispersion (Theil) | 0.011 | | | |
| | (0.013) | | | |
| Industry fixed effect | X | X | X | X |
| Year fixed effect | X | X | X | X |
| Time-varying industry characteristics | X | X | X | X |
| Interactions between Post02 and tariff determinants | X | X | X | X |
| Observations | 1,210 | 1,210 | 818 | 818 |
| R-squared | 0.366 | 0.150 | 0.042 | 0.041 |
| Number of industries | 154 | 154 | 147 | 147 |

Note: Standard errors, clustered at 3-digit industry level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1;

Table 6: Heterogeneous effects

| Dependent variable: Theil dispersion of markups (in log) | (1) Surviving | (2) Entry/ Exit | (3) SOEs | (4) Non-SOEs | (5) Coastal | (6) Inland |
|---|---|---|---|---|---|---|
| $Tariff_{2001}*Post_{2002}$ | -0.157 | -0.335* | -0.253* | -0.203 | -0.262*** | -0.359** |
| | (0.103) | (0.194) | (0.135) | (0.132) | (0.117) | (0.124) |
| Industry fixed effect | X | X | X | X | X | X |
| Year fixed effect | X | X | X | X | X | X |
| Time-varying industry characteristics | X | X | X | X | X | X |
| Interactions between Post02 and tariff determinants | X | X | X | X | X | X |
| Observations | 1,226 | 1,196 | 1153 | 1,218 | 1,226 | 1,202 |
| R-squared | 0.157 | 0.371 | 0.105 | 0.130 | 0.197 | 0.312 |
| Number of products/industries | 154 | 154 | 150 | 153 | 154 | 153 |

Note: Standard errors, clustered at 3-digit industry level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1;