How Did China's WTO Entry Benefit U.S. Consumers?*

Mary Amiti
Federal Reserve Bank of New York

Mi Dai
Beijing Normal Univesity

Robert C. Feenstra
University of California, Davis

John Romalis
University of Sydney

Preliminary

November 28, 2016

Abstract

China's rapid rise in the global economy following its 2001 WTO entry has raised questions about its economic impact on the rest of the world. In this paper, we focus on the U.S. market and potential consumer benefits. We find that the China trade shock reduced the U.S. manufacturing price index by 6.4 percent between 2000 and 2006. In principle, this consumer welfare gain could be driven by two distinct policy changes that occurred with WTO entry. The first, which has received much attention in the literature, is the U.S. granting permanent normal trade relations (PNTR) to China, effectively removing the threat of China facing very high tariffs on its exports to the U.S. A second, new channel we identify is China reducing its own input tariffs. Our results show that China's lower input tariffs increased its imported inputs, boosting Chinese firm's productivity and their export values and varieties. Lower input tariffs also reduced Chinese export prices to the U.S. market. In contrast, PNTR only increased Chinese exports to the U.S. through its effect on new entry, but had no effect on Chinese productivity nor export prices. We find that at least 60% of the China WTO effect on U.S. price indexes was through China lowering its own tariffs on intermediate inputs.

*Amiti: Federal Reserve Bank of New York. 33 Liberty Street, New York, NY 10045 (email: mary.amiti@ny.frb.org); Di: School of Economics and Business Administration (SEBA), Beijing Normal University, Beijing 100875, China (email: daimi002@gmail.com); Feenstra: UC Davis (rcfeenstra@ucdavis.edu); Romalis: University of Sydney (john.romalis@sydney.edu.au. We thank Gordon Hanson, Pablo Fajgelbaum, Dan Trefler and David Weinstein for insightful comments. We are grateful to Tyler Bodine-Smith and Preston Mui for excellent research assistance. The views expressed in this paper are those of the authors and do not necessarily represent those of the Federal Reserve Bank of New York or the Federal Reserve System.

1 Introduction

China's manufacturing export growth in the last 20 years has produced a dramatic realignment of world trade, with China emerging as the world's largest exporter. China's export growth was especially rapid following its World Trade Organization (WTO) entry in 2001, with the 2001–2006 growth rate of 30 percent per annum being more than double the growth rate in the previous five years. This growth has been so spectacular that it has attracted increasing attention to the negative effects of the China "trade shock" on other countries, such as employment and wage losses in import-competing U.S. manufacturing industries. Surprisingly, given the traditional focus of international trade theory, little analysis has been made of the potential gains to consumers in the rest of the world, who could benefit from access to cheaper Chinese imports and more imported varieties. Our focus is on potential benefits to consumers in the U.S., where China accounts for more than 20 percent of imports. In principle, consumer gains could be driven by two distinct policy changes that occurred with China's WTO entry. The first, which has received much attention in the literature, is the U.S. granting permanent normal trade relations (PNTR) to China, effectively removing the threat of China facing very high tariffs on its exports to the U.S. A second, new new channel we identify through which China's WTO entry lowered U.S. price indexes, is China reducing its own input tariffs. In this paper, we quantify how much U.S. consumer welfare improved due to China's WTO entry; and we identify that the key mechanism by which China's WTO entry reduced U.S. price indexes was through China lowering its own tariffs on intermediate inputs.

To measure China's impact on U.S. consumers (by which we mean both households and firms importing from China), we utilize Chinese firm-product-destination level export data for the years 2000 to 2006, during which China's exports to the U.S. increased nearly four-fold. One striking feature is that the extensive margin of China's U.S. exports accounts for 85 percent of this growth, and most of it is due to new firms entering the export market (70 percent of total growth) rather than incumbents exporting new products (15 percent of total growth). To ensure we properly incorporate new varieties in measuring price indexes, we construct an exact CES price index, as in Feenstra (1994), which comprises a "price" and a "variety" component.[1] We find that the China import price index in the U.S. falls by 44 percent over the period 2000 to 2006 due to the growth in exported product varieties. But of course this number needs to be adjusted by China's share in U.S. manufacturing industries to get a measure of U.S. consumer welfare. We supplement the Chinese data with U.S. reported trade data from other countries as well as U.S. domestic sales to construct overall U.S. manufacturing price indexes. With these data, we explicitly take into account that the China shock can affect prices of competitor firms as well as net entry in the U.S. market.

We model Chinese firm behavior by generalizing the Melitz (2003) model to allow firms to import intermediate inputs as in Blaum, Lelarge, and Peters (2015). We expect that the reduction in China's

[1]Broda and Weinstein (2006) built on this methodology to estimate the size of the gains from importing new varieties into the U.S. In contrast to that paper, we define a Chinese variety at the firm-product-destination level (rather than product-country).

tariffs on intermediate inputs has expanded the international sourcing of these inputs, as in Antras, Fort, and Tintelnot (2014), Gopinath and Neiman (2014), and Halpern, Koren, and Szeidl (2015). Expanded sourcing of imported inputs raises Chinese firms' productivity, which makes it possible for them to increase their exports on both the intensive and extensive margins. Lower tariffs on Chinese imported inputs also lowers the marginal costs of Chinese firms producing goods, thus reducing export prices. We also extend the theory to allow the China shock to be driven by a reduction in uncertainty due to PNTR, which we model as a simplified version of Handley and Limão (2013).

Within this theoretical framework, we estimate an equation for Chinese firms' U.S. export participation and export prices, from which we construct fitted values due to WTO entry that we link to U.S. price indexes. We estimate these equations using highly disaggregated Chinese firm-product level international trade data, which we combine with tariff data and firm-level Chinese industrial data. A major challenge in the estimation is measuring marginal costs, which appears in the export price equation. Our proxy for marginal costs is the inverse of a Chinese firm's total factor productivity (TFP), for which we construct a novel instrument that targets the channel through which input tariffs affect TFP directly. More specifically, we estimate an importing equation of Chinese firms' inputs at the firm-product level and use the fitted import values from these estimates to construct theoretically consistent instruments of the intensive and extensive margins of importing. The results from the importing equations show that reductions in Chinese input tariffs lead to higher import values and more imported varieties, with proportionally bigger effects for large firms. In first-stage estimates of the export price equation, we find that lower input tariffs, by increasing imported intermediate inputs, boost firm-level productivity. Our specifications allow for export prices to be influenced by input tariffs and the effect of PNTR, which we estimate by utilizing the "gap" between the US column 2 tariff and the US MFN tariff as in Pierce and Schott (2016) and Handley and Limão (2013).

Our results show that China's WTO entry drove down the U.S. manufactured goods price index by 6.4 percent, an average of 1 percent annually between 2000 and 2006, due to a lower conventional price index and increased variety. Lower tariffs on Chinese firms' imported inputs resulted in lower prices on their U.S. exports, mostly arising from the direct effect of lower input tariffs.[2] In contrast, we find no effect at all from PNTR on China's export prices, as expected from the theory where Chinese firms set prices after the tariff is known. We find that PNTR has no effect on TFP but does have a significant effect on Chinese entry into U.S. exporting, with more entry in the higher gap industries post-WTO entry.

Interestingly, our results show that most of the effect of the China WTO shock on U.S. price indexes is due to China reducing its own input tariffs rather than the PNTR: we find that at least 60 percent of China's WTO effect comes via China's conventional price index, which PNTR has no

[2]The input tariff effect on export prices through TFP was difficult to identify, possibly due to a quality bias, which we address in section 4.3. In a related paper, Fan, Li, and Yeaple (2014) find that Chinese export prices are increasing in productivity and decreasing in tariffs due to quality upgrading. Also consistent with quality upgrading, Manova and Zhang (2012) show that Chinese firms charge higher prices to more distant richer countries. However, Khandelwal, Schott, and Wei (2013) find that the removal of quotas in China's textile industry led to lower export prices.

effect on. Our analysis explicitly takes into account how the China trade shock affects competitor prices and entry. We find that lower Chinese export prices due to China's WTO entry, constructed from the fitted values of our export price equation, reduced both the China price index and the prices of competitor firms in the U.S.; and led to exit of Chinese competitors and other competitors in the U.S. These effects could be due to less efficient firms exiting the U.S. market, lower marginal costs or lower markups. The China-WTO variety instrument, constructed from the fitted values of the export share equation, works almost entirely through the China variety component with hardly any effect on competitor prices and varieties. Both PNTR and lower input tariffs contribute to the reduction in the U.S price index due to the Chinese variety component. However, since most of the effects work through the conventional price index it becomes clear that the overall WTO effect is primarily driven by lower Chinese input tariffs.

Our paper builds on a literature that finds lower input tariffs increase firms' TFP (see, for examples, Amiti and Konings (2007) for Indonesia; Goldberg, Khandelwal, Pavcnik, and Topalova (2010) for India; Yu (2015) and Brandt, Van Biesebroeck, Wang, and Zhang (2012) for China). All of these studies only consider the effect of a country's own tariff reduction on firms in their own countries. In contrast, our focus is on how China's lower input tariffs generated gains to households and firms in another country — these are additional sources of gains from trade.[3]

The impact of China's enormous growth on the rest of the world is an increasingly active area of study. Focusing on the United States, Autor, Dorn, and Hanson (2013) find evidence that China's strong export growth has caused negative employment and wage effects in import-competing industries, and Acemoglu, Autor, Dorn, Hanson, and Price (2014) find that China's export growth reduced overall U.S. job growth.[4] Pierce and Schott (2016) attribute the fall in U.S. manufacturing employment from 2001 to 2007 to the change in U.S. trade policy, whereby China was granted PNTR after its WTO entry. Feng, Li, and Swenson (2015) use firm-level data on Chinese exporters to show that the reduced policy uncertainty had a positive impact on the count of exporters, through simultaneous entry and exit. Handley and Limão (2013) argue that the granting of permanent MFN status to China is a reduction in U.S. policy uncertainty, which leads to greater entry and innovation by those exporters. They measure the positive effects on U.S. consumers, and attribute a 0.8 percent gain in U.S. consumer income due to the reduced policy uncertainty. Our focus is on a different channel — China's lower input tariffs — and we also take account of the PNTR policy for which we find a relatively small role.

A limitation of our study is that we consider only the potential consumer benefits, and do not attempt to evaluate the overall welfare gains to the U.S. from China's WTO entry. That broader question requires a computable model. For example, Hsieh and Ossa (2011) calibrate a multi-country model with aggregate industry data at the two-digit level, and find that China transmits small gains

[3] A number of papers have shown a connection between importing varieties and exporting. See Feng, Li, and Swenson (2012) on China, Bas (2012) on Argentina, and Bas and Strauss-Kahn (2014) on France.

[4] These type of channels have also been studied for other countries (for example, Bloom, Draca, and Van Reenen (2011) on European countries and Iacovone, Rauch, and Winters (2013) on Mexico).

to the rest of the world.[5] More recently, Caliendo, Dvorkin, and Parro (2015) combine a model of heterogeneous firms with a dynamic labor search model. Calibrating this to the United States, they find that China's export growth created a loss of about 1 million jobs, effectively neutralizing any short-run gains, but still increasing U.S. welfare by 6.7 percent in the long-run. Both of these papers rely on the assumption of the Arkolakis, Costinot, and Rodriguez-Clare (2012) (ACR) framework (i.e. a Pareto distribution for firm productivities). Our approach does not rely on a particular distribution of productivities, and we shall argue that the sources of consumer gains from trade that we measure are *additional to* the gains from reducing iceberg trade costs in the ACR framework, and additional to the the gains from reducing uncertainty over U.S. tariffs in Handley and Limão (2013).

The rest of the paper is organized as follows. Section 2 develops the theoretical framework. Section 3 previews key features of the data, including estimates of variety, elasticities of substitution, and total factor productivity (TFP). Section 4 estimates export share and export price equations for Chinese firms. Section 5 estimates the impact of China's WTO accession on U.S. manufacturing price indexes. Section 6 concludes.

2 Theoretical Framework

2.1 Consumers

In order to measure the impact of China's export growth on the U.S. consumer price index, we shall assume a nested CES utility function for the representative consumer. At the upper level, we can write utility from consuming goods $g \in G$ in country j (the United States) and period t as:

$$U_t^j = \left(\sum_{g \in G} \alpha_g^j \left(Q_{gt}^j \right)^{\frac{\kappa-1}{\kappa}} \right)^{\frac{\kappa}{\kappa-1}}, \tag{1}$$

where g denotes an industry that will be defined at an HS 6-digit code or some other broad product grouping, and G denotes the set of HS 6-digit codes; Q_{gt}^j is the aggregate consumption of good g in country j and period t; $\alpha_g^j > 0$ is a taste parameter for the aggregate good g in country j; and κ is the elasticity of substitution across goods.

Consumption of g is comprised of varieties from each country within that HS code:

$$Q_{gt}^j = \left(\sum_{i \in I_{gt}} \left(Q_{gt}^{ij} \right)^{\frac{\sigma_g-1}{\sigma_g}} \right)^{\frac{\sigma_g}{\sigma_g-1}}, \tag{2}$$

where Q_{gt}^{ij} is the aggregate industry quantity in industry g sold by countries $i \in I_{gt}$ to country j in period t, and σ_g denotes the elasticity of substitution across these aggregate country varieties in industry g.

[5] In a multi-country general equilibrium model, di Giovanni, Levchenko, and Zhang (2014) find that the welfare impact of China's integration is larger when its growth is biased toward its comparative disadvantage sectors.

We suppose that there is a number of disaggregate varieties N_{gt}^{ij} sold in industry g by country i to country j in year t. In practice, these varieties of products will be measured for China by firm-level data in country i across all HS 8-digit level products within an HS 6-digit industry. Denoting consumption of these product varieties by $q_g^{ij}(\omega)$, aggregate sales in industry g by country i to country j are:

$$Q_{gt}^{ij} = \left(\sum_{\omega \in \Omega_{gt}^{ij}} \left(z_g^{ij}(\omega) q_{gt}^{ij}(\omega) \right)^{\frac{\rho_g - 1}{\rho_g}} \right)^{\frac{\rho_g}{\rho_g - 1}}, \tag{3}$$

where $z_g^{ij}(\omega) > 0$ is a *taste or quality* parameter for the variety ω of good g sold by country i to country j;[6] Ω_g^{ij} is the set of varieties; and ρ_g denotes the elasticity of substitution across varieties in sector g. We can expect that the elasticity of substitution ρ_g at the firm-product level exceeds the elasticity σ_g across countries in sector g.

Our goal is to compute a price index that accurately reflects consumer utility given this nested CES structure. We begin with the exports of a foreign country i (think of China) to country j (think of the U.S.). The CES price index that is dual to (3) is:

$$P_{gt}^{ij} = \left(\sum_{\omega \in \Omega_{gt}^{ij}} \left(p_{gt}^{ij}(\omega) / z_g^{ij}(\omega) \right)^{1-\rho_g} \right)^{\frac{1}{1-\rho_g}}, \tag{4}$$

from which it follows that share of product variety ω within the exports of country j is,

$$s_{gt}^{ij}(\omega) \equiv \left(\frac{p_{gt}^{ij}(\omega) q_{gt}^{ij}(\omega)}{\sum_{\omega \in \Omega_{gt}^{ij}} p_{gt}^{ij}(\omega) q_{gt}^{ij}(\omega)_{gt}^{ij}} \right) = \left(\frac{p_{gt}^{ij}(\omega) / z_g^{ij}(\omega)}{P_{gt}^{ij}} \right)^{1-\rho_g}. \tag{5}$$

Consider two equilibria with theoretical price indexes P_{gt}^{ij} and P_{g0}^{ij}, which reflect different prices $p_{gt}^{ij}(\omega)$ and $p_{g0}^{ij}(\omega)$ and also differing sets of varieties Ω_{gt}^{ij} and Ω_{g0}^{ij}. We assume that these two sets have a non-empty intersection of varieties, denoted by $\overline{\Omega}_g^{ij} = \Omega_{gt}^{ij} \cap \Omega_{g0}^{ij}$. We refer to the set $\overline{\Omega}_g^{ij}$ as the "common" varieties, available in periods t and 0. Feenstra (1994) shows how the ratio of P_{gt}^{ij} and P_{g0}^{ij} can be measured, as:

$$\frac{P_{gt}^{ij}}{P_{g0}^{ij}} = \left[\prod_{\omega \in \overline{\Omega}_g^{ij}} \left(\frac{p_{gt}^{ij}(\omega)}{p_{g0}^{ij}(\omega)} \right)^{w_{gt}^{ij}(\omega)} \right] \left(\frac{\lambda_{gt}^{ij}}{\lambda_{g0}^{ij}} \right)^{\frac{1}{\rho_g - 1}}, \tag{6}$$

where $w_{gt}^{ij}(\omega)$ are the Sato-Vartia weights at the variety level, defined using the shares $\bar{s}_{gt}^{ij}(\omega)$ within the common set,

$$w_{gt}^{ij}(\omega) \equiv \frac{\left(\bar{s}_{gt}^{ij}(\omega) - \bar{s}_{g0}^{ij}(\omega) \right) / \left(\ln \bar{s}_{gt}^{ij}(\omega) - \ln \bar{s}_{g0}^{ij}(\omega) \right)}{\sum\limits_{\omega \in \overline{\Omega}_g^{ij}} \left(\bar{s}_{gt}^{ij}(\omega) - \bar{s}_{g0}^{ij}(\omega) \right) / \left(\ln \bar{s}_{gt}^{ij}(\omega) - \ln \bar{s}_{g0}^{ij}(\omega) \right)}, \quad \bar{s}_{gt}^{ij}(\omega) \equiv \frac{p_{gt}^{ij}(\omega) q_{gt}^{ij}(\omega)}{\sum\limits_{\omega \in \overline{\Omega}_g^{ij}} p_{gt}^{ij}(\omega) q_{gt}^{ij}(\omega)}, \tag{7}$$

[6] We will allow some of the taste parameters $z_g^{ij}(\omega)$ to vary with time, as explained below.

and

$$\lambda_{gt}^{ij} \equiv \frac{\sum_{\omega \in \overline{\Omega}_g^{ij}} p_{gt}^{ij}(\omega) q_{gt}^{ij}(\omega)}{\sum_{\omega \in \Omega_{gt}^{ij}} p_{gt}^{ij}(\omega) q_{gt}^{ij}(\omega)} = 1 - \frac{\sum_{\omega \in \Omega_{gt}^{ij} \setminus \overline{\Omega}_g^{ij}} p_{gt}^{ij}(\omega) q_{gt}^{ij}(\omega)}{\sum_{\omega \in \Omega_{gt}^{ij}} p_{gt}^{ij}(\omega) q_{gt}^{ij}(\omega)}, \tag{8}$$

and likewise for $\bar{s}_{g0}^{ij}(\omega)$ and λ_{g0}^{ij}, defined as above for $t = 0$.

The first term in equation (6) is constructed in the same way as a conventional Sato-Vartia price index – it is a geometric weighted average of the price changes for the set of varieties $\overline{\Omega}_g^{ij}$, with log-change weights. The second component comes from Feenstra (1994) and takes into account net variety growth: λ_{gt}^{ij} equals one minus the share of expenditure on new products, in the set Ω_{gt}^{ij} but not in $\overline{\Omega}_g^{ij}$, whereas λ_{g0}^{ij} equals one minus the share of expenditure on disappearing products, in the set Ω_{g0}^{ij} but not in $\overline{\Omega}_g^{ij}$.

Note that the quality of the products in the "common" set $\overline{\Omega}_g^{ij}$, as reflected by their taste parameters $z_g^{ij}(\omega)$, is assumed to be constant over time, but products outside this set and appearing within the λ_{gt}^{ij} terms can have changing quality. To achieve this in theory we can chose $\overline{\Omega}_g^{ij}$ as any non-empty *subset* of $\Omega_{gt}^{ij} \cap \Omega_{g0}^{ij}$ for which the products have constant quality, and the price index formulas above continue to hold true (see Feenstra (1994)). In practice, however, it is hard to know which products have constant quality, so we shall simply use $\overline{\Omega}_g^{ij} = \Omega_{gt}^{ij} \cap \Omega_{g0}^{ij}$ and then correct those prices in the common set for changing quality, using the empirical methods of Hallak and Schott (2011) and Khandelwal (2010); see section 4.

While (6) provides us with an exact price index for varieties sold from country i (China) to country j (the U.S.), we also want to incorporate all other countries selling good g. This could be done in principle by using the exact price index for every other country, as we have done for China. But we will not be able to implement that approach because we do not have the firm-level export data for all other countries. Rather, we will use the HS-10 digit U.S. import data with unit-values from each country, which are defined over the firms $\omega \in \Omega_{gt}^{ij}$ within an industry g as:

$$UV_{gt}^{ij} = \left(\frac{\sum_{\omega \in \Omega_{gt}^{ij}} p_{gt}^{ij}(\omega) q_{gt}^{ij}(\omega)}{\sum_{\omega \in \Omega_{gt}^{ij}} q_{gt}^{ij}(\omega)} \right). \tag{9}$$

We can relate these unit values to the CES index in (4) by using the symmetric CES demand in (5):

LEMMA 1. The unit value is related to the CES index by $UV_{gt}^{ij} = P_{gt}^{ij} / H_{gt}^{ij}$, where $H_{gt}^{ij} \equiv \sum_{\omega \in \Omega_{gt}^{ij}} \left(s_{gt}^{ij}(\omega) \right)^{\frac{\rho_g}{\rho_g - 1}}$.

This result is proved in Appendix A. To interpret it, H_{gt}^{ij} is a modified Herfindahl index depending on the elasticity ρ_g, and if $\rho_g = 2$ then it is the usual Herfindahl index. To apply this Lemma to our data, for countries exporting to the U.S. other than China we will use their unit-values at the HS 10-digit level. In a slight abuse of notation, let ω in (6) refer to the HS 10-digit goods within each HS 6-digit industry, and let $p_{gt}^{ij}(\omega)$ in (6) denote the CES price indexes at the HS 10-digit level. Applying the Lemma, we will replace the CES indexes $p^{ij}_{gt}(\omega)$ by $uv^{ij}_{gt}(\omega)H^{ij}_{gt}$, where $uv^{ij}_{gt}(\omega)$ denotes the unit values at the HS 10-digit level. In principle we should be using the Herfindahl indexes of exporters at the HS 10-digit level, too, but in practice due to data limitations we use these indexes at the HS 6-digit level (see Appendix B). For each HS 6-digit industry, we can construct the variety terms λ^{ij}_{gt} for the products exported by those countries and the change in variety using (8). We also construct the Sato-Vartia index for each HS 6-digit industry and country using the unit-values times their Herfindahls, $uv^{ij}_{gt}(\omega)H^{ij}_{gt}$.[7]

We will aggregate over these U.S. import price indexes from all source countries i, including the U.S. itself, using Sato-Vartia price weights defined over countries. Denoting the non-empty intersection of countries selling to j in period t and period 0 by $\bar{I}^j_g = I^j_{gt} \cap I^j_{g0}$, which we call the "common" countries, the Sato-Vartia weights at the country-industry level are

$$W^{ij}_{gt} = \frac{\left(\bar{S}^{ij}_{gt} - \bar{S}^{ij}_{g0}\right) / \left(\ln \bar{S}^{ij}_{gt} - \ln \bar{S}^{ij}_{g0}\right)}{\sum\limits_{k \in \bar{I}^j_{gt}} \left(\bar{S}^{kj}_{gt} - \bar{S}^{kj}_{g0}\right) / \left(\ln \bar{S}^{kj}_{gt} - \ln \bar{S}^{kj}_{g0}\right)}, \text{ with } \bar{S}^{ij}_{gt} \equiv \frac{P^{ij}_{gt}Q^{ij}_{gt}}{\sum\limits_{k \in \bar{I}^j_{gt}} P^{kj}_{gt}Q^{kj}_{gt}}. \tag{10}$$

The share of countries selling in both period t and period 0 is,

$$\Lambda^j_{gt} \equiv \frac{\sum_{i \in \bar{I}^{ij}_g} P^{ij}_{gt}Q^{ij}_{gt}}{\sum_{i \in I^{ij}_{gt}} P^{ij}_{gt}Q^{ij}_{gt}}. \tag{11}$$

Then we can write the change in the overall U.S. price index for industry g as,

$$\frac{P^j_{gt}}{P^j_{g0}} = \left[\prod_{k \in \bar{I}^j_g} \left(\frac{P^{kj}_{gt}}{P^{kj}_{g0}}\right)^{W^{kj}_{gt}}\right] \left(\frac{\Lambda^j_{gt}}{\Lambda^j_{g0}}\right)^{\frac{1}{\sigma_g - 1}}. \tag{12}$$

The term (11) accounts for countries that begin exporting the the U.S. in industry g during the 2000-2006 period, or who drop out due to competition from China, for example. If a country k selling to the U.S. in the base period drops out entirely and no longer sells in period t, then that will lower Λ^j_{g0} and raise the price index in (12). Provided that the loss in variety from exiting firms and exiting countries is not greater than the gain in variety due to entering Chinese firms, then there will still be consumer variety gains due to the expansion of Chinese trade following its WTO entry. The overall price index (12) accounts for all these offsetting effects, and it will be the basis for our calculations of U.S. consumer welfare.

Using all the above equations, we can decompose this industry g price index as,

[7]The only exception to these methods is for the United States itself, where we will use the Producer Price Index (PPI) in each industry to measure the Sato-Vartia index. For the U.S. variety terms we follow Feenstra and Weinstein (2016) and use the share of sales accounted for by the largest four firms, which is a valid measure of λ^{ij}_{gt} if these are the same firms over time in each industry.

$$\ln \frac{P^j_{gt}}{P^j_{g0}} = \underbrace{\ln\left[\prod_{\omega\in\overline{\Omega}^{ij}_g}\left(\frac{p^{ij}_{gt}(\omega)}{p^{ij}_{g0}(\omega)}\right)^{W^{ij}_{gt}w^{ij}_{gt}(\omega)}\right]}_{ChinaP_g} + \underbrace{\ln\left[\prod_{k\in\overline{I}^j_g\setminus i}\ \prod_{\omega\in\overline{\Omega}^{kj}_g}\left(\frac{uv^{kj}_{gt}(\omega)}{uv^{kj}_{g0}(\omega)}\right)^{W^{kj}_{gt}w^{kj}_{gt}(\omega)}\right]}_{OtherP_g}$$

$$+ \underbrace{\ln\left(\frac{\lambda^{ij}_{gt}}{\lambda^{ij}_{g0}}\right)^{\frac{W^{ij}_{gt}}{\rho_g-1}}}_{ChinaV_g} + \underbrace{\ln\left\{\left[\prod_{k\in\overline{I}^j_g\setminus i}\left(\frac{H^{kj}_{gt}}{H^{kj}_{g0}}\right)^{W^{kg}_{gt}}\left(\frac{\lambda^{kj}_{gt}}{\lambda^{kj}_{g0}}\right)^{\frac{W^{kj}_{gt}}{\rho_g-1}}\right]\left(\frac{\Lambda^j_{gt}}{\Lambda^j_{g0}}\right)^{\frac{1}{\sigma_g-1}}\right\}}_{OtherV_g}. \tag{13}$$

The first term on the right is a conventional Sato-Vartia price index for Chinese imports, constructed over common goods in industry g available both years. The second term is the Sato-Vartio index constructed over the unit-values $uv^{kj}_{gt}(\omega)$ in industry g for all other countries, where we allow ω to measure the HS 10-digit products within each HS 6-digit industry, and using the PPI for the U.S. The third term is the gain from increased varieties from China, constructed using Chinese firm-level export data. The fourth term is the combined welfare effect (potentially a loss) of changing variety and Herfindahl index of exporters H^{kj}_{gt} at the HS 6-digit level from country k and from the U.S. itself.

To aggregate over goods, we follow Broda and Weinstein (2006) and again use the Sato-Vartia weights, now defined as:

$$W^j_{gt} = \frac{\left(S^j_{gt} - S^j_{g0}\right) / \left(\ln S^j_{gt} - \ln S^j_{g0}\right)}{\sum_{g\in G}\left(S^j_{gt} - S^j_{g0}\right) / \left(\ln S^j_{gt} - \ln S^j_{g0}\right)},\ \text{with } S^j_{gt} \equiv \frac{P^j_g Q^j_g}{\sum_{g\in G} P^j_g Q^j_g}.$$

Then we can write the change in the overall U.S. price index as

$$\frac{P^j_t}{P^j_0} = \prod_{g\in G}\left(\frac{P^j_{gt}}{P^j_{g0}}\right)^{W^j_{gt}}. \tag{14}$$

This completes our description of the consumer side of the model, but we still need to investigate the behavior of firms. If we find a substantial increase in the product variety of Chinese firms exporting to the U.S., it will be important to determine what amount of this increase is actually due to China's entry to the WTO, whether this increase comes from reduced uncertainty over U.S. tariffs or from the reduction in Chinese tariffs. Introducing heterogeneous firms will allow us to develop structural equations to determine how variety in our model is related to U.S. and Chinese tariff changes. This investigation will also clarify the various sources of gains from trade in our model and, in particular, how these sources are related to the gains from trade in ACR (2012).

2.2 Firms

We focus on Chinese firms (country i) exporting to the United States (country j). Within each industry g, firms randomly draw a productivity φ. The production structure is as in Melitz (2003), but we also incorporate the imports of intermediate inputs by firms engaged in exporting. That generalization is particularly important as China's accession to the WTO reduced its own import tariffs. We let τ^i_{nt} denote one plus the *ad valorem* tariff that China charges on its imports of intermediate input n, with the vector of inputs tariffs denoted by τ^i_t. Choosing the Chinese wage as the numeraire, the marginal cost of a Chinese firm with productivity φ in industry g is written as,

$$c^i_g(\tau^i_t, \varphi), \quad \frac{\partial c^i_g}{\partial \varphi} < 0. \tag{15}$$

Marginal costs depend negatively on the productivity of the firm, φ, in addition to depending on the input tariffs. Productivity will influence both marginal costs and the quality of the firm, and firms of differing productivity might have different sourcing strategies for their inputs, as in Blaum, Lelarge, and Peters (2015).[8] For example, we expect that more productive and therefore larger firms will source from more suppliers, leading to lower costs, as we discuss in section 4.[9]

Given this structure of costs, the rest of model is similar to Melitz (2003) but differing in three respects. First, we also allow for *ad valorem* tariffs on the output of firms, but for the moment we ignore any uncertainty about the U.S. tariff that is applied to China so that the U.S. tariff τ^{ij}_{gt} does not change. Second, we allow for a simple treatment of quality, which we have denoted above by $z^{ij}_g(\varphi)$. Specifically, we shall suppose that each productivity level corresponds to a quality $z^{ij}_g(\varphi)$, and that the marginal costs in (15) are needed to produce one unit of the *quality-adjusted quantity* $z^{ij}_g(\varphi)q^{ij}_{gt}(\varphi)$. Finally, for simplicity we do not consider the entry of firms into each sector, but normalize the mass of potential entrants at unity.

The *quality-adjusted prices* $p^{ij}_{gt}(\varphi)/z^{ij}_g(\varphi)$ of an individual product variety are inclusive of tariffs, and are obtained as a markup over marginal costs:

$$\frac{p^{ij}_{gt}(\varphi)}{z^{ij}_g(\varphi)} = \frac{\rho_g}{(\rho_g - 1)} c^i_g(\tau^i_t, \varphi)\tau^{ij}_{gt}. \tag{16}$$

The revenue of the firm must be divided by τ^{ij}_{gt} to reflect tariff payments, and then is further divided by the elasticity ρ_g to obtain firm profits. These profits are set equal to the fixed labor costs of f^{ij}_g to give us the zero-profit-cutoff (ZPC) condition:

[8] By *sourcing strategy* we mean the list of industries and countries that the firm sources from. This strategy is endogenous and determining it requires the solution of a complex problem for the firm, as illustrated by Antras, Fort, and Tintelnot (2014), Gopinath and Neiman (2014) and Halpern, Koren, and Szeidl (2015). We do not attempt to solve that problem here, but we discuss in section 4.1 how the sourcing strategy is measured empirically. See also Amiti and Davis (2011) and Kasahara and Lapham (2013) for models of firm importing and exporting.

[9] We are influenced by Amiti, Itskhoki, and Konings (2014), who find that large exporters have a greater share of imported intermediate inputs in their costs than do small exporters; in other words, there appears to be a non-homothetic feature to the production structure, as we shall also find for China.

$$\frac{p_{gt}^{ij}(\varphi)q_{gt}^{ij}(\varphi)}{\rho_g} \geq \tau_{gt}^{ij} f_g^{ij}. \tag{17}$$

To solve for the cutoff productivity, we can combine the above two equations with the CES demand equation for product varieties from country i,

$$z_g^{ij}(\varphi)q_{gt}^{ij}(\varphi) = \left(\frac{p_{gt}^{ij}(\varphi)/z_g^{ij}(\varphi)}{P_{gt}^{ij}}\right)^{-\rho_g} \frac{X_{gt}^{ij}}{P_{gt}^{ij}}, \tag{18}$$

where X_{gt}^{ij} is the expenditure on all varieties sold from country i to j in sector g. Multiplying this equation by the quality-adjusted price $p_{gt}^{ij}(\varphi)/z_g^{ij}(\varphi)$ and using (16) and (17), we can solve for firm exports as:

$$p_{gt}^{ij}(\varphi)q_{gt}^{ij}(\varphi) = X_{gt}^{ij}\left(\frac{\rho_g c_g^i(\tau_t^i, \varphi)}{(\rho_g - 1)P_{gt}^{ij}} \frac{\tau_{gt}^{ij}}{z_g^{ij}(\varphi)}\right)^{1-\rho_g}, \tag{19}$$

with the ZCP condition as in (17).

Note that the U.S. tariff on Chinese firms, τ_{gt}^{ij}, enters in two places in the above equations. First, it enters into the price of exports on the right of (16), which also then appears on the left of (17). A reduction in U.S. tariffs *only* as they appear on the right of (16) and on the left of (17) would have the same impact on U.S. welfare as a reduction in iceberg trade costs in ACR (2012). Namely, if the marginal costs defined in (15) are distributed as Pareto, then the gains to the U.S. from a reduction in its tariff on Chinese imports would be inversely proportional to the fall in the U.S. share of expenditure on home varieties. As we explain in the next section, however, the tariff that appears on the right of (16) and on the left of (17) is the MFN tariff; since this tariff changed very little over the period, this source of gains from trade for the U.S. is correspondingly small.

A second place that the U.S. tariff enters is on the right of (17), where it multiplies the fixed costs. The tariff enters there because we have modeled the tariffs as applying to the import revenue, so the revenue on the left of (17) must be divided by τ_{gt}^{ij} to obtain the net-of-tariff revenue remaining for the Chinese firm; we have moved that tariff term to the right. This means that a reduction in tariffs *only* on the right of (17) will have the same impact on the selection of Chinese firms into exporting as a reduction in the *fixed costs* of exporting. This welfare gain for the U.S. does not rely on any distributional assumption for firm productivity and is *distinct from* a reduction in iceberg trade costs in the ACR framework. Indeed, as illustrated in a two-sector, two-country model by Caliendo, Feenstra, Romalis, and Taylor (2015), when tariffs are reduced there is a welfare gain due to the reduction in the home share of varieties (the ACR gain), and in addition, another potential gain due to the entry of firms and expansion in varieties (reflecting the fall in τ_{gt}^{ij} on the right of (17) as well as the change in tariff revenue). As we explain in the next section, the tariff that appears on the right of (17) is actually the "gap" between the U.S. column 2 and MFN tariff, which was eliminated once

China joined the WTO. This is the source of the gains from trade identified in Handley and Limão (2013), and is distinct from the gains in ACR (2012).

Of course, there is a third way that tariffs enter the above equations, and that is through the Chinese tariffs on intermediate inputs, τ_t^i. As we explain in more detail in section 4, a reduction in those tariffs can be expected to a *direct* impact on reducing marginal costs $c_g^i(\tau_t^i, \varphi)$ due to the reduction in input prices, and also an *indirect* impact through potentially improving the firms' productivity. The productivity gains to the firm from expanding suppliers are exactly as we have described in the previous section, i.e. the gains depend on the share of expenditure on new suppliers, just as in Feenstra (1994) and ACR (2012). These gains will translate into lower Chinese prices and also more Chinese exporters due to the term $c(\tau_t^i, \varphi)$ appearing in (16) and on the left of (17). Furthermore, we can expect that the resulting drop in Chinese prices will lead to the exit of some domestic U.S. producers, as well as the exit of firms exporting to the U.S. from other countries. These various effects will lead to gains for the U.S. that are similar in spirit to those in ACR (2012), though they originate from a reduction in Chinese input tariffs rather than from a fall in iceberg transport costs.

2.3 Uncertainty in Tariffs

We now extend the model to incorporate tariff uncertainty, using a simplified version of Handley and Limão (2013).[10] Suppose that the Chinese firm faces two possible values of the U.S. tariff $\tau_{gt}^{ij} \in \left\{\tau_g^{MFN}, \overline{\tau}_g\right\}$, which are at either the MFN level or the alternative column 2 level denoted by $\overline{\tau}_g > \tau_g^{MFN}$. We suppose that some component of the fixed costs of exporting is sunk, which we denote by F_g^{ij}, with the remaining per-period fixed costs of exporting denoted by f_g^{ij}. The firm's decision about its price is made after that tariff is known, while the decision about whether to participate in the export market or not is made before the tariff is known. The pricing decision is shown in (16). The revenue and variable profits for the firm are as before, and deducting the fixed costs of exporting, the one-period value of the firm is,

$$v(\varphi, \tau_{gt}^{ij}) = \frac{p_{gt}^{ij}(\varphi) q_{gt}^{ij}(\varphi)}{\tau_{gt}^{ij} \rho_g} - f_g^{ij} = \frac{X_{gt}^{ij}(\varphi)}{\tau_{gt}^{ij} \rho_g} \left(\frac{c_{gt}^i(\tau_t^i, \varphi)}{C_{gt}^{ij}} \right)^{1-\rho_g} - f_g^{ij}, \tag{20}$$

where we have substituted for the export revenue from (18). Note because all Chinese firms have the same markup and face the same tariffs, then $p_{gt}^{ij}(\varphi)/P_{gt}^{ij} = c_{gt}^{ij}(\tau_t^i, \varphi)/C_{gt}^{ij}$, where C_{gt}^{ij} denotes the CES index as in (4) but over marginal costs $c_{gt}^{ij}(\tau_t^i, \varphi)$.

We suppose for simplicity that if the tariff starts at its MFN level then it remains there in the next period with probability π, and with probability $(1 - \pi)$ the tariff moves to its column 2 level; whereas if the tariff starts at its column 2 level then it stays there forever. This Markov process applies to all industries simultaneously. We need to keep track of what happens to overall Chinese exports under the differing tariffs, so let $\overline{X}_g$ (X_g^{MFN}) denote the value of Chinese exports X_{gt}^{ij} when all tariffs are at

[10] Our simplified treatment here draws on Feng, Li, and Swenson (2015).

their column 2 (MFN) level and let $\overline{C}_g$ (C_g^{MFN}) denote the index of Chinese marginal costs C_{gt}^{ij} when all tariffs are at their column 2 (MFN) level.

With a discount rate $\delta < 1$, the present discounted value of the Chinese firm facing the MFN tariff is

$$V(\varphi, \tau_g^{MFN}) = v(\varphi, \tau_g^{MFN}) + \delta \left[\pi V(\varphi, \tau_g^{MFN}) + (1-\pi) V(\varphi, \overline{\tau}_g) \right].$$

Since $V(\varphi, \overline{\tau}_g) = v(\varphi, \overline{\tau}_g)/(1-\delta)$ by our assumption that the column 2 tariff is an absorbing state, we obtain the *entry condition* for a Chinese firm facing the MFN tariffs,

$$V(\varphi, \tau_g^{MFN}) = \frac{v(\varphi, \tau_g^{MFN})}{(1-\delta\pi)} + \frac{\delta(1-\pi) v(\varphi, \overline{\tau}_g)}{(1-\delta)(1-\delta\pi)} \geq F_g^{ij}. \tag{21}$$

We can simplify this condition by using (20) to obtain,[11]

$$v(\varphi, \overline{\tau}_g) + f_g^{ij} = \left[v(\varphi, \tau_g^{MFN}) + f_g^{ij} \right] \left(\frac{\overline{X}_g / \overline{\tau}_g}{X_g^{MFN} / \tau_g^{MFN}} \right) \left(\frac{\overline{C}_g}{C_g^{MFN}} \right)^{\rho_g - 1}.$$

Substituting this term into (21), we obtain the export participation condition written in terms of one-period profits:

$$v(\varphi, \tau_g^{MFN}) \geq (T_g - 1) f_g^{ij} + T_g (1-\delta) F_g^{ij}, \tag{22}$$

where,

$$T_g \equiv \left\{ \frac{(1-\delta)}{(1-\delta\pi)} + \frac{\delta(1-\pi)}{(1-\delta\pi)} \left(\frac{\overline{X}_g / \overline{\tau}_g}{X_g^{MFN} / \tau_g^{MFN}} \right) \left(\frac{\overline{C}_g}{C_g^{MFN}} \right)^{\rho_g - 1} \right\}^{-1}. \tag{23}$$

These conditions hold in the presence of tariff uncertainty. After China's entry to the WTO, U.S. tariffs are permanently at their MFN level, and the export participation condition for Chinese firms becomes $v(\varphi, \tau_g^{MFN}) \geq (1-\delta) F_g^{ij}$. The right-hand side of that condition differs from (22) by the term $(T_g - 1)[f_g^{ij} + (1-\delta) F_g^{ij}]$, which we interpret as the "effective" tariff term $(T_g - 1)$ multiplied by fixed costs and amortized sunk costs. The effective tariff we have obtained is similar to the results in Handley and Limão (2013) and Feng, Li, and Swenson (2015), except that in (23) we also keep track of what happens to overall Chinese exports. If discounting is small so that $\delta \to 1$, then we see that

$$\ln T_g \to \left(\ln \overline{\tau}_g - \ln \tau_g^{MFN} \right) - \left(\ln \overline{X}_g - \ln X_g^{MFN} \right) - (\rho_g - 1) \left(\ln \overline{C}_g - \ln C_g^{MFN} \right). \tag{24}$$

The first term on the right of (24) is the "gap" between the column 2 and MFN *ad valorem* tariffs, as first used by Pierce and Schott (2016). That variable will influence the export participation decision of Chinese firms. The second and third terms keep track of what happens to the real value of Chinese exports to the U.S. market. We will not attempt to measure these additional terms, and to the extent

[11] We suppose that there is no change in the Chinese input tariffs τ_t^i even if column 2 tariffs are imposed.

that they are correlated with the "gap" between the column 2 and MFN *ad valorem* tariffs, we will be taking that correlation into account by using only the "gap" itself.

3 Data and Preliminary Estimates

The key variables required for our analysis are China's export prices, measures of variety, estimates of elasticities of substitution, and total factor productivity. For these, we utilize a number of different data sources. The first is from China Customs, providing annual trade data on values and quantities at the HS 8-digit level by firm-destination for the period 2000 to 2006. This covers the universe of Chinese exporters. We restrict the sample to manufacturing products, which we identify using a mapping to SITC 1-digit codes in the range 5 to 8. We use these data to construct price indexes as described in section 2.1.

Second, we supplement the China-reported trade data with U.S.-reported data in order to incorporate all other foreign countries and domestic U.S. firms in the construction of the U.S. price index for manufacturing industries. For U.S. imported goods from countries other than China we use customs data at the HS 10-digit-country level from the U.S. Census; for domestic sales by U.S. producers we use the U.S. producer price indexes (PPI) for the common goods component of the price index, and domestic sales shares of the top 4 or 8 U.S. firms, also available from U.S. Census, for the variety component of the price index.

Third, to construct measures of total factor productivity (TFP), we draw on the Annual Survey of Industrial Firms (ASIF) from the National Bureau of Statistics. This is a survey of manufacturers, available for the same period as the customs data. It contains firm-level information on output, materials cost, employment, capital and wages. Each firm's main industry is recorded at the 4-digit Chinese Industrial Classification (CIC). We keep all manufacturing industries, being CIC 2-digit industry codes 13 to 44. To combine the customs and industrial data sets, we relied on information on firm names, addresses, and zip codes because the firm codes are not consistent across the two data sets. For the merged data sets, we end up with industrial data for a third of exporting firms, which account for 50 percent of China's total U.S. exports over this period. We keep the set of firms that exported to the U.S. at any point between 2000 to 2006 and that appear in the industrial data for at least one overlapping year. We will refer to this as our "overlapping sample" and we will make comparisons with the complete data set whenever possible. The data show that the number of U.S. exporters more than tripled over the sample period. See Appendix B for more details on the data construction.

3.1 China's Export Variety

China's exports to the U.S. grew a spectacular 286 percent over the sample period, with growth rates of around 30 percent every year except in 2001 (see Table 1). Most important for our study is how much of this growth comes from new varieties. Denoting the value of Chinese exports to the U.S. by

X_{fgt} for firm f and product g in year t, where the products are now defined at the Chinese HS 8-digit level, and dropping the earlier superscripts ij, we can decompose China's export growth to the U.S. as follows:

$$\frac{\sum_{fg}(X_{fgt}-X_{fg0})}{\sum_{fg}X_{fg0}}=\frac{\sum_{fg\in\overline{\Omega}}X_{fgt}-\sum_{fg\in\bar{\Omega}}X_{fg0}}{\sum_{fg}X_{fg0}}+\frac{\sum_{fg\in\Omega_t\setminus\bar{\Omega}}X_{fht}-\sum_{fg\in\Omega_0\setminus\overline{\Omega}}X_{fg0}}{\sum_{fg}X_{fg0}}, \tag{25}$$

where $\overline{\Omega}=\Omega_t\cap\Omega_0$ is the set of varieties (at the firm-product level) that were exported in t and $t=0$, $\Omega_t\setminus\overline{\Omega}$ is the set of varieties exported in t but not in 0 and $\Omega_0\setminus\overline{\Omega}$ is the set of varieties exported in t_0 but not in t. This equation is an identity that decomposes the total export growth into the intensive margin (the first term on the right) and the extensive margin (the last term), which we report in Table 1. Surprisingly, most of this growth arises from new net variety growth. From the bottom of column 3, we see that the extensive margin accounts for 85 percent of export growth to the U.S. over the whole sample period (columns 2 and 3 sum to 100 percent of the total growth). It is often the case in many other countries that new entrants do not account for a large share of their export growth because new firms typically start off small. But for Chinese exporters, even in the year-to-year changes the extensive margin accounts for around 40 percent of export growth. We can further break down the extensive margin to see if it is driven by incumbent exporters shipping new products or new firms exporting to the U.S. We see from columns 4 and 5 that the extensive margin is almost entirely driven by new exporters — 70 percent of the total export growth over the sample period comes from new firms and the other 15 percent is by incumbent firms exporting new products (columns 4 and 5 sum to the total extensive margin in column 3).

The results in Table 1 clearly show that most of the growth in China's exports to the U.S. is due to new entrants into the U.S. export market. Given that some firms change their identifier over time due to changes of firm type or legal person representatives, we tracked firms over time (using information on the firm name, zip code and telephone number) to ensure that the firm maintains the same identifier over time. This affects 5 percent of the firms and hardly changes the size of the extensive margin (see Table D1 in the Appendix).[12] Even if our algorithm for tracking reclassifications has missed some identifier changes for incumbent firms due to, say, mergers and acquisitions, our approach to measuring the gains from China's entry into the U.S. market using equation (6) is largely unaffected by reclassifications of product codes or firm codes, as the new entry would be offset by the exit.

Those measures also support the finding of a large extensive margin, as we see from column 6, where we report the year-to-year variety adjustment in the China price index and the variety gain over the whole sample period, 2000-2006, i.e. the second term in equation (6). The lambda ratios are raised to a power that includes the elasticity of substitution ρ_g, which we describe in the next subsection, and then weighted as in equation (14). So column 6 reports the effective drop in the U.S.

[12] In the Appendix Table D1, we also show that the very high extensive margin is present when we use alternative ways to define a variety, including HS6-firm, HS4-firm, and HS2-firm, provided we keep the firm dimension. Furthermore, this large extensive margin is present in various subsamples of the data, including nonprocessing trade, consumer goods, nontraders and private firms.

Table 1: Decomposition of China's Export Growth to the U.S.

		Proportion of export growth due to different margins:					
		Variety at the HS8-firm level				Equivalent Price Change	
Year	Total Export Growth %	Intensive Margin	Extensive Margin	Extensive Margin new firms	Extensive Margin incumbents	Due to Chinese Variety	Weighted by China Share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2001	4.2	0.09	0.91	0.75	0.17	-0.025	-0.010
2002	29.8	0.56	0.44	0.21	0.22	-0.036	-0.014
2003	32.2	0.61	0.39	0.23	0.16	-0.076	-0.019
2004	35.1	0.65	0.35	0.23	0.12	-0.019	-0.009
2005	29.4	0.57	0.43	0.22	0.21	-0.066	-0.011
2006	25.6	0.65	0.35	0.20	0.15	0.000	-0.003
2000-2006	286.3	0.15	0.85	0.70	0.15	-0.441	-0.032

Notes: All these margins are calculated using data concorded to HS 8-digit codes at the beginning of the sample. The sum of the intensive margin (column 2) and the extensive margin (column 3) equal 100 percent. The sum of the extensive margin of new firms (column 4) and the extensive margin of incumbent firms (column 5) equals the total extensive margin (column 3). Column 6 converts the variety gain in column 3 to the equivalent change in the price index i.e. the second term on the right of equation (6) and column 7 computes the third term on the right of equation (13), both weighted using the Sato-Vartia weights in equation (14).

import price index from China due to the new varieties, which amounts to -44 percent over 2000-2006. Notice that this total change at the bottom of column 6 is not the same as what is obtained by summing the year-to-year changes in the earlier rows, because the calculation for 2000-2006 is done on the exports that are "common" to those two years. If there is a new variety exported from China in 2001, for example, then its growth in exports up to 2006 is attributed to variety growth; whereas in the earlier rows, only its initial value of exports in 2001 is attributed to variety growth. This method of using a "long difference" to measure variety growth is consistent with the theory outlined in section 2, as it allows for increases in the U.S. taste parameter for that Chinese export in the intervening years, as it penetrates the U.S. market.

To see China's effect on the overall U.S. manufacturing sector index, we need to adjust the values of the variety index in column 6 by China's share in each industry g, which averaged 6 percent in 2006 and 3 percent in 2000. (These are the weights in the entire U.S. market, and not just the import shares.) We do this using the Sato-Vartia weights as in the third term in equation (13), before weighting across industries g as in equation (14). Column 7 shows that the effective price drop due to variety gain from China is reduced to 3.4 percent — this will be the starting point in our consumer welfare calculations, in section 5.[13]

[13]The common goods price index for China — the first component in equation (6) — increased 1.7 percent per year on average.

Table 2: Distribution of Elasticities of Substitution

	China ρ_g	Other countries ρ_g	σ_g
Percentile 5	1.55	1.40	1.51
Percentile 25	2.61	2.21	2.21
Percentile 50	4.21	2.98	3.08
Percentile 75	8.26	5.08	4.31
Percentile 95	31.72	28.76	15.83
Mean	10.98	9.27	5.42
Standard Deviation	32.49	53.70	11.29

Notes: The China ρ_g are estimated using Chinese firm-HS8 level US export data for each HS 6-digit industry g. The "Other countries" $\rho_g$are estimated using U.S. import data at the HS 10-digit country level. And the $\sigma_g$are elasticities of substitution across different countries' HS6 digit goods exported to the U.S.

3.2 Elasticities

We estimate the elasticity of substitution between varieties as in Feenstra (1994), Broda and Weinstein (2006), and Soderbery (2015). For China's exports to the U.S., we estimate the elasticities of substitution across varieties, defined at the firm-HS 8-digit level and within an HS 6-digit industry, ρ_g.[14] This parameter enters in the variety adjustment in the price index — the second term in equation (6) and the third term in equation (13). The median ρ_g, reported in Table 2, is 4.2. We see that there is a big range in the elasticities, with some very large numbers. Variety growth in industries with low elasticities will generate the largest gains whereas variety growth in industries with high elasticities will have a negligible effect on the U.S. price index. For countries other than China, we do not have data at the firm-product level so we estimate elasticities of substitution across varieties defined at the HS 10-digit level within an HS 6-digit industry for each country. The methodology is otherwise the same as for China, except that we constrain the elasticities to be the same for all these other countries within an industry g.[15] We see that the median elasticity for "other countries" is lower at 3.0. This is to be expected because a variety is defined at a more aggregate level. Finally, we estimate σ_g, which is the elasticity of substitution between varieties in industry g produced in different countries. This elasticity appears in the last term of equation (13), and its median estimate is also around 3.

3.3 Total Factor Productivity

We estimate total factor productivity (TFP) using data on all manufacturing firms in the ASIF sample for the period 1998 to 2007. We follow Olley and Pakes (1996), by taking account of the simultaneity between input choices and productivity shocks using firm investment. To estimate the production

[14]We trim the top and bottom 1 percentiles. If there were insufficient observations to estimate an elasticity for an HS 6-digit industry, we used the median in the next level of aggregation.

[15]For the baseline specifications, we estimate the elasticities using U.S. import data for the top 40 countries which account for 98 percent of U.S. manufacturing imports.

Table 3: China's Productivity Growth

Year	Total factor productivity			Real value added per worker		
	All exporters	Overlapping sample		All exporters	Overlapping sample	
	Simple av	Simple av	Weighted av	Simple av	Simple av	Weighted av
2001	-0.01	0.00	0.02	0.01	0.00	-0.03
2002	0.07	0.09	0.11	0.10	0.11	0.12
2003	0.12	0.13	0.11	0.16	0.16	0.10
2004	0.09	0.11	0.13	0.12	0.13	0.15
2005	0.16	0.16	0.11	0.21	0.20	0.19
2006	0.10	0.10	0.10	0.15	0.14	0.15
Average	0.09	0.10	0.10	0.13	0.12	0.11

Notes: Total factor productivity are estimated at the firm level as in Olley and Pakes (1996).

coefficients, we use real value added rather than gross output as the dependent variable because of the large number of processing firms present in China. These processing firms import a large share of their intermediate inputs and have very low domestic value added (Koopman, Wang, and Wei (2012)). Real value added is constructed as deflated production less deflated materials. We use industry level deflators from Brandt, Biesebroeck, and Zhang (2012), where output deflators draw on "reference prices" from China's Statistical Yearbooks and the input deflators are constructed using China's 2002 national input-output table.[16] For the firm's investment measure, we construct a capital series using the perpetual inventory method, with real investment calculated as the time difference in the firm's capital stock deflated by an annual capital stock deflator. The firm's real capital stock is the fixed capital asset at original prices deflated by capital deflators. We begin with the firm's initial real capital stock and construct subsequent periods' real capital stock as $K_{ft} = (1-\delta)K_{f,t-1} + I_{ft}$, where δ is the firm's actual reported depreciation rate. The production coefficients for each 2-digit CIC industry, reported in Table D2 in the Appendix, are used to calculate each firm's log TFP as follows:[17]

$$\ln(TFP_{ft}) = \ln(VA_{ft}) - \gamma_l \ln L_{ft} - \gamma_k \ln K_{ft}$$

The TFP measures are all normalized relative to the firm's main 2-digit CIC industry. From Table 3, we see that average TFP growth of Chinese exporters has been very high. For the average exporter in the full sample it has grown 9 percent per year and only slightly more, at 10 percent per year, in the overlapping sample. For comparison, we also report the average growth in real value added per worker, which shows a similar pattern to TFP growth though at slightly higher average rates of between 11 and 13 percent per annum.

[16] see http://www.econ.kuleuven.be/public/N07057/CHINA/appendix

[17] For the TFP estimation, we clean the ASIF data on the top and bottom one percentile changes in real value added, output, materials, and investment rates; and drop any firm with less than 10 employees.

Table 4: Average Tariffs

	China's tariffs on intermediate inputs					
	HS8 digit		IO category		$ln(Gap)$	
Year	Average	Std Dev	Average	Std Dev	Average	Std Dev
	(1)	(2)	(3)	(4)	(5)	(6)
2000	0.15	0.10	0.13	0.05	0.24	0.15
2001	0.14	0.09	0.12	0.05	0.24	0.15
2002	0.11	0.08	0.09	0.03	0	0
2003	0.10	0.07	0.08	0.03	0	0
2004	0.10	0.07	0.08	0.03	0	0
2005	0.09	0.06	0.07	0.03	0	0
2006	0.09	0.06	0.07	0.02	0	0

Notes: All tariffs are defined as the log of the ad valorem tariff so a 5 percent tariff appears as $ln(1.05)$. The first column presents the simple average of China's import tariffs on HS 8-digit industries. Column 3 represents the simple mean of the cost-weighted average of China's input tariffs within an IO industry code, using weights from China's 2002 input-output table. Column 5 represents the simple average of the gap defined as the difference between the U.S. column 2 tariff and the U.S. MFN tariff in 2000.

3.4 Trade Liberalization

China joined the WTO in December 2001, when it gave a commitment to bind all import tariffs at an average of 9 percent.[18] Although China had already begun the process of reducing tariffs long before then, average tariffs in 2000 were still high at 15 percent, with a large standard deviation of 10 percent.[19] Our main interest is in how China's lower import tariffs on intermediate inputs affected Chinese firms' TFP. Identifying what is an input is not straightforward in the data so we approach this in two ways. First, we follow Amiti and Konings (2007) in the way we construct tariffs on intermediate inputs, using China's 2002 input-output (IO) tables. The most disaggregated IO table available is for 122 sectors, with only 72 of these in manufacturing. We take the raw Chinese import tariff data, which are MFN ad valorem rates at the HS 8-digit level, and calculate the simple average of these at the IO industry level. The input tariff for industry g is the weighted average of these IO industry tariffs, using the cost shares in China's IO table as weights.[20] Average tariffs for each year are reported in Table 4.

Our second approach utilizes the more disaggregated HS 8-digit raw tariff data directly where possible and targets the main channel through which we expect input tariffs to affect TFP. We do this by estimating Chinese exporters fitted imported inputs that are due to lower import tariffs at the HS8 level. In this case we identify an input as any HS 8-digit import that the firm does not also export. These fitted import values will form the basis for our main instruments for firm-level TFP.

Upon China's WTO entry, China benefited from another form of trade liberalization with the

[18]See wto.org for more details.

[19]SeeBai, Krishna, and Ma (2016) for a discussion of other reforms in China.

[20]We thank Rudai Yang from Peking university for the mapping from IO to HS codes, which he constructed manually based on industry descriptions. We include both manufacturing and nonmanufacturing inputs and drop "waste and scrapping".

U.S. Congress granting Permanent Normal Trade Relations (PNTR). It is important to realize that the PNTR did not actually change the tariffs that China faced on its exports to the U.S. The U.S. had applied the low MFN tariffs on its Chinese imports since 1980, but they were subject to annual renewal, with the risk of tariffs reverting to the much higher non-NTR tariff rates assigned to some non-market economies. These non-NTR tariffs are set at the 1930 Smoot-Hawley Tariff Act levels and can be found in "column 2" of the U.S. tariff schedule. Studies by Pierce and Schott (2016) and Handley and Limão (2013) argue that the removal of the uncertainty surrounding these tariff rates helped boost China's exports to the U.S. economy. Following this literature, we refer to this measure as the "gap" and define it as the difference between the column 2 tariff and the U.S. MFN tariff rate in 2000. We see from the last two columns in Table 4 that the gap was very high with a large standard deviation. We will exploit this cross-industry variation to analyze its effect on China's U.S. exports by interacting the gap with a WTO dummy that equals one post-2001.

4 Estimation

Taking the log of the pricing equation (16) we obtain,

$$\ln[p^{ij}_{gt}(\varphi)/z^{ij}_{g}(\varphi)] = \ln[c^{i}_{g}(\tau^{i}_{t}, \varphi)\tau^{ij}_{t}]. \tag{26}$$

The tariff variable τ^{ij}_{t} is the U.S. MFN tariff on China's exports to the U.S., which differs hardly at all over our sample period and is absorbed into firm and year fixed effects, β_f and β_t. In practice when we estimate (26) we shall rewrite it as:

$$\ln p_{fgt} = \beta_f + \beta_t + \beta_1 \ln TFP_{ft} + \beta_2 \ln(Input\tau_{gt}) + \beta_3 IMR(X_{fgt}) + \beta_4 IMR(M_{ft}) + \epsilon_{1fgt}. \tag{27}$$

In this estimating equation, f denotes Chinese firms (country i) selling to the United States (country j) in industry g, where we now drop the country superscripts. We absorb unobserved quality $\Delta \ln z^{ij}_{g}(\varphi)$ from the left of (26) into the error term in (27), $\epsilon_{1fgt} \equiv \ln z_{fgt}$, where we now recognize that firms' qualities might be changing over time.[21] This quality change will likely be correlated with the change in firms' prices, as we shall need to account for in the estimation. We use the total factor productivity of the firms as an inverse measure of marginal costs, which has a coefficient of $\beta_1 = -1$.[22] The next variable on the right is an industry-level measure of the tariffs on imported inputs faced by the firms in industry g, denoted by $Input\tau_{gt}$,which also influences marginal costs.[23]

[21]The error will also incorporate terms reflecting the fact that the pass-through coefficientβ_1 differs across firms, as discussed in the next footnote. As analyzed by Murtazashvili and Wooldridge (2008), pooling across firms to obtain a single coefficient means that additional terms are introduced into the error.

[22]Amiti, Itskhoki, and Konings (2015) show that in the nested CES framework we are using, if the market shares of firms are not infinitesimally small then a pass-through of less than unity (in absolute value) that differs across firms is obtained. We allow for any magnitude of β_1 in our empirical analysis, but we find that $\beta_1 = -1$ is not rejected.

[23]We use the input-output table for China to construct the average input tariff that each firm *producing* in industry g faces.

The final two variables appearing in (27) are the inverse Mills ratio from the export and import participation equations. There is a potential bias due to selection into exporting since the dependent variables are undefined for zero export values. To address this, we include the inverse mills ratio $IMR(X_{fgt})$, which we construct from estimating the following export participation equation:

$$prob(X_{fgt} > 0) = \alpha_g + \alpha_t + \alpha_1 \ln(Input\tau_{gt}) + \alpha_2 Gap_g + \alpha_3 Gap_g \times WTO_t \\ + \alpha_4 Age_{ft} + \alpha_5 Foreign_f + \epsilon_{2fgt}. \tag{28}$$

The dependent variable is binary, equal to 1 if the firm had positive export value in industry g, and zero otherwise. We view this export participation equation as the empirical counterpart to the ZCP condition (17). The explanatory variables include industry fixed effects α_g, time fixed effects α_t, and the input tariffs . We also include the "gap" between the column 2 U.S. tariff and the MFN tariffs and also its interaction with a WTO dummy that takes the value of one after 2001. In addition to these tariff variables, we include the age of the firm — so long as the firm's age is reported for at least one year we can include it for all of the sample years. This variable is used to represent the fixed costs of exporting which appears in (17), using the argument that more experienced exporters will have lower fixed costs. We also include an indicator variable for whether the firm is foreign owned, since we expect that such firms will have lower fixed costs and a higher probability of exporting.

4.1 TFP Instrument

In the pricing equation (27) the marginal costs of the firm appear on the right. We shall measure marginal costs inversely by the TFP of each firm, which is influenced by its sourcing strategy. To develop instruments for TFP, we start with the same production function used by Blaum, Lelarge, and Peters (2015), which is similar to the nested CES utility functions shown in (2) and (3):

$$Y_{ft} = \varphi_{ft} L_{ft}^{\gamma} \left(\alpha_D \left(Q_{ft}^D \right)^{\frac{\sigma-1}{\sigma}} + \alpha_M \left(Q_{ft}^M \right)^{\frac{\sigma-1}{\sigma}} \right)^{(1-\gamma)\frac{\sigma}{\sigma-1}}, \tag{29}$$

where Y_{ft} is the output of firm f in year t with productivity φ_{ft}, using labor L_{ft}, the domestic intermediate input Q_{ft}^D, and the imported intermediate input Q_{ft}^M. We assume that the imported input is a CES aggregate of the various inputs $n \in \Sigma_{ft}$ that are purchased across different HS-8 level categories:

$$Q_{ft}^M = \left(\sum_{n \in \Sigma_{ft}} \alpha_n (q_{fnt})^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}}. \tag{30}$$

Blaum, Lelarge, and Peters (2015) refer to $n \in \Sigma_{ft}$ as the *sourcing strategy* of the firm, and the goal of their analysis it to show how the productivity of the firm is enhanced as the range of imported inputs expands. To achieve this, let C_{ft} denote the unit-cost function for the production function in (29). Following Blaum, Lelarge, and Peters (2015), we suppose the price of the domestic input does not change and we also set the wage at unity as the numeraire. We can choose that input as the "common" intermediate input that is sold in both periods 0 and t, and because its price does not change, an immediate application of Theorem 2 in Feenstra (1994) is that:[24]

$$\frac{C_{ft}}{C_{f0}} = \left(\frac{S_{Dft}}{S_{Df0}}\right)^{\frac{1-\gamma}{\sigma-1}}, \tag{31}$$

where S_{Dft} is the share of total expenditure on intermediate inputs that is devoted to *domestic* inputs in period t. Blaum, Lelarge, and Peters (2015) proceed by measuring the change in unit-costs using this domestic share variable, which endogenously reflects the sourcing strategy of the firm.

There is another way to write the change in unit costs, however, that focuses more directly on the sourcing strategy. Let $c_{ft} \equiv c_f(\tau_t, \Sigma_f(\tau_t))$ denote the unit-cost function corresponding to the CES production function over inputs in (30), where now we make it explicit that the set of inputs $n \in \Sigma_f(\tau_t)$ purchased in period t depends on the vector of input tariffs τ_t.[25] Let $\overline{\Sigma}_f \subseteq \Sigma_f(\tau_0) \cap \Sigma_f(\tau_t)$ denote a non-empty subset of the "common" imported inputs purchased in periods 0 and t. Then analogous to the consumer CES indexes discussed in section 2, the index of firm costs between period t and period 0 is

$$\frac{c_{ft}}{c_{f0}} = \left[\prod_{n \in \overline{\Sigma}_f} \left(\frac{\tau_{nt}}{\tau_{n0}}\right)^{w_{nt}}\right] \left(\frac{\lambda_{ft}}{\lambda_{f0}}\right)^{\frac{1}{\rho-1}}, \tag{32}$$

where w_{nt} is the Sato-Vartia weight for input n, and λ_{ft} denotes the expenditure on imported inputs in the *common* set $\overline{\Sigma}_f$ relative to *total* expenditure on imported inputs in period t.[26] The first term on the right of (32) is the direct effect of tariffs on costs, or the Sato-Vartia index of input tariffs.[27] The second term is the efficiency gain from expanding the range of inputs, resulting in $\lambda_{ft} < \lambda_{f0} \leq 1$.

We can easily relate the efficiency gain in (32) back to the unit-costs of the production function C_{ft}. The ratio of unit-costs can be written as a Sato-Vartia index over the ratio of wages in period t relative to period 0, the ratio of the price of the domestic intermediate input, and the ratio of the price of imported intermediate inputs. Since we are treating the wage and domestic input price as unchanged over time, we simply obtain:

$$\frac{C_{ft}}{C_{f0}} = \left(\frac{c_{ft}}{c_{f0}}\right)^{W_{Mft}(1-\gamma)}, \tag{33}$$

where W_{Mft} is the Sato-Vartia weight of imported inputs within total expenditure on intermediate

[24]Theorem 2 in Feenstra (1994) states that the ratio of unit-costs for a CES function is the Sato-Vartia price index over "common" goods available in both periods (in our case we have only one common domestic input and its price does not change), times the ratio of expenditure on the common goods in the two periods, raised to the inverse of the elasticity of substitution minus one as shown by the final term in in (31).

[25]We are holding constant the net-of-tariff prices of imported inputs, so these prices are suppressed in the notation.

[26]Goldberg, Khandelwal, Pavcnik, and Topalova (2010) adopt a similar approach to estimate the effect of trade liberalization of intermediate inputs on the number of domestic products produced in India.

[27]In principle, this first term corresponds to the index of input tariffs that we denoted by $Input\tau_{gt}$. Because Chinese firms often produce multiple products and we cannot disentangle which of the firm's imports are used to produce each of its export goods, we cannot accurately measure the input weight w_{nt} for each input and output. So in practice, we construct the index of input tariffs as the industry level g, using the weights from an input-output table.

inputs.[28] We see that a reduction in the unit-cost of the import bundle in (32) corresponds directly to a reduction in overall unit costs in (33), by an amount that depends on the Sato-Vartia share of imported inputs. We expect that larger firms would have a greater share of expenditure inputs, for example, and would therefore experience a greater efficiency gain from an expanded sourcing strategy. In our specifications below, we use the firms' employment to reflect their size.

We can combine (31), (32) and (33) and readily solve for the change in the domestic share of intermediate inputs:

$$\frac{S_{Dft}}{S_{Df0}} = \left[\prod_{n\in\bar{\Sigma}_f}\left(\frac{\tau_{nt}}{\tau_{n0}}\right)^{w_{nt}}\right]^{W_{Mft}(\sigma-1)}\left(\frac{\lambda_{ft}}{\lambda_{f0}}\right)^{\frac{W_{Mft}(\sigma-1)}{\rho-1}}. \tag{34}$$

This equation emphasizes that the change in the domestic share is endogenous to the sourcing strategy of the firm and to the tariffs that it faces. Reflecting this, our approach will be to construct instruments that capture the right-hand side variables in (34), focusing on constructing instrument for λ_{ft}, while also including the change in tariffs in our structural equations.

We exploit the highly disaggregated raw tariffs and directly target the channel through which these operate. We regress Chinese firms' imports on China's import tariffs at the HS 8-digit level — the most disaggregated level available — and use those fitted values to construct our instrument for λ_{ft}. More specifically, we estimate the following equation with OLS:

$$\ln M_{fnt} = \delta_f + \delta_t + \delta_1 \ln\tau^i_{nt} + \delta_2 Process_f \times \ln\tau^i_{nt} + \delta_3 \ln L_f \times \ln\tau^i_{nt} + \delta_4 IMR(M_{fnt}) + \epsilon_{3fnt}. \tag{35}$$

The dependent variable is the value of firm f imports in HS 8-digit category n, and the Chinese tariffs τ_{nt} are those that apply to those imports. In order to ensure that these imports are actually intermediate inputs and not final goods, we exclude any import of an HS 8-digit good that the firm exports to any country in that year. We also interact the tariff with a processing dummy to indicate whether the input is imported under processing trade, and with the average firm size, measured by employment. Processing imports already enjoyed duty-free access so a lower tariff on those imports would not reduce the cost of importing and thus should not have a direct impact on the quantity imported.[29] We expect that firms with some portion of ordinary trade to benefit from trade liberalization. We also include year fixed effects and firm fixed effects. Given that the dependent in equation (35) is undefined for zero imports, we address the potential selection bias by including an inverse mills ratio $IMR(M_{fnt})$. The variable is calculated as the ratio of the probability density function to the cumulative distribution function, from the following import participation equation:

[28]Denoting the import share of intermediate input purchases by S_{Mft}, then $W_{Mft} = \left[(S_{Mft}-S_{Mf0})/(\ln S_{Mft}-\ln S_{Mf0})\right]/\left[(S_{Dft}-S_{Df0})/(\ln S_{Dft}-\ln S_{Df0})+(S_{Mft}-S_{Mf0})/(\ln S_{Mft}-\ln S_{Mf0})\right]$.

[29]We classify a firm as processing if it imports more than 99% of its total intermediate inputs under processing trade over the sample period.

$$prob(M_{fnt} > 0 | X_{ft} = 1) = \theta_g + \theta_t + \theta_1 \ln \tau^i_{nt} + \theta_2 Process_f \times \ln \tau^i_{nt} + \theta_3 Process_f \qquad (36)$$
$$+ \theta_4 \ln L_{ft} \times \ln \tau^i_{nt} + \theta_5 \ln L_{ft} + \theta_6 Age_{ft} + \theta_7 Foreign_f + \epsilon_{4ft}.$$

The dependent variable equals one if the firm imports an intermediate input in HS 6-digit category n and zero otherwise.[30] In addition to the variables included in equation (35), we also include the firm's age and a foreign ownership indicator to proxy for the firm's ability to cover the fixed costs of importing, as in the export participation equation — these will provide our exclusion restrictions.

Table 5: Import Participation

Dependent variable Y_{fnt} =1 if Imports>0	probit				OLS
	(1)	(2)	(3)	(4)	(5)
$ln(\tau_{nt})$	-0.248	-0.037	0.549***	0.050*	0.324***
	(0.244)	(0.246)	(0.063)	(0.029)	(0.101)
$ln(\tau_{nt}) \times Process_f$	1.387***	1.161***	1.088***	1.085***	0.396***
	(0.177)	(0.132)	(0.034)	(0.034)	(0.046)
$ln(\tau_{nt}) \times ln(L_f)$	-0.105***	-0.076**	-0.092***		-0.037***
	(0.030)	(0.030)	(0.009)		(0.011)
$ln(L_f)$	0.071***	0.077***	0.083***		0.030***
	(0.004)	(0.004)	(0.001)		(0.001)
$ln(\tau_{nt}) \times Large_f$				-0.300***	
				(0.023)	
$Large_f$				0.198***	
				(0.003)	
Age_{ft}	-0.007***	-0.007***	-0.007***	-0.006***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Foreign_f$	-0.013	0.017***	0.033***	0.030***	0.012***
	(0.008)	(0.006)	(0.001)	(0.001)	(0.002)
$Process_f$	0.084***	0.007	-0.002	-0.006	-0.000
	(0.024)	(0.018)	(0.004)	(0.004)	(0.006)
industry effects	no	HS4 fe	HS6 re	HS6 re	HS6 fe
N	5,685,934	5,685,928	5,685,934	5,685,934	5,685,934

Notes: Observations are at the HS6-firm-year level. The sample includes all firms that import an input at any time during the sample and appear in the ASIF sample at least once and export at least once. Standard errors clustered at HS6 level. Process dummy equals one if more than 99% of the firm's imports were processing over the sample. The dependent variable equals one in 36.1% of the observations.

We present the results from estimating the import participation equation in Table 5 and the results from estimating the import value equation in Table 6.[31] In the import participation equation, we

[30] The dimensionality became too large to estimate at the more disaggregated HS 8-digit level.

[31] We experimented with including the gap variable and its interaction with the WTO dummy in equation (36); however it had an insignificant coefficient with the wrong sign.

also introduce industry effects to take account that some industries may be high-growth industries due to factors such as technological progress, however, in a probit we need to be mindful of the incidental parameters problem induced by too many fixed effects. In Table 5, the first four columns are estimated using probit, with no industry fixed effects in column 1, HS 4-digit industry fixed effects in column 2, and HS 6-digit industry random effects in columns 3 and 4. In the first three columns we interact tariffs with the firm's log mean number of workers and in column 4 we use a dummy indicator equal to one if the firm has more than 1,000 employees. We find that lower tariffs reduce the probability of importing for processing firms[32] and they increase that probability for large nonprocessing firms. The magnitude of the effect of lower tariffs for large nonprocessing firms is equal to $\theta_1 + \theta_4 \times ln(L_f)$, which is negative above the threshold of 790 employees — the median number in the sample is 590. In column 4, we include a "Large" indicator and we see that the coefficient for firms with more than 1,000 workers is equal to -0.25 (summing $\theta_1$and θ_4), which suggests that large exporting firms are more likely to import their inputs. We would expect larger firms to be more likely to be able to cover the fixed costs of importing as they probably have better access to capital markets to finance fixed costs and working capital. As expected, we find foreign firms are more likely to import their inputs; however older exporters are less likely to import their inputs. This might appear surprising but we need to bear in mind that this result is conditional on exporting. We use this equation to construct the inverse mills ratio for importing, $IMR(M_{fnt})$, in the import value equation.

All of the specifications from estimating the import value equation (35) include firm fixed effects and year fixed effects. In column 1 of Table 6, we include China's import tariffs and the tariffs interacted with a processing dummy. We find that the the coefficient on tariffs is negative and significant, $\delta_1 = -5.72$, showing that trade liberalization results in higher imports for firms that import under ordinary trade. In contrast, the coefficient on the interacted processing term is positive, $\delta_2 = 4.88$. The sum of δ_1 and δ_2 is not significantly different from zero, suggesting that processing imports are not greatly affected by lower tariffs. In column 2, we include the interacted employment variable and in column 3 we use the large-firm dummy instead. Both specifications show there is a non-homothetic effect of sourcing imported inputs, with the effects significantly bigger for large firms. In column 4, where we include $IMR(M_{fnt})$, we show that the results are robust to selection bias — the coefficients in column 4 are close to those in column 2.

We use the results from estimating (35) to construct two instruments for λ_{ft} and therefore for TFP. The first is the firm's fitted *total* imports at time t — we take the exponential of the fitted import values $\ln \hat{M}_{tot,ft}$, summed across all of the firm's imports n in each year to get the firm's total and then take the log. That instrument corresponds to the denominator of λ_{ft}. The numerator is the expenditure on inputs that are common in period t and period 0, or any non-empty subset of these common inputs, such as the imports in the median HS category.[33] In practice, we found that many

[32] This is consistent with Kee and Tang (2016).

[33] The use of a non-empty subset of common goods rather than the entire set was shown to be valid by Feenstra (1994) when constructing the λ terms, and it also applies to the variety results in section 2, though we did not use it there.

Table 6: Import Value

Dep. Var: $ln(M_{fnt})$	(1)	(2)	(3)	(4)
$ln(\tau_{nt})$	-5.715***	-2.310*	-5.537***	-2.590*
	(0.790)	(1.204)	(0.796)	(1.480)
$ln(\tau_{nt}) \times Process_f$	4.879***	4.626***	4.808***	4.102***
	(0.878)	(0.887)	(0.881)	(1.365)
$ln(\tau_{nt}) \times ln(L_f)$		-0.512***		-0.466**
		(0.145)		(0.199)
$ln(\tau_{nt}) \times Large_f$			-0.568*	
			(0.297)	
$IMR(M_{ft})$				-0.700
				(1.799)
N	1,908,374	1,908,374	1,908,374	1,908,374
R^2	0.142	0.142	0.142	0.142

Notes: Observations are at the firm-HS8-year level. All columns include firm fixed effects and year fixed effects. Standard errors are clustered at the HS8 level. Process dummy equals one if more than 99% of imports are processing. Large dummy equals one if average firm employment is greater than 1000.

firms did not have common imported inputs over the entire sample period, so the median HS category over the entire sample could not be constructed. We have instead used imports in the median HS 8-digit category on a year-by-year basis, and denote the fitted value of those median imports by $ln(\hat{M}_{med,ft})$. Then the difference between these two instruments, $\ln \hat{M}_{tot,ft} - \ln \hat{M}_{med,ft}$, is meant to capture the expansion of imported inputs on the extensive margin. One limitation of these instruments is that they are only defined for firms that import their inputs, which means that we can only estimate the export share equation for the set of exporters that also import their inputs. This could potentially induce an additional selection bias if exporters that are nonimporters behave differently from exporters that also import. We address this by including the inverse mills ratios constructed from the import selection equation (36).

4.2 Selection into Exporting

For the baseline estimation of the export particpation equation (36), we include all firm-HS6 observations for the period 2000 to 2006 for the set of firms that have at least one nonzero U.S. export observation and that can be mapped to the industrial ASIF data in at least one year. The dependent variable is equal to 1 if the firm had positive export value in industry g, defined at the HS 6-digit level, and zero for all fgt observations where the firm did not export in those HS 6-digit categories. We present the results in Table D3, with columns 1 through 4 presenting estimates from probit models with different sets of industry effects. All of the equations have year fixed effects to take account of macro factors that affect overall entry and exit.

In column 1, with no industry effects, we find that all of the coefficients have the expected signs. The coefficient on China's input tariffs on imported inputs — our main variable of interest — suggests that lower Chinese import tariffs on intermediate inputs increase the probability of exporting. We also find that the probability of entry into the U.S. export market is high in industries where the gap between the U.S. column 2 tariffs and MFN tariffs is high in the post-WTO period, consistent with the literature (see Pierce and Schott (2016)). Once China entered the WTO, the threat of raising U.S. import tariffs to the high column 2 tariffs was removed, increasing the expected profitability of exporting in those industries. The positive coefficients on the foreign firm indicator and the firm's age are consistent with the idea that older firms and foreign firms are in a better position to cover the fixed costs of entering export markets.

Table 7: Export Participation

Dependent variable Y_{fgt} =1 if exports>0	probit				OLS
	(1)	(2)	(3)	(4)	(5)
$ln(Input\tau_{gt})$	-2.325***	-7.242***	-7.242***	-6.651***	-1.934***
	(0.475)	(0.868)	(2.299)	(0.127)	(0.195)
$ln(Gap_g)$	-0.381**	-0.208	-0.208		
	(0.190)	(0.129)	(0.239)		
$ln(Gap_g) \times WTO_t$	0.388***	0.298**	0.298	0.329***	0.052**
	(0.127)	(0.119)	(0.244)	(0.022)	(0.026)
$Foreign_f$	0.028***	0.041***	0.041***	0.047***	0.015***
	(0.006)	(0.005)	(0.009)	(0.003)	(0.002)
Age_{ft}	0.002***	0.003***	0.003***	0.003***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
industry effects	no	HS4 fe	HS4 fe	HS6 re	HS6 fe
clustering	HS6	HS6	IO	no	HS6
# obs.	1,370,885	1,370,885	1,370,885	1,370,885	1,370,885

Notes: Standard errors are clustered at the HS6 industry level in all columns except column 3 where they are clustered at the more aggregate IO industry level. The dependent variable equals one in 31% of the observations.

In the next three columns, we also introduce industry effects to take account that some industries may be high-growth industries due to factors such as technological progress or increased world demand: columns 2 and 3 include industry fixed effects at the more aggregated HS 4-digit level; and column 4 includes random industry effects at the HS 6-digit level. We see that the signs on all of the coefficients remain the same as in column 1, which has no industry effects, but some of the point estimates are different. In particular, the magnitude of the negative coefficient on the input tariff becomes much larger with industry effects, and of similar size in the fixed effects and random effects estimations. For comparison, we present OLS estimates with HS 6-digit level fixed effects in column 5, and again find the same signs on all of the coefficients.[34]

4.3 Pricing Equation

We now have all the variables we need to estimate the export price equation, (27). The dependent variable in the pricing equation is the log unit value of each Chinese exporting firm f in each industry g, inclusive of freight, insurance, and duties.[35] The variables on the right of (27) are total factor productivity of the firm and the input tariff, as well as inverse mills ratios to take account of selection bias into exporting and importing. The results from estimating equation (27) are shown in Table (8).

We present the results from estimating the export price equation 27 in Table 8, where we instrument for TFP and control for potential selection bias in importing and exporting. To address the endogeneity of $ln(TFP_{fgt})$ we instrument using the fitted import values from equation (35) for the *total* and *median* imported input expenditure; and the log gap interacted with the WTO dummy equal to 1 post-2001. The instrument set pass the weak instrument test and the overidentification tests. The first thing to note is that the coefficient on input tariffs is positive and significant — lower input tariffs lead to lower export prices as we would expect. This is one of the main channels through which input tariffs affect the U.S. price index. Another channel we would expect input tariffs to affect prices is by increasing TFP. We find this channel is difficult to detect in the export price regression. The coefficient on the TFPvariable is positive and insignificant. Another potential issue is that the dependent variable is at the firm-industry level whereas the TFP is at the firm level. Ideally, we would have a measure of TFP for each product that a firm produces but that is not possible. This could be problematic because our average firm-level TFP measure is unlikely to be representative of the smaller export share products, which would result in a downward bias. So in column 2, we exclude any observation for which the Chinese firm's exports within its total U.S. exports that year were less than 5 percent. We see that this changes the sign on the coefficient on TFP to negative, but it is still insignificant. We also experimented with alternative cutoffs, with all giving similar results, and in column 3 we report the case where we only include the firm's largest value HS 6-digit U.S. export.

In the first stage results in the lower panel of Table (8) we see that the total fitted import values have a positive effect on TFP but the gap variable is again insignificant (and close to zero) and switches signs across specifications. We further experiment with including the gap term in the second stage equation in the upper panel of column 4 and find it is insignificant and close to zero. These results suggest that the gap variable does not affect prices consistent with our theory, as prices are chosen after the tariff is known.

[34]We show that these results are robust to expanding the sample to include the universe of exporters (rather than just those in the overlapping sample) in Table **??** in Appendix D. Although the age variable is unavailable for the full sample, we find that the coefficients on all of the other variables are the same as in the overlapping sample.

[35]The results are unchanged for firm unit values that exclude duties because the U.S. MFN import tariffs are very low and have hardly changed over the sample period. For this reason, the MFN tariff is not included on the right of (27).

Table 8: Export Price

Dependent variable	$ln(price_{fgt})$				$ln(s_{fgt})/(1-\rho)$
	(1)	(2)	(3)	(4)	(5)
Sample	All	>0.05 share	Major	>0.05 share	>0.05 share
$ln(TFP_{ft})$	0.167	-0.117	-0.108	-0.118	-1.244***
	(0.161)	(0.130)	(0.138)	(0.131)	(0.198)
$ln(Input\tau_{gt})$	4.890***	4.285**	4.707***	4.353**	
	(1.305)	(1.809)	(1.633)	(2.050)	
$IMR(X_{fgt})$	-0.360	-0.831**	-0.598	-0.826*	0.144
	(0.324)	(0.399)	(0.520)	(0.414)	(0.644)
$IMR(M_{ft})$	-2.438**	-2.937**	-3.860***	-2.998**	3.097
	(1.009)	(1.255)	(1.150)	(1.389)	(8.627)
$ln(Gap_g) \times WTO_t$				-0.011	
				(0.112)	
Overid. J-test	5.989	0.012	0.142	0.000	0.091
[p-value]	[0.050]	[0.994]	[0.932]	[0.992]	[0.763]
Weak instrument F-test	91.816	53.822	20.128	80.548	80.006
First Stage: Dep var: $ln(TFP_{ft})$					
$ln(\hat{M}_{tot,ft})$	0.038***	0.046***	0.043***	0.046***	0.045***
	(0.005)	(0.008)	(0.010)	(0.008)	(0.007)
$ln(\hat{M}_{med,ft})$	0.011	-0.045	-0.059	-0.045	-0.025
	(0.023)	(0.030)	(0.043)	(0.030)	(0.035)
$ln(Gap_g) \times WTO_t$	0.028	-0.022	0.066	-0.022	
	(0.175)	(0.276)	(0.410)	(0.276)	
# obs.	197,292	90,789	45,711	90,789	87,034

Notes: Standard errors are clustered at the IO level. The first stage also includes all second stage variables, not reported to save space.

We suspect that the small and insignificant coefficient on TFP is due to the quality bias. We can correct for this by substituting the pricing equation (27) into the share equation for Chinese firms in 5, while dropping the country supercripts *ij* and replacing the firm index ω with f, we obtain:

$$\begin{aligned} \frac{\ln s_{fgt}}{(1-\bar{\rho})} &= \ln\left(\frac{p_{fgt}}{z_{fgt}}\right) - \ln P_{gt} \\ &= \beta_{gt} + \beta_1 \ln TFP_{ft} + \beta_2 IMR(X_{fgt}) + \beta_3 IMR(M_{fgt}) + (\epsilon_{1fgt} - \ln z_{fgt}) \end{aligned} \tag{37}$$

where $\beta_{gt} \equiv \beta_t - \ln P_{gt}$ are introduced as year×industry fixed effects. We allow firm quality z_{fgt} to vary over time in this specification. But recall that the eror term in the pricing equation (27) is $\epsilon_{1fgt} \equiv \ln z_{fgt}$, so this error cancels out in (37).We should really think of the dependent variable in (37) as reflecting the *quality-adjusted price* of ehch firm (relative to the industry price index) , analogous to Hallak and Schott (2011) and Khandelwal (2010). The dependent variable is the value of a Chinese firms exports in industry g to the U.S. relative total Chinese exports to the U.S. in g divided by one minus the median ρ in industry g (equal to 4.2). We present the results in column 5, where we see that the coefficient on $ln(TFP_f)$is negative and significant equal to -1.2, and we cannot reject that it is equal to -1. We will use this unbiased estimate of β_1 to predict China's export prices.

To summarize this section, the results show that lower Chinese input tariffs increase Chinese firms' imports of intermediate inputs, both on the intensive and extensive margins, and thus increase their TFP. This, in turn, increases their export shares relative to the average firm, as well as increasing the probability of entry into the U.S. market. Lower input tariffs also reduce Chinese firms export prices in the U.S. market. The effect from PNTR is more limited. There is no direct effect on export prices — the only effect we detected was through new entry into exporting. We now turn to evaluate how these effects feed into the U.S. price index.

5 The Impact of China's WTO Entry

With our regression results in hand, we turn to estimating the impact of China's WTO entry on U.S. consumer welfare (treating U.S. importing households and firms as "consumers"). Our starting point is equation (13). We call the first term on the right of (13) the U.S. import price index of common Chinese goods, or $ChinaP_g$; the second term is the common goods price index from all other countries, including the U.S.; the third term is the Chinese variety component of imports, or $ChinaV_g$; and the fourth term is the variety component from other countries (including the U.S.), or $OtherV_g$. We shall regress each of these terms on two instruments that we construct, based on China's entry to the WTO.

The first China WTO instrument is the change in Chinese exporter prices predicted from equation (27), with β_1 from the quality adjusted equation (37). Letting the year 2000 represent the base period 0, we predict prices in 2006 relative to 2000. Then we construct the instrument as follows:

$$Chi\hat{na}P_g \equiv \ln \frac{\hat{P}_{gt}}{\hat{P}_{g0}} = W_{gt}^{ij} \ln \left[\prod_{fh \in \overline{\Omega}_g} \left(\frac{\hat{p}_{fht}}{\hat{p}_{fh0}} \right)^{w_{fht}} \right], \tag{38}$$

where $\overline{\Omega}_g = \Omega_{gt} \cap \Omega_{g0}$ is the set of varieties (at the firm-product level, fh) that were exported in industry g during both year 2006 and 2000, and w_{fht} are the Sato-Vartia weights over these varieties. Note that this instrument uses only the predicted export prices for Chinese firms due to China's WTO entry, and does not include any prices from other exporters to the U.S. nor prices of U.S. domestic producers.

The second WTO instrument uses fitted values from the China firm export participation equation (28), with predicted value $pr\hat{o}b_{fgt}$. We instrument for the Chinese variety term λ_{gt} with the predicted number of exporters obtained by summing the predicted probability of exporting from the participation equation (28). The instrument for the China variety component in (13) is given by:

Table 9: Decomposition of WTO Effect on U.S. Price Index

Independent Variable		US Price Index (1)	$ChinaP_g$ (2)	$OtherP_g$ (3)	$ChinaV_g$ (4)	$OtherV_g$ (5)
	Growth, 2000-2006	0.048	0.014	0.060	-0.032	0.006
$Chi\hat{n}aP_g$	-0.021	1.918**	0.635***	3.087***	-1.080***	-0.724**
		(0.763)	(0.119)	(0.580)	(0.223)	(0.356)
growth x regression coefficient		-0.040	-0.013	-0.064	0.022	0.015
contribution		*61.6%*	*20.4%*	*99.1%*	*-34.7%*	*-23.3%*
$Chi\hat{n}aV_g$	-0.049	0.504***	-0.056***	-0.101**	0.666***	-0.004
		(0.066)	(0.010)	(0.050)	(0.019)	(0.031)
growth x regression coefficient		-0.025	0.003	0.005	-0.032	0.000
contribution		*38.4%*	*-4.3%*	*-7.7%*	*50.7%*	*-0.3%*
Total WTO effect		-0.0641	-0.010	-0.059	-0.010	0.015
N		989	989	989	989	989
R^2		0.113	0.314	0.072	0.625	0.005

Notes: The first row growth rates are the weighted averages of the U.S. price index and each of its four components in equation (13) with the Sato-Vartia weights in equation (14). The first column growth rates are the weighted average of each instrument in equations (38) and (39), using the same Sato-Vartia weights from equation (14). The total WTO effect in the last row is the sum of the China WTO price and variety effects on the U.S. price index, with each effect calculated as the growth rate times the regression coefficient: the price component is 1.918 x $-0.021 = -0.040$; the variety component is 0.504 x$-0.049 = -0.025$.

$$Chi\hat{n}aV_g \equiv \frac{W_{gt}^{ij}}{\hat{\rho}_g - 1}\left[\ln\left(\frac{\sum_{fh\in\overline{\Omega}_g}\widehat{prob}_{fht}}{\sum_{fh\in\Omega_{gt}}\widehat{prob}_{fht}}\right) - \ln\left(\frac{\sum_{fh\in\overline{\Omega}_g}\widehat{prob}_{fh0}}{\sum_{fh\in\Omega_{g0}}\widehat{prob}_{fh0}}\right)\right]. \tag{39}$$

The terms in the square brackets are meant to reflect the terms $\left[\ln(\lambda_{gt}^{ij}) - \ln(\lambda_{g0}^{ij})\right]$ that appears in (13).[36] That term is raised to a power, W_{gt}^{ij}, reflecting China's share of overall U.S. expenditure in industry g, and then dividing by the estimated industry elasticity $\hat{\rho}_g - 1$.

We regress each of the four terms on the right of (13) on these two instruments. By construction, summing coefficients obtained on each instrument across the four regression will give the same results as if we regressed the left-hand side of (13) on these two instruments. In this way, we obtain the *overall* impact of China's entry to the WTO on the U.S. manufacturing price index. We report the results in Table 9. From column 1, we see that a lower China price index due to China's WTO entry (lower $Chi\hat{n}aP_g$) reduces the U.S. price index, and more Chinese export variety due to WTO entry (lower $Chi\hat{n}aV_g$) also lowers the U.S. price index, so U.S. consumers gain due to both lower Chinese export prices and more varieties. To convert these regression coefficients into aggregate effects, we multiply them by the aggregate growth in the two instruments (reported on the left of column 1). The

[36]The estimated probabilities of exporting from (28)are meant to reflect estimated export shares.

sum of these two values indicates that the total WTO effect on the U.S. price index is -0.064 i.e. the U.S. manufacturing price index was 6.4 percent lower in 2006 relative to 2000 due to China joining the WTO.

Now, let's consider how each of these instruments affected each of the components of the U.S. price index. The largest effect is coming through $Chi\hat{n}aP_g$. As expected, the lower China price instrument lowers the China common goods price index (column 2). Interestingly, it also has a very big effect on competitor prices in column 3, which may reflect exit of inefficient competitor firms, lower marginal costs or lower markups.[37] Further, a lower $Chi\hat{n}aP_g$ causes exit of some competitor Chinese firms (column 4) and other competitor firms (column 5).

To interpret the effect of $Chi\hat{n}aP_g$ appearing in Table 10, consider the impact on $OtherP_g$ in column 3, which is the the second term on the right of (13). To be explicit, the coefficients appearing in column 3 are obtained from the regression (ignoring the constant term which is also used):[38]

$$OtherP_g = \sum_{k \in \overline{I}_g^j \setminus i} W_{gt}^{kj} \ln \left(\frac{UV_{gt}^{kj}(\omega)}{UV_{g0}^{kj}(\omega)} \right) = 3.087 Chi\hat{n}aP_g - 0.101 Chi\hat{n}aV_g. \tag{40}$$

Notice that the dependent variable in this regression has the weights W_{gt}^{kj} on each country, but that these weights sum to *less than unity* over the countries $k \in \overline{I}_g^j \setminus i$. Specifically, $\sum_{k \in \overline{I}_g^j \setminus i} W_{gt}^{kj} = 1 - W_{gt}^{ij}$, where W_{gt}^{ij} is the Chinese share in U.S. consumption within industry g. On the other hand, the instruments $Chi\hat{n}aP_g$ and $Chi\hat{n}aV_g$ defined in (38) and (39) have the weights W_{gt}^{ij}, which are just the Chinese share. The coefficients estimates obtain in column (3) are certainly influenced by having weights on the left and the right of (40) that differ from unity.

To illustrate, suppose that we divide $Chi\hat{n}aP_g$ and $Chi\hat{n}aV_g$ by $\overline{W}_t^{ij}$, by which we mean the average over industries g of the Chinese shares W_{gt}^{ij}. That will ensure that the weights $W_{gt}^{ij}/\overline{W}_t^{ij}$ average to unity over the industries used in the regression (40). Analogously, we divide the dependent variable in (40) by the weight $1 - \overline{W}_t^{ij}$, so that the industry weights $W_{gt}^{kj}/(1 - \overline{W}_t^{ij})$ average to unity. With this re-normalization of the left and right-side variables in (40), the regression becomes,

$$\frac{OtherP_g}{1 - \overline{W}_t^{ij}} = 0.582 \frac{Chi\hat{n}aP_g}{\overline{W}_t^{ij}} - 0.019 \frac{Chi\hat{n}aV_g}{\overline{W}_t^{ij}}, \tag{41}$$

which is obtained directly from (40) because the average Chinese share of U.S. consumption over 2000-2006 across manufacturing industries is $\overline{W}_t^{ij} = 0.157$, so that $0.582 = 3.087(0.157/0.843)$ and $-0.019 = -0.101(0.157/0.843)$. With this re-normalization of weights, we see that the actual impact of the Chinese price instrument on the prices of other country's exporters and U.S. firms selling in the U.S. market is a pass-through coefficient of 0.582. That is still a very sizable impact of Chinese

[37] Atkin, Faber, and Gonzalez-Navarro (2016) also find strong effects on competitor prices from entry of foreign retail in Mexico.

[38] We simplify the notation slightly by defining $UV_{gt}^{kj} \equiv \sum_{\omega \in \overline{\Omega}_g^{kj}} w_{gt}^{kj}(\omega) uv_{gt}^{kj}(\omega)$ as the Sato-Vartia index of the unit values $uv_{gt}^{kj}(\omega)$ at the HS 10-digit level ω, aggregated up to HS 6-digit level g.

prices on other prices in the U.S. market, when we consider that the Chinese share is only 0.157 on average. Still, this re-normalization helps us to properly interpret the rather large coefficient of 3.087 appearing in column (3) of Table 10.[39]

Turning to the variety instrument in the lower half of the table, we see that increased Chinese variety due to WTO increases the China variety component in column 4, though this is somewhat muted with a coefficient of 0.67, and it has no effect on other competitors (column 5). We also find that a lower $\hat{ChinaV_g}$ leads to higher Chinese prices in column 2 and competitor prices in column 3, though these are both very small effects, and we suspect that the negative coefficients are due to a quality bias.

The decomposition in Table 9 shows that 62 percent of the China WTO effect on the U.S. price index comes through China's price instrument. The finding that most of the gains come through a lower $\hat{ChinaP_g}$ rather than a lower $\hat{ChinaV_g}$ is somewhat surprising given the large extensive margin of exporting we documented above. In fact, most of the consumer gain comes from China's impact on competitor prices. Given that the gap does not affect $\hat{ChinaP_g}$ at all, since it had no effect on Chinese export prices, as shown in Table 8, we can conclude that at least 60 percent of the WTO effect was indeed due to lower Chinese tariffs on intermediate inputs.

6 Conclusion

In this paper, we quantify the effect of China's WTO entry on U.S. consumers. We construct U.S. manufacturing price indexes by combining highly disaggregated Chinese firm-product data for the period 2000 to 2006, with U.S. import data from other countries and U.S. domestic sales. To take account of new product varieties, we construct exact CES price indexes, which comprise both a price and variety component. We find that China's WTO entry reduced the U.S. manufactured goods price index by 6.4 percent. At least 60 percent of this effect arose through the conventional price index component, despite there being a huge extensive margin of exporting. Importantly, our analysis explicitly takes account of China's trade shock on competitor prices and entry. Our results indicate that lower Chinese export prices due to WTO entry reduced both the China price index and the prices of competitor firms in the U.S., which also led to exit of Chinese competitors and other competitors in the U.S. These effects could be due to less efficient firms exiting the U.S. market, lower marginal costs or lower markups.

Our paper is the first to show that the key mechanism underlying the China WTO effect on U.S. price indexes is China lowering its own import tariffs on intermediate inputs. Indeed, we find that lower Chinese tariffs on its intermediate inputs increased its import values and varieties, which in turn boosted Chinese firms productivity. This higher productivity meant that new firms could enter

[39]The coefficient reported in column (5) can be re-interpreted in the same way, i.e. by multiplying them by (0.157/0.843)=0.186. The coefficients reported in column (2) and (4) do not need any re-interpretation, however, since if we divide the dependent variables and the instruments all by the appropriate weight $\overline{W}_t^{ij} = 0.157$, then the coefficient estimates do not change at all.

the U.S. market and increase their market shares. In addition, lower input tariffs have a direct effect on Chinese export prices. We also allow for China's access to PNTR under WTO to affect Chinese exports — a channel that has received a lot of attention — and consistent with the literature we show that PNTR does result in higher entry into exporting. However, we find no effect of PNTR on Chinese firms TFP or export prices. As such, most of the WTO effect on U.S. price indexes comes through China's lower input tariffs; accounting for at least 60 percent of the overall effect. These gains are additional to the standard gains from trade that arise from reducing iceberg trade costs and in addition to the gains from reducing uncertainty.

References

ACEMOGLU, D., D. AUTOR, D. DORN, G. H. HANSON, AND B. PRICE (2014): "Import Competition and the Great U.S. Employment Sag of the 2000s," NBER Working Papers 20395.

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

AMITI, M., O. ITSKHOKI, AND J. KONINGS (2014): "Importers, Exporters, and Exchange Rate Disconnect," *American Economic Review*, 104(7), 1942–78.

——— (2015): "International Shocks and Domesic Prices: How Large are Strategic Complementarities?," unpublished mimeo.

AMITI, M., AND J. KONINGS (2007): "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia," *American Economic Review*, 97(5), 1611–1638.

ANTRAS, P., T. C. FORT, AND F. TINTELNOT (2014): "The Margins of Global Sourcing: Theory and Evidence from U.S. Firms," Working Paper 154856, Harvard University OpenScholar.

ARKOLAKIS, C., A. COSTINOT, AND A. RODRIGUEZ-CLARE (2012): "New Trade Models, Same Old Gains?," *American Economic Review*, 102(1), 94–130.

ATKIN, D., B. FABER, AND M. GONZALEZ-NAVARRO (2016): "Retail Globalization and Household Welfare: Evidence from Mexico," *Journal of Political Economy*, forthcoming.

AUTOR, D. H., D. DORN, AND G. H. HANSON (2013): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103(6), 2121–68.

BAI, X., K. KRISHNA, AND H. MA (2016): "How You Export Matters: Export Mode, Learning, and Productivity in China," *NBER*, 21164.

BAS, M. (2012): "Input-trade liberalization and firm export decisions: Evidence from Argentina," *Journal of Development Economics*, 97(2), 481–493.

BAS, M., AND V. STRAUSS-KAHN (2014): "Does importing more inputs raise exports? Firm-level evidence from France," *Review of World Economics (Weltwirtschaftliches Archiv)*, 150(2), 241–275.

BLAUM, J., C. LELARGE, AND M. PETERS (2015): "The Gains From Input Trade in Firm-Based Models of Importing," .

BLOOM, N., M. DRACA, AND J. VAN REENEN (2011): "Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity," CEPR Discussion Papers 8236.

BRANDT, L., J. V. BIESEBROECK, AND Y. ZHANG (2012): "Creative Accountng or creative destruction? Firm-level productivity growth in Chinese manufacturing," *Journal of Development Economics*, 97, 339–351.

BRANDT, L., J. VAN BIESEBROECK, L. WANG, AND Y. ZHANG (2012): "WTO Accession and Performance of Chinese Manufacturing Firms," CEPR Discussion Papers 9166.

BRODA, C., AND D. E. WEINSTEIN (2006): "Globalization and the Gains from Variety," *The Quarterly Journal of Economics*, 121(2), 541–585.

CALIENDO, L., M. DVORKIN, AND F. PARRO (2015): "The Impact of Trade on Labor Market Dynamics," NBER Working Paper 21149.

CALIENDO, L., R. C. FEENSTRA, J. ROMALIS, AND A. M. TAYLOR (2015): "Tariff Reductions, Entry and Welfare: Theory and Evidence from the last Two Decades," NBER Working Paper 21768.

DI GIOVANNI, J., A. A. LEVCHENKO, AND J. ZHANG (2014): "The Global Welfare Impact of China: Trade Integration and Technological Change," *American Economic Journal: Macroeconomics*, 6(3), 153–83.

FAN, H., Y. A. LI, AND S. R. YEAPLE (2014): "Trade Liberalization, Quality, and Export Prices," NBER Working Papers 20323.

FEENSTRA, R. C. (1994): "New Product Varieties and the Measurement of International Prices," *American Economic Review*, 84(1), 157–77.

FEENSTRA, R. C., AND D. E. WEINSTEIN (2016): "Globalization, Markups and U.S. Welfare," *Journal of Political Economy*, forthcoming.

FENG, L., Z. LI, AND D. L. SWENSON (2012): "The Connection between Imported Intermediate Inputs and Exports: Evidence from Chinese Firms," NBER Working Papers 18260.

——— (2015): "Trade Policy Uncertainty and Exports: Evidence from China's WTO Accession," .

GOLDBERG, P. K., A. K. KHANDELWAL, N. PAVCNIK, AND P. TOPALOVA (2010): "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India," *The Quarterly Journal of Economics*, 125(4), 1727–1767.

GOPINATH, G., AND B. NEIMAN (2014): "Trade Adjustment and Productivity in Large Crises," *American Economic Review*, 104(3), 793–831.

HALLAK, J. C., AND P. K. SCHOTT (2011): "Estimating Cross-Country Differences in Product Quality," *Quarterly Journal of Economics*, 126(1), 417–474.

HALPERN, L., M. KOREN, AND A. SZEIDL (2015): "Imported Inputs and Productivity," *American Economic Review*, 105(12), 3660–3703.

HANDLEY, K., AND N. LIMÃO (2013): "Policy Uncertainty, Trade and Welfare: Theory and Evidence for China and the U.S," CEPR Discussion Papers 9615, C.E.P.R. Discussion Papers.

HSIEH, C.-T., AND R. OSSA (2011): "A Global View of Productivity Growth in China," Global COE Hi-Stat Discussion Paper Series gd10-166, Institute of Economic Research, Hitotsubashi University.

IACOVONE, L., F. RAUCH, AND L. A. WINTERS (2013): "Trade as an engine of creative destruction: Mexican experience with Chinese competition," *Journal of International Economics*, 89(2), 379–392.

KASAHARA, H., AND B. LAPHAM (2013): "Productivity and the Decision to Import and Export: Theory and Evidence," *Journal of International Economics*, 89(2), 297–316.

KEE, H. L., AND H. TANG (2016): "Domestic Value Added in Exports: Theory and Firm Evidence from China," *American Economic Review*, 106(6), 1402–36.

KHANDELWAL, A. (2010): "The Long and Short (of) Quality Ladders," *Review of Economics Studies*, 77, 1450–1476.

KHANDELWAL, A. K., P. K. SCHOTT, AND S.-J. WEI (2013): "Trade Liberalization and Embedded Institutional Reform: Evidence from Chinese Exporters," *American Economic Review*, 103(6), 2169–95.

KOOPMAN, R., Z. WANG, AND S.-J. WEI (2012): "Estimating Domestic Conent in Exports When Processing is Pervasive," *Journal of Development Economics*, 99, 178–189.

MANOVA, K., AND Z. ZHANG (2012): "Export Prices Across Firms and Destinations," *The Quarterly Journal of Economics*, 127(1), 379–436.

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

MURTAZASHVILI, I., AND J. M. WOOLDRIDGE (2008): "Fixed Effects Instrumental Variables Estimation in Correlated Random Coefficient Panel Data Models," *Journal of Econometrics*, 142(1), 539–552.

OLLEY, S. G., AND A. PAKES (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64(6), 1263–1297.

PIERCE, J. R., AND P. K. SCHOTT (2016): "The Surprisingly Swift Decline of U.S. Manufacturing Employment," *American Economic Review*, 106(7), 1632–1662.

SODERBERY, A. (2015): "Estimating Import Supply and Demand Elasticities: Analysis and Implications," *Journal of International Economics*, 96(1), 1–17.

YU, M. (2015): "Processing Trade, Tariff Reductions, and Firm Productivity: Evidence from Chinese Firms ," *Economic Journal*, 125, 943–988.

A Proof of Lemma

Making use of the revenue in (18), where we replace φ by ω, we can rewrite the unit value as,

$$UV_{gt}^{ij} \equiv \left(\frac{\sum_{\omega \in \Omega_{gt}^{ij}} p_{gt}^{ij}(\omega)^{1-\rho_g}}{\sum_{\omega \in \Omega_{gt}^{ij}} p_{gt}^{ij}(\omega)^{-\rho_g}} \right). \tag{42}$$

From (5) it follows that,

$$p_{gt}^{ij}(\omega) = P_{gt}^{ij} \left(s_{gt}^{ij}(\omega) \right)^{\frac{1}{1-\rho_g}} \iff \sum_{\omega \in \Omega_{gt}^{ij}} p_{gt}^{ij}(\omega)^{-\rho_g} = (P_{gt}^{ij})^{-\rho_g} \sum_{\omega \in \Omega_{gt}^{ij}} \left(s_{gt}^{ij}(\omega) \right)^{\frac{\rho_g}{\rho_g - 1}}.$$

Substituting into (42) and using (4), we obtain,

$$UV_{gt}^{ij} \equiv \left(\frac{\sum_{\omega \in \Omega_{gt}^{ij}} p_{gt}^{ij}(\omega)^{1-\rho_g}}{(P_{gt}^{ij})^{-\rho_g} \sum_{\omega \in \Omega_{gt}^{ij}} \left(s_{gt}^{ij}(\omega) \right)^{\frac{\rho_g}{\rho_g - 1}}} \right) = P_{gt}^{ij} / H_{gt}^{ij}.$$

B Data construction

Herfindahl Indexes:

ASIF data: The firm-level data in this paper comes from Annual Survey of Industrial Firms (ASIF) conducted by the National Bureau of Statistics of China for 1998 to 2007. The survey includes all state-owned enterprises and non-state-owned enterprises with annual sales of RMB five million (or equivalently, about \$800,000) or more. The data set includes information from balance sheets of profit and loss and cash flow statements of firms, and provides detailed information on firms' identity, ownership, export status, employment, capital stock, and revenue. There is a large entry spike of 43 percent in 2004 (more than double in other years). This has been attributed to improvements in the business registry in the industrial census in 2004 so more privately owned firms were included in the survey.

The ASIF data records the firm's main industrial activity at the CIC 4-digit level, which comprise over 500 industrial codes. The ASIF has a firm indicator called id. Some firms change their id because of changes in name, location, or ownership type, yet they are still the same firm As such, these have been mapped to a consistent "panelid" so that each firm maintains a unique identifier. The mapping is done through a two-step procedure. We first link firms by name. For those not linked by name, we then link by zip code, telephone number, and legal person representatives (i.e. two observations are linked if they have the same zip code, telephone number and legal person representative). The number of firms shrinks by 7 percent after the mapping.

Customs data: The customs data includes the universe of firms, reporting import and export values and quantities at the HS 8-digit level. Each firm has a firm identifier called partyid, different from the one assigned in the ASIF data. Some firms change their partyid because of changes in location, firm type, or trade mode. Thus, we also link firms in the customs data to create a firm identifier that is unique over time, as a robustness check. The linking procedure is similar to the one for the ASIF data, except that with the customs data we link firms using the monthly trade data. The number of firms shrinks by 5 percent due to this mapping.

Matching firm id's in customs and ASIF data: Although both the customs data and the ASIF report firm codes, they come from different administrative systems and have no common elements. Thus we construct a concordance between the two data sets using information on firm name as the main matching variable and the zip code and telephone number as a supplement, as in Yu (2015). Using this methodology, we were able to match 32-36 percent of firms in the customs data, which account for 46 percent of the value of exports to the world and 51 percent of the value of U.S. exports.

The details of the matching procedure are as follows. When the firm name is identical in the two data sets the match is straightforward. However, sometimes firm names recorded in the two data sets can be slightly different, and in each data set, the name for the same firm might change slightly over time. This may arise because of typos, inaccurate recordings by the administrative staff, and other reasons. For example, a firm may report its name in the customs data as "Beijing ABC Steel Company" in year 2001 and "Beijing ABC Steel Co. Ltd" in year 2002, while in the ASIF, they report the opposite. In this case, no observations will be matched if the matching is based solely on firm name and year.

To address this issue, we match firms based on all names ever used in the two data sets. More precisely, an id_customs will be matched to an id_asif if one of the names ever used by the id_customs is also ever used by the id_asif. In our previous example, since "Beijing ABC Steel Company" appeared in the customs data in year 2001 and in ASIF in year 2002, firm codes pertaining to this name will be matched. In other words, we exhaust all combinations of id_customs and id_asif if their names are identical at least once in the sample period.

We further use information on zip codes and telephone numbers to aid in the matching process, given that the telephone number is unique within a region. We match two firm codes if they have identical zip codes and the last seven digits of the telephone number. Using the last seven digits of the telephone number is based on several considerations. First, telephone numbers are reported in different formats in the two data sets. Area codes are included at the beginning of each telephone number in the customs data, but not in ASIF. Taking the last seven digits makes the formats in the two data sets comparable. Second, during the sample period, some large cities changed their telephone numbers from seven to eight digits by adding one digit before the original seven-digit number. As a result, the reported telephone number for the same firm may change over time. Taking the last seven digits solves this problem. We dropped the zip codes and telephone numbers that are inconsistent with China's regulations (e.g. ones that only have one digit, or contain non-numerical symbols). Finally, similar to the matching by name, we also exhaust all combinations of codes if they have the same zip code and telephone numbers at least once in the sample period.

Table B1: Number of Firms

year	# firms in ASIF	# exporters in ASIF	# exporters in customs	# US exporters in customs	share of US export value in overlapping sample
2000	138,431	38,854	62,746	23,437	0.41
2001	151,017	43,978	68,487	26,172	0.44
2002	162,780	49,824	78,613	31,835	0.47
2003	179,151	56,737	95,690	39,556	0.50
2004	252,540	81,435	120,590	49,878	0.55
2005	250,909	84,251	144,031	63,193	0.53
2006	277,863	89,329	171,205	76,081	0.53

The total number of exporters is reported in Table B1. The striking pattern to emerge from Table B1 is the massive net entry into exporting. First, note that the number of firms in the ASIF data doubled over the sample period, with 278,000 firms by 2006. But since only firms with at least 5 million RMB are included in the sample, some of this increase in firm numbers in the sample is due to firms crossing this threshold. It does, however, comprise a large portion of the manufacturing sector. Comparing the ASIF data with the 2004 census, we find these data cover 91 percent of the manufacturing sector in terms of output, 71 percent in terms of employment, and 98 percent in terms of export value (see Brandt, Biesebroeck, and Zhang (2012) for more details.) Of more relevance for our study is the pattern for exporters. In the customs data, we see that the number of U.S. exporters more than tripled over the sample period. This represents actual net entry into the market since the customs data represents the universe of exporters. This pattern is also mirrored for exporting to the world, and in the overlapping sample.

Product concordances: We make the China HS 8-digit codes consistent over time, using a concordance from China Customs. We map all HS8 codes to their earliest code in the sample.

The Chinese Industrial Classifications (CIC) were revised in 2003, so we used a concordance from the China National Bureau of Statistics (NBS) to bridge the two sets of codes, which we mapped to the new codes. As usual with concordances, we found that some of these mappings were not one-to-one so this required some groupings of the codes. Our concordance is consistent with Brandt, Biesebroeck, and Zhang (2012). The manufacturing codes comprise those CIC codes that begin with 13 to 44: there are 502 distinct CIC manufacturing codes in the pre-2003 revision and 432 after we group some industry codes to take care of the many-to-many mappings.

We mapped the HS8 codes to CIC codes using a partial concordance from NBS, and completed the rest manually. The mapping between HS8 and IO codes uses the HS2002 version so we converted that to HS 2000 codes. We built on a concordance from HS6 2002 to IO from one constructed manually by Rudai Yang, Peking University, using a mapping from HS to SITC to IO. The mappings from IO_2002 to CIC_2003 and IO_2002-CIC_2002 were downloaded from Brandt, Biesebroeck, and Zhang (2012) (http://www.econ.kuleuven.be/public/n07057/China/).

C U.S. Consumption

To construct China's share in total U.S. consumption in equations (12) and (13), at the HS 6-digit level, we combine production, import and export data, and define S_{gt}^{ij}= (China's imports to U.S. / (production – exports + imports). However, the production data is only available in NAICS 6-digit (from NBER database) and the trade data is at HS10 level. There is a concordance that maps HS10 to NAICS, but it is more difficult to go from NAICS to HS6 because of the usual many-to-many issue. To overcome this problem, we follow Feenstra and Weinstein (2016) (p45 Appendix) as follows. First, denote a NAICS industry by k, we define

$$share^k = \left(\frac{supply^k}{supply^k + imports^k} \right)$$

which can easily be constructed at NAICS (with supply defined as production less exports).

Second, we assume that the share in NAICS equals the share in HS10. Imposing that equality, and rearranging we get

$$supply^{HS10} = \left(\frac{share^k}{1 - share^k} \right) * Imports^{HS10}.$$

Once we have U.S. supply at HS10, we can then construct China's share in U.S. consumption at any level of HS. Note that in some cases there were some negative supply values at the NAICS levels (15 out of 447), in which case we replaced those negative values with zeros. Once we have supply at the HS10 level, it is straightforward to construct China's share, which equals imports from China into the U.S./(U.S. supply + total imports).

D Additional Results

Table D1: Export Decompositions

Sample	Type of trade	Total export growth	EM proportion	Share in total US	Share of total US growth
Full sample	Variety defined at HS6-firm	2.86	0.83		
	Variety defined at HS4-firm	2.86	0.79		
	Variety defined at HS2-firm	2.86	0.74		
	Time consistent id (unconcorded hs8)	2.86	0.87		
	Time consistent id (concorded hs8)	2.86	0.84		
Sub-samples	Nonprocessing trade	3.39	0.91	0.37	0.39
	Consumer goods	4.00	0.86	0.39	0.33
	Nontraders	3.66	0.88	0.85	0.90
	Private firms	4.32	0.89	0.84	0.92
	Foreign firms	4.00	0.87	0.67	0.73
	ASIF overlap	3.30	0.83	0.53	0.57

Notes: Private firms exclude SOE. ASIF overlap is the set of exporters for which we could match in the industrial data

Table D2: Production Coefficients for Chinese Plants: 1998-2006

	Chinese Industrial Classification	Olley-Pakes Labor β_l	Olley-Pakes Capital β_k
13	Processing of Foods	0.38	0.37
14	Manufacture of Foods	0.46	0.44
15	Manufacture of Beverages	0.46	0.48
17	Manufacture of Textile	0.44	0.36
18	Manufacture of Apparel, Footwear & Cap	0.52	0.31
19	Manufacture of Leather, Fur, & Feather	0.47	0.33
20	Processing of Timber, Manufacture of Wood, Bamboo, Rattan, Palm & Straw Products	0.40	0.37
21	Manufacture of Furnitur	0.55	0.32
22	Manufacture of Paper & Paper Product	0.41	0.37
23	Printing, Reproduction of Recording Medi	0.42	0.44
24	Manufacture of Articles For Culture, Education, & Sport Activities	0.48	0.32
25	Processing of Petroleum, Coking, & Fue	0.23	0.47
26	Manufacture of Raw Chemical Materials	0.33	0.44
27	Manufacture of Medicines	0.44	0.39
28	Manufacture of Chemical Fibers	0.41	0.55
29	Manufacture of Rubber	0.40	0.52
30	Manufacture of Plastics	0.38	0.38
31	Manufacture of Non-metallic Mineral goods	0.33	0.46
32	Smelting & Pressing of Ferrous Metals	0.45	0.37
33	Smelting & Pressing of Non-ferrous Metals	0.34	0.34
34	Manufacture of Metal Products	0.41	0.38
35	Manufacture of General Purpose Machiner	0.41	0.40
36	Manufacture of Special Purpose Machinery	0.42	0.42
37	Manufacture of Transport Equipment	0.46	0.42
39	Electrical Machinery & Equipment	0.41	0.43
40	Computers & Other Electronic Equipment	0.49	0.41
41	Manufacture of Measuring Instruments & Machinery for Cultural Activity & Office Work	0.37	0.38
42	Manufacture of Artwork	0.44	0.31
	Average: all manufacturing	0.42	0.40

Notes: We estimate the production coefficient following Olley and Pakes (1996).

Table D3: Export Participation- Full Sample

Dependent variable Y_{fgt} =1 if exports>0	probit				OLS
	(1)	(2)	(3)	(4)	(5)
$ln(Input\tau_{gt})$	-1.266***	-5.669***	-5.600***	-5.792***	-1.626***
	(0.236)	(0.675)	(0.061)	(0.064)	(0.163)
$ln(Gap_g)$	-0.353***	-0.252***			
	(0.108)	(0.087)			
$ln(Gap_g) \times WTO_t$	0.417***	0.306***	0.309***	0.314***	0.057***
	(0.096)	(0.094)	(0.010)	(0.010)	(0.019)
$Foreign_f$	0.196***	0.197***	0.197***	0.197***	0.063***
	(0.008)	(0.008)	(0.002)	(0.002)	(0.003)
$ln(Output\tau_{gt})$				0.227***	
				(0.022)	
industryeffects	no	HS4 fe	HS6 re	HS6 re	HS6 fe
#obs.	7,366,709	7,366,709	7,366,709	7,365,585	7,366,709

Notes: This includes the universe of Chinese exporters to the U.S. All equations include year fixed effects. Standard errors are clustered at the firm level.