# DISCUSSION PAPER SERIES

DP13904
(v. 2)

## Markets and Markup: A New Empirical Framework and Evidence on Exporters from China

Meredith A Crowley, Giancarlo Corsetti, Lu Han and Huasheng Song

INTERNATIONAL MACROECONOMICS AND FINANCE

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

CEPR

# Markets and Markup: A New Empirical Framework and Evidence on Exporters from China

*Meredith A Crowley, Giancarlo Corsetti, Lu Han and Huasheng Song*

# Markets and Markup: A New Empirical Framework and Evidence on Exporters from China

## Abstract

Exporting firms frequently change the set of foreign markets they serve: changes in the pattern of destinations for a firm's product convey information on unobservable factors affecting pricing and market participation. Building on this insight, we show how to construct a "trade-pattern" fixed effect estimator that helps reduce omitted variable and selection biases in analyses of pricing-to-market. Using this estimator and a new product classification, we document substantial markup elasticities to the exchange rate among exporters of highly differentiated goods, accounting for half of China's exports. Conversely, we find little evidence of pricing-to-market in the trade of less differentiated goods.

Meredith A Crowley - crowley.meredith@gmail.com
*Faculty of Economics, University of Cambridge and CEPR*

Giancarlo Corsetti - giancarlo.corsetti@gmail.com
*Faculty of Economics, University of Cambridge and CEPR*

Lu Han - lu.han2@liverpool.ac.uk
*School of Management, University of Liverpool and CEPR*

Huasheng Song - songzju@zju.edu.cn
*Zhejiang University*

# Markets and Markups:

## A New Empirical Framework and Evidence on Exporters from China

Giancarlo Corsetti

Cambridge, Cambridge-INET and CEPR

Meredith Crowley

Cambridge, Cambridge-INET and CEPR

Lu Han

Liverpool and CEPR

Huasheng Song

Zhejiang University

28 May 2021

### Abstract

Exporting firms frequently change the set of foreign markets they serve: changes in the pattern of destinations for a firm's product convey information on unobservable factors affecting pricing and market participation. Building on this insight, we show how to construct a "trade-pattern" fixed effect estimator that helps reduce omitted variable and selection biases in analyses of pricing-to-market. Using this estimator and a new product classification, we document substantial markup elasticities to the exchange rate among exporters of highly differentiated goods, accounting for half of China's exports. Conversely, we find little evidence of pricing-to-market in the trade of less differentiated goods.

Giancarlo Corsetti, Faculty of Economics, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DD, United Kingdom. Email: giancarlo.corsetti@gmail.com. Meredith Crowley, Faculty of Economics, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DD, United Kingdom. Email: crowley.meredith@gmail.com. Lu Han, Management School, University of Liverpool, Chatham St, Liverpool, L69 7ZH, United Kingdom. Email: hanlulong@gmail.com. Huasheng Song, CRPE and School of Economics, Zhejiang University, 38 Zheda Road, Hangzhou, Zhejiang 310028, China. Email: songzju@zju.edu.cn.

# Contents

# 1  Introduction

Firms that export to more than one country account for the lion's share of cross-border trade. Serving multiple markets, these firms face demand conditions and costs shocks that may be specific to an export destination and are inherently time-varying. From the perspective of an exporter, a changing local economic environment systematically creates opportunities to raise profits, or induces the need to contain losses, through destination-specific adjustment of export prices, i.e., by engaging in pricing-to-market (see, e.g., Krugman (1986), Dornbusch (1987), Goldberg and Knetter (1997) and, for a recent reconsideration, Burstein and Gopinath (2014)).[1]

The increasing availability of high-dimensional administrative customs databases has provided a wealth of new insights about the pricing behavior of exporters, stressing that larger, more highly productive firms adjust markups more (see, e.g., Berman, Martin and Mayer (2012), Chatterjee, Dix-Carneiro and Vichyanond (2013), Fitzgerald and Haller (2014), De Loecker, Goldberg, Khandelwal and Pavcnik (2016), Amiti, Itskhoki and Konings (2014, 2019, 2020)). This literature has broken new ground in documenting significant heterogeneity in markups and markup elasticities across firms by directly employing estimates of the firm's (unobservable) productivity and marginal costs, or by indirectly controlling for unobservables with fixed effects. At the same time, the wealth of information on prices at fine levels of firm, product, and market disaggregation has created opportunities for new solutions to the problem of controlling for unobservable determinants of pricing, as well as for investigating heterogeneity in pricing behavior along new dimensions.

The point of departure for our study is a stylized fact in international trade that has been overlooked in micro studies of pricing to market. Namely, the *set of foreign markets* in which a firm sells a particular product(s), which we define as the "trade pattern" for a firm and product, changes from year to year. Moreover, certain trade patterns tend to *reappear time and again over the course of a firm's export life.*[2] The theory of firms in international trade has long recognized that firms may choose to enter a new market or discontinue exports to a destination in response to variation in idiosyncratic demand or marginal cost, the very factors that most likely affect pricing (Melitz and Redding (2014)). This suggests that the changes in a firm's trade patterns, and the reoccurrence of certain trade patterns over time that we observe in the data, are not purely random. Rather, they contain valuable information about unobservable economic variables that, in addition to determining market participation, affect pricing in all the foreign markets where

---

[1]Pricing-to-market is a standard feature in open macro models, which increasingly feature firm dynamics and competition (see, e.g., Bergin and Feenstra (2001) and Atkeson and Burstein (2008)), vertical interactions of exporters with local producers and distributors (see, e.g., Corsetti and Dedola (2005)), and nominal rigidities in either local or a third-country vehicle currency (Corsetti, Dedola and Leduc (2008), Gopinath (2015) and Gopinath, Boz, Casas, Díez, Gourinchas and Plagborg-Møller (2020)).

[2]Han (2018) offers a pioneering study of systematic changes in the extensive margin, documenting the same stylized fact in UK data.

the firm continues to make sales. Our main methodological contribution consists of showing how a firm's trade patterns can be integrated into the design of a fixed effect estimator of markup elasticities – the Trade Pattern Sequential Fixed Effect (TPSFE) estimator – in order to reduce omitted variable and selection biases.

Our substantive contribution is a novel set of estimates of pricing to market by exporters from China, drawing distinctions across estimates for different types of products and firms. Specifically, our estimates unveil significant evidence of markup adjustment in response to movements in the bilateral exchange rate not only for large, State-Owned, and Foreign-Invested enterprises, but also for firms exporting highly differentiated products. For these exporters, our estimated elasticities are up to 50% higher than those from fixed effects estimators used in the literature. In light of the prevailing view that most Chinese exporters set the prices of their products in dollars, our findings suggest that, for highly differentiated goods and large exporters, pricing in a dominant currency does not necessarily preclude markup and price adjustments to local economic conditions.[3] Conversely, for private firms and exporters of less differentiated merchandise, we find little or no evidence of pricing to market. For these exporters, covering roughly half of China's global exports, our study provides independent evidence of *global* pricing, closely in line with the properties of an international price system dominated by the US dollar analyzed by Gopinath (2015) and Gopinath, Boz, Casas, Díez, Gourinchas and Plagborg-Møller (2020).

Our study is organized in three core chapters. First, to gain theoretical insight into our estimator, we rely on a standard open economy macro model with multiple countries, heterogeneous firms, and variable markups solved in partial equilibrium. The model shows how the interaction of the firms' exporting and pricing decisions creates multiple and complex sources of bias in estimators of pricing to market. Our main innovation is the construction of a firm-product-destination-specific fixed effect indexed by trade patterns. This novel fixed effect restricts the estimator to compare observations conditional on a firm selling to the *same* set of export markets. We show both analytically and based on model-simulated data that our new estimator enhances transparency of identification, and can significantly reduce biases associated with unobservable time-varying factors at the firm, product and even destination level, relative to conventional fixed effect estimators.[4]

Second, we introduce a new product classification to obtain a better empirical proxy for market power. In addition to varying with an exporter's size, market power can also be expected to vary

---

[3]Although the invoicing currency of Chinese exports is not recorded in our dataset, the value of exports is reported in US dollars and the US dollar is widely-held to have been the principal invoicing currency for Chinese exports throughout this period. See appendix G.8 for evidence on dollar invoicing. We report results from applying our TPSFE estimator conditional on price changes in dollars.

[4]We provide a formal discussion on how our approach can be thought as a variation of the control function approach by Heckman (1979) and Kyriazidou (1997) in Appendix B. We show how our fixed effect approach connects to more conventional fixed effects approaches in Appendix C, and how it is related to the productivity estimation approach by De Loecker, Goldberg, Khandelwal and Pavcnik (2016) in Appendix E.

with the type of product and the market segment in which a firm operates. Most intuitively, everything else equal, firms selling highly differentiated products may be better able to segment the market and exercise monopoly power than firms exporting less differentiated goods. We exploit information contained in Chinese customs records – specifically, Chinese linguistic particles that reflect a good's physical attributes and act as measures for numbers of items – to construct a comprehensive, general, and exogenous product classification that distinguishes between goods with high versus low degrees of differentiation. A key advantage of our classification is that it articulates the large class of differentiated goods obtained by following the approach of Rauch (1999) into two large subgroups. These subgroups, of high- and low-differentiation products, can then be combined with other criteria (e.g., firm size) and classifications (e.g., functional end-use of a product) to further refine firms into smaller groups according to their (potential) market power.[5]

Finally, we apply our TPSFE estimator together with our linguistics-based product classification to estimate markup elasticities to the exchange rate using the universe of Chinese exports over 2000-2014. To reduce the incidence of nominal rigidities, we follow the method pursued by Gopinath, Itskhoki and Rigobon (2010) and carry out our analysis conditional on detecting a change in the dollar price of a product. Our framework allows us to study markup elasticities according to the types of products exporters sell, by firm ownership, and by exporter size (measured by global export revenue). We are thus able to document significant heterogeneity in pricing-to-market across firms and products.

According to our estimates, over the period of the managed float of the renminbi (2006-2014), in response to a 10% appreciation of a destination country's currency against the Chinese currency, Chinese exporters of highly differentiated goods raised their markups by 2% while exporters of goods characterized by little or no differentiation raised their markups by a mere 0.6%. This means that, conditional on a dollar price change following an exchange rate movement, firms exporting highly differentiated goods kept their prices *measured in local currency* significantly more stable than firms exporting less differentiated products. Other things equal, pricing-to-market is more pervasive in markets in which firms face less direct competition because of the highly differentiated nature of their output. Consistently, we detect even larger markup elasticities when we examine subsamples of our data that include large (as opposed to small) exporters, firms selling consumption (as opposed to intermediate) goods, and firms with complex corporate structures such as State-Owned or Foreign-Invested Enterprises (as opposed to private firms). In these subsamples, our

---

[5]Applying Rauch (1999)'s categories to the Chinese Customs Database, we find about 80 percent of Chinese export value is classified as differentiated because these products are not traded on organized exchanges or in markets with published reference lists. According to our linguistics-based classification, about half of this, amounting to 39 percent of Chinese export value, is actually highly differentiated, while 41 percent exhibits low differentiation. Furthermore, we find that many products which are left unclassified by Rauch can be classified as high or low differentiation goods according to our classification.

estimated markup elasticities for highly differentiated products almost double, ranging from 0.3 to 0.4. Conversely, for low-differentiation goods, markup elasticities typically remain close to zero.

In estimating markup elasticities to bilateral exchange rates, we appeal to the classical view in international economics that, after controlling for a firm's costs, currency movements provide a measure of destination-specific shifts in the demand faced by exporters (see, e.g., Goldberg and Knetter (1997)). Consequently, consistent with the classical view, changes in relative export volumes across markets should be positively correlated with predicted changes in relative markups arising from exchange rate changes. We complete our analysis with an empirical exercise showing that this correlation, as expected, is positive in the data. In addition, we show that the strength of the correlation is weaker for products and groups of firms with higher markup elasticities to the exchange rate. This suggests that exporters with market power that respond to currency movements by adjusting their markups substantially also keep their export quantities relatively stable across destinations, especially when compared to firms operating in more competitive markets. For State-Owned and Foreign-Invested Enterprises (SOEs and FIEs), each percentage increase in markups in a destination corresponds to a rise in exports by mere 0.3%; in contrast, for private firms, the corresponding elasticity is as high as 5.2%.

As a final point, we note that SOEs, FIEs, and, more generally, large firms selling highly differentiated products are those most likely to be integrated into complex supply chain structures. They arguably set prices responding to a host of administrative, taxation, and institutional demands. Altogether, this suggests that their demand and costs are more exposed to variation in destination-specific factors relative to the other firms in the sample. Since controlling for destination-specific time-varying unobservable factors is precisely the comparative advantage of our TPSFE procedure, we would expect our TPSFE estimates for these groups of firms to be distinctively different from those obtained from conventional fixed effects. Our evidence squares remarkably well with this conjecture. A comparison among the TPSFE and fixed effects estimators used in leading studies shows that divergences are especially large and significant for SOEs and FIEs, while very small and insignificant for smaller exporters and exporters of low differentiation goods.

In addition to the contributions referred to above, our paper is also closely related to the literature that examines the effects of extensive margin adjustments of aggregate, product- and firm-level exports on trade elasticities and exchange rate pass through (Helpman, Melitz and Rubinstein (2008), Nakamura and Steinsson (2012), Bas, Mayer and Thoenig (2017) and Fitzgerald and Haller (2018)) as well as studies specifically devoted to assess how Chinese firms respond to changes in foreign trade policy (Khandelwal, Schott and Wei (2013), Crowley, Meng and Song (2018)) and exchange rates (Dai and Xu (2017)). Finally, our paper naturally complements the empirical study by Manova and Zhang (2012), who establish a set of stylized facts on exporters from China pointing to systematic price differences across countries and the potential relevance of

destination specific variations in demand and costs.

The rest of the paper is organized as follows. Section 2 draws on the Chinese dataset to present the general stylized fact that motivates our analysis: the variability in the set of destinations reached by exporters over time. Section 3 carries out a model-based analysis of biases that potentially plague studies of pricing to market, and discusses the theoretical foundations of our approach to its estimation. Section 4 introduces our product classification and discusses its properties relative to alternative classifications. Section 5 presents our empirical results. Section 6 concludes.

# 2   The Variable Trade Patterns of Exporting Firms

We motivate our empirical analysis of international pricing by highlighting a key salient feature of firms' engagement in international trade: the set of foreign markets reached by an exporting firm changes frequently over time, with particular sets of foreign markets featuring repeatedly for a given firm. In this section, we briefly describe our dataset and document that most of China's exports originate from firms selling to multiple destinations. We then present statistics on the variability of firms' product-level trade patterns, focusing on the number of unique trade patterns observed during a firm's export history. We conclude with a stylized example, introducing at an intuitive level the idea we exploit in our analysis.

## 2.1   Data

Our analysis uses the Chinese Customs Database, the universe of annual import and export records for China from 2000 to 2014 along with annual macroeconomic data from the World Bank.[6] The final estimation dataset consists of over 200,000 multi-destination exporters, around 8,000 HS08 products, and 152 foreign markets over 15 years. The Chinese Customs Database reports values and quantities of exports in US dollars by firm (numerical ID and name) and foreign destination country at the 8-digit Harmonized System product level over 2000-2014.[7]

Chinese exports are thus structured as a panel with four dimensions—firm, product, destination

---

[6]Details regarding the macroeconomic data and information about the Chinese Customs Database are given in appendix G.

[7]The database is available at the monthly frequency during the period 2000-2006 and annual frequency during the period 2007-2014. We aggregate the monthly data for 2000-2006 to the annual level in this study. Because no information on the currency of invoicing is reported in the Chinese Customs Database, we turn to administrative data from Her Majesty's Revenue and Customs (HMRC) in the UK to provide information about the currency of invoicing of Chinese exports to the UK so that we can place our results in context. See Appendix G.8. We should nonetheless note upfront that, because our TPSFE estimator differences out the common components across destinations, using prices denominated in dollars with dollar-destination exchange rates versus using prices denominated in renminbi with renminbi-destination exchange rates in the estimation procedure yields exactly the same estimates.

market, and time. However, some specific characteristics of the Chinese customs data allow us to obtain a classification of types of products by their differentiation and types of firms by the nature of their commerce. Most notably for our purposes, each observation in the database contains (a) the Chinese measure word in which quantity is reported, (b) an indicator of the form of commerce for tax and tariff purposes, and (c) a categorization based on the registration type of the exporting firm.[8] We will see that all these entries can be exploited to obtain information on the firm's market power in its export markets.

Like other firm-level studies using customs databases, we use unit values as a proxy for prices. However, the rich information on forms of commerce and Chinese measure words enables us to build more refined product-variety categories than prior studies have used. Specifically, we define the product identifier as an 8-digit HS code plus a form of commerce dummy. The application of our product-variety definition generates 14,560 product-variety codes in our final estimation dataset as opposed to 8,076 8-digit HS codes reported in the database.[9] Throughout our study, we will use the term "product" to refer to these 14,560 product-varieties. This refined product measure allows us to get a better proxy of prices for two reasons. First, the inclusion of the information on form of commerce helps to distinguish subtle differences of goods being sold under the same 8-digit HS code. Second, as discussed later on in the text, the extensive use of a large number of measure words as quantity reporting units makes unit values in Chinese data conceptually closer to transactions prices than unit values constructed with other national customs datasets.[10]

## 2.2 Quantitative Importance of Multi-destination Exporters

An overwhelming majority of Chinese exporters serve multiple foreign destinations. A similar pattern has been documented for other markets, most notably for France by Mayer, Melitz and Ottaviano (2014), suggesting that this is a core feature of foreign market participation by exporting

---

[8]The form of commerce indicator records the commercial purpose of each trade transaction including "general trade," "processing imported materials," and "assembling supplied materials." Essentially, a firm can produce the same HS08 product under different tax regulations depending on the exact production process used. We simplify different tax treatments into a form of commerce dummy equal to 1 if the transaction is "general trade" and 0 otherwise. The registration type variable contains information on the capital formation of the firm by eight mutually-exclusive categories: state-owned enterprise, Sino-foreign contractual joint venture, Sino-foreign equity joint venture, wholly foreign-owned enterprise, collective enterprise, private enterprise, individual business, and other enterprise. In our analysis, we aggregate the three types of foreign-invested firms, namely wholly foreign-owned enterprises, Sino-foreign contractual joint ventures and Sino-foreign equity joint ventures, into one category dubbed "foreign-invested enterprises." We group minority categories including collective enterprises, individual businesses and other enterprises into one category and refer to them as "other enterprises."

[9]When we clean the data, the number of HS08 products and HS08 product-varieties declines with the number of observations. These numbers refer to products and product-varieties in the final estimation dataset.

[10]Important previous studies have constructed unit values (export value/export quantity) from data in which quantity is measured by weight (Berman, Martin and Mayer (2012)) or in a combination of weights and units (Amiti, Itskhoki and Konings (2014)).

Table 1: Multi-destination exporters (2007)

| | Number of Foreign Destinations | | | | Total |
| --- | --- | --- | --- | --- | --- |
| | 1 | 2-5 | 6-10 | 10+ | |
| (a) by Share of Exporters | 27.2 | 33.1 | 14.7 | 25.0 | 100.0 |
| (b) by Share of Export Values | 5.4 | 11.9 | 10.4 | 72.3 | 100.0 |
| (c) by Share of Number of Annual Transactions | 3.0 | 8.0 | 7.8 | 81.2 | 100.0 |

Note: Each cell in the top row is the proportion of exporters in the Chinese customs data in 2007 that fall under the relevant description. The middle and bottom rows present the corresponding proportions for export value and count of annual export transactions, respectively.

firms. Based on our dataset, table 1 presents a breakdown of the proportion of exporting firm, export values, and count of annual transactions according to the number of destinations served in 2007. Overall, we see that around three-quarters of exporters reach more than one destination (row a). These firms are responsible for 94.6% of export value (row b) and 97.0% of annual transactions (row c). Conversely, the 27.2% of exporters that sell to a single destination, comprised only 5.4% of Chinese export value and 3.0% of export transactions in 2007. While we present a single year snapshot from our dataset in the table, the patterns in year 2007 are by no means special: the shares of exporters, export value, and export transactions by count of destination markets remain relatively stable over our sample period, 2000-2014.

## 2.3 The Frequently Changing Trade Patterns of a Firm's Product Sales: Stylized Facts

The key stylized fact revealed by the customs records of multi-destination exporters is that the sets of foreign markets a firm reaches with a given product are highly volatile. While this has remained relatively under-researched by the literature, Han (2018) has provided detailed comparative evidence using Chinese and UK data.[11]

To gain insight and introduce concepts that we will use extensively in our study, in Figure 1 we provide the simplest stylized example of an exporter selling a specific product to different sets of foreign markets over a five-year span—similar to what can be found in the dataset. The figure refers to a firm exporting a product to three destination countries, A, B and C, in different combinations. Empty elements indicate that there is no trade in a period. We define the set of markets active at a firm-product level in one period as a *trade pattern*. In our stylized example,

---

[11]A number of papers have studied the within-year volatility in trade flows in a given destination due to transaction costs (see e.g., Alessandria, Kaboski and Midrigan (2010) and Kropf and Sauré (2014)). We inspect the volatility in trade flows across years stemming from firms selecting in and out of destination markets and illustrate how the time-varying trade pattern of firms can be used to construct a novel fixed effect estimator to improve the estimation of markup elasticities.

the firm has three *unique* trade patterns, A-B, A-C, A-B-C over the course of its five year trade in that product. Notably, however, two of these firm's product-level trade patterns repeat. The pattern A-C repeats in periods 2 and 4; A-B-C repeats in periods 3 and 5.
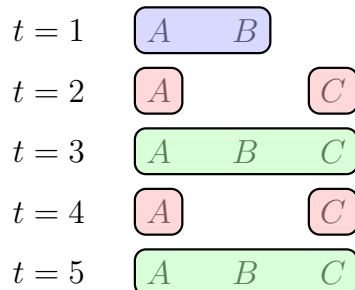


Figure 1: Example of an observed trade pattern

The volatility of firm-product trade patterns in the Chinese Customs Database is summarized in Table 2. To construct this table, we begin with the universe of firm-product pairs in the Chinese Customs Database over the sample period 2000-2014. To build our table, we first drop all firm-product pairs that appear only once in the 15 year timespan of our dataset—since there is no time variation associated with these pairs. We next place firm-product pairs into bins according to the total number of years $(x)$ for which foreign sales were observed. In the last row of the table, we report the share of firm-product pairs with observed foreign sales in 2, 3,...,15 years. Unsurprisingly, firm-year pairs with observed sales in only a few years are the most common; in our dataset (after removing the single period pairs) about 60% of firm-product pairs are observed for between two and four years (29.3+17.9+12.0). At the other extreme, only 1.1% of firm-product pairs are observed in every year.

In the columns of the table, for each number of exporting years, we calculate the share of firm-product pairs associated with a specified number of unique trade patterns, $y$. For example, the firm-product pair in the stylized example of Figure 1 has three unique trade patterns, {A-B, A-C, A-B-C}, over five years of sales abroad. In the table, our stylized firm-product would be included in the cell reporting that 14.1% of firm-product pairs observed for five years have three unique trade patterns. More generally, entries in the first row of the table report the share of firm-product pairs that have perfectly stable trade patterns over the course of their entire export life. At the other extreme, the diagonal elements contain firm-product pairs with extremely volatile trade patterns – these firm-products have a different, unique trade pattern in every year of export life. Most crucially for our purposes, the statistics above the diagonal shows that the majority of firm-product pairs have a smaller number of unique trade patterns than their total number of exporting years. This means these firms export a particular product to the same set of destinations

Table 2: Number of Unique Trade Patterns

|  | Total Number of Exporting Years ($x$) | | | | | | | | | | | | | | |
| Number of Unique Trade Patterns ($y$) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 35.9 | 26.6 | 22.4 | 19.3 | 16.7 | 14.0 | 11.8 | 10.3 | 8.8 | 7.7 | 6.2 | 5.5 | 5.1 | 4.7 | 23.4 |
| 2 | 64.1 | 23.2 | 16.5 | 13.0 | 10.8 | 9.1 | 7.7 | 6.7 | 6.0 | 5.4 | 4.6 | 4.3 | 3.8 | 3.8 | 28.5 |
| 3 |  | 50.2 | 20.3 | 14.1 | 11.0 | 8.9 | 7.1 | 6.3 | 5.4 | 4.7 | 3.9 | 3.5 | 3.0 | 3.1 | 15.0 |
| 4 |  |  | 40.8 | 17.6 | 12.2 | 9.3 | 7.3 | 6.2 | 5.1 | 4.3 | 3.6 | 2.9 | 2.6 | 2.7 | 8.9 |
| 5 |  |  |  | 35.9 | 15.8 | 11.1 | 8.3 | 6.6 | 5.3 | 4.5 | 3.7 | 2.9 | 2.7 | 2.3 | 6.1 |
| 6 |  |  |  |  | 33.4 | 14.9 | 10.1 | 7.7 | 6.2 | 5.0 | 3.8 | 3.0 | 2.4 | 2.2 | 4.5 |
| 7 |  |  |  |  |  | 32.7 | 13.8 | 9.6 | 7.3 | 5.5 | 4.5 | 3.7 | 2.9 | 2.2 | 3.5 |
| 8 |  |  |  |  |  |  | 33.9 | 13.7 | 9.4 | 7.0 | 5.2 | 4.2 | 3.3 | 2.3 | 2.8 |
| 9 |  |  |  |  |  |  |  | 33.0 | 13.5 | 9.1 | 6.7 | 5.0 | 3.7 | 2.7 | 2.0 |
| 10 |  |  |  |  |  |  |  |  | 33.3 | 13.2 | 8.9 | 6.8 | 5.1 | 3.2 | 1.6 |
| 11 |  |  |  |  |  |  |  |  |  | 33.6 | 13.1 | 9.0 | 6.5 | 3.5 | 1.1 |
| 12 |  |  |  |  |  |  |  |  |  |  | 35.9 | 13.7 | 8.4 | 5.1 | 0.9 |
| 13 |  |  |  |  |  |  |  |  |  |  |  | 35.6 | 13.6 | 7.1 | 0.6 |
| 14 |  |  |  |  |  |  |  |  |  |  |  |  | 36.9 | 12.1 | 0.5 |
| 15 |  |  |  |  |  |  |  |  |  |  |  |  |  | 42.9 | 0.5 |
| Total | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| Share | 29.3 | 17.9 | 12.0 | 9.1 | 7.3 | 5.8 | 5.0 | 3.7 | 2.9 | 2.2 | 1.6 | 1.2 | 0.9 | 1.1 | 100.0 |

Note: The statistics are constructed as follows. We start from the whole sample of all firms and drop firm-product pairs that only exported once in their lifetime. For each firm-product pair, we calculate its total number of exporting years and the number of unique trade patterns in its lifetime and then put it into the relevant cells of the table. The last row "Share" indicates the share of firm-product pairs with the total number of exporting years equal to $x$. The last column gives the share of firm-product pairs with $y$ number of unique trade patterns.

for two or more years in their lifetime. For example, consider the firm-product pairs being observed for 5 years: 64.1% (100-35.9%) of them have at least one repeated trade pattern in their exporting life.

At the core of our study is the recognition that the time-varying patterns of market participation are informative about economically important but unobservable factors that drive exporters' trade strategies. By way of example, looking at Figure 1, a plausible hypothesis is that the time-varying unobservables (in demand and production costs) that drive a firm to sell to destinations A and C in periods 2 and 4 are very similar to each other; and that time variation in these unobservables may also drive the firm's choice of destinations A, B and C in period 3 and 5.

Intuitively, by constructing a fixed effect that controls for a destination market *when it appears as part of a larger trade pattern*, one can restrict the comparison of observations to circumstances in which the underlying time-varying unobservables take similar values.[12] Restricting the analysis of price and exchange rate variation by comparing observations for a destination conditional on the same (repeated) trade patterns allows us to construct a difference-in-difference estimator that offers a potentially stronger control in unbalanced panels, compared to alternative popular approaches, effectively limiting the variation of unobserved confounding factors. We provide a theory-based discussion of these points in the next section.

# 3 Pricing to Market by Heterogeneous Firms and Products: Model-based Analysis

In this section, we lay out the theoretical underpinnings of trade-patterns as a fixed effect that improves on the control of unobservable variables. First, we will use a well-known model to gain fundamental insight on how and why empirical analyses of pricing to market may suffer from bias arising from omitted variables and market selection, as well as other econometric issues raised by firm heterogeneity. Second, we will show that this model and the recent statistical literature lend theoretical support to the hypothesis that trade patterns are informative about crucial factors impinging on firms' behavior. Finally, we will use simulated data to gauge the performance of alternative fixed effects estimators in the presence of multiple sources of bias, and show that comparing results across estimators can provide an informative diagnostic.

We use a variant of the model by Corsetti and Dedola (2005) and Berman, Martin and Mayer (2012), specifying a multi-country framework that allows for heterogeneous firms, variable markups

---

[12]To be concrete, in our example this implies a set of fixed effects which interact each country with each of its observed trade patterns; this set could be captured by a series of dummies: one for destination A interacted with the trade pattern A-C that takes the value 1 in periods 2 and 4, but 0 in periods 1, 3, and 5; a second dummy for destination A interacted with the trade pattern A-B-C that takes the value of 1 in periods 3 and 5, and a third for destination A interacted with the (non-repeating) trade pattern A-B that is equal to 1 in period 1.

and endogenous entry, and solve the model in partial equilibrium. The model calls attention to vertical interactions between upstream producers and downstream distributors—we thus label it CD, for Cost and Distribution. As a result, markups are heterogeneous across firms depending on a firm's size and the incidence of distribution costs in the product market between distributors and producers, all of which impinge on the trade elasticity. We also allow for heterogeneity in the type(s) of products firms sell. The model will thus be useful to understand the multi-layered econometric problems arising from the fact that product-level marginal costs are not observed and firms' responses to shocks are likely to be heterogeneous.[13]

## 3.1 The Model Setup

The world consists of $H$ countries. In each country, there is a continuum of industries indexed by $i$. The final consumption $Y_{dt}$ and price $P_{dt}$ in each destination $d \in H$ are CES aggregated across the industries and defined as:

$$Y_{dt} = \left[ \int_i (Y_{idt})^{\frac{\eta-1}{\eta}} di \right]^{\frac{\eta}{\eta-1}}, \quad P_{dt} = \left[ \int_i (P_{idt})^{1-\eta} di \right]^{\frac{1}{1-\eta}} \tag{1}$$

where $\eta$ is the elasticity of substitution across industries. Within each industry, there are domestic and foreign firms producing different varieties with the substitutability of these varieties indicated by $\rho_i$. The industry-level consumption $Y_{idt}$ and price $P_{idt}$ are defined as:

$$Y_{idt} = \left[ \sum_o \sum_f \phi_{fiodt} \left( \alpha_{fiodt} \right)^{\frac{1}{\rho_i}} \left( Y_{fiodt} \right)^{\frac{\rho_i-1}{\rho_i}} \right]^{\frac{\rho_i}{\rho_i-1}},$$

$$P_{idt} = \left[ \sum_o \sum_f \phi_{fiodt} \alpha_{fiodt} (P_{fiodt})^{1-\rho_i} \right]^{\frac{1}{1-\rho_i}} \tag{2}$$

where $\alpha_{fiodt}$ is a preference shifter of the demand for firm $f$'s product $i$ from origin $o$ in destination $d$ at time $t$ and $\phi_{fiodt}$ indicates whether the firm $f$ from origin $o$ is active in industry $i$ and destination $d$ at time period $t$.

**Firms' Problem**. Markets are assumed to be segmented and each firm makes its pricing and entry decisions independently in each market.[14] A firm decides whether to sell to a market

---

[13]Progress in estimating markups for multi-product firms has been made in contributions including De Loecker and Warzynski (2012) and De Loecker, Goldberg, Khandelwal and Pavcnik (2016), but these approaches require strong assumptions on input ratios in production functions as well as detailed balance sheet data on firm-level inputs which, for many countries, is not readily available to researchers or is only available for some firms. See Appendix E for a formal discussion.

[14]The independent market decisions are implied by the assumption of a constant return to scale production function. As the marginal cost in one destination does not depend on that in another destination, the optimization

(indicated by $\phi_{fiodt}$) by calculating its expected operating profit in that market $\pi_{fiodt}$ and comparing it to the entry cost $E_{iod}$:

$$
\phi_{fiodt} = \begin{cases} 1 & \text{(observed)} & \text{if} & \pi_{fiodt} \geq E_{iod} \\ 0 & \text{(missing)} & \text{if} & \pi_{fiodt} < E_{iod} \end{cases} \tag{3}
$$

Upon entry into market $d$, the firm's operating profit is given by:

$$
\pi_{fiodt} = \max_{P_{fiodt}} Y_{fiodt} \left[ \mathcal{E}_{odt}(P_{fiodt} - \chi_i P_{dt}^N) - \tau_{odt} \mathcal{MC}_{fiodt} \right]
$$

subject to

$$
Y_{fiodt} = \alpha_{fiodt} \left( \frac{P_{fiodt}}{P_{idt}} \right)^{-\rho_i} \left( \frac{P_{idt}}{P_{dt}} \right)^{-\eta} Y_{dt} \tag{4}
$$

where $\mathcal{MC}_{fiodt}$ is the marginal cost of firm $f$ from industry $i$ and origin $o$ selling to destination $d$ at time $t$; $\tau_{odt}$ is the bilateral trade cost between origin $o$ and destination $d$ at time $t$; $\mathcal{E}_{odt}$ is the bilateral exchange rate defined as the units of currency $o$ per unit of currency $d$ at time $t$. We assume that $\chi_i$ units of destination nontradable goods are needed to distribute the product to local consumers, resulting in an additional "wedge", $\chi_i P_{dt}^N$, in the profit maximisation problem.[15]

At the retail level, the firm's optimal price expressed in the destination's currency is given by:

$$
P_{fiodt} = \frac{\rho_i}{\rho_i - 1} \left( \frac{\tau_{odt} \mathcal{MC}_{fiodt}}{\mathcal{E}_{odt}} + \chi_i P_{dt}^N \right). \tag{5}
$$

At the border, the firm's optimal price denominated in the producer's currency (i.e., origin country $o$'s currency) is given by:

$$
P_{fiodt}^B \equiv \mathcal{E}_{odt} \left( P_{fiodt} - \chi_i P_{dt}^N \right) = \underbrace{\frac{\rho_i}{\rho_i - 1} \left( \tau_{odt} + \frac{1}{\rho_i} \frac{\chi_i P_{dt}^N \mathcal{E}_{odt}}{\mathcal{MC}_{fiodt}} \right) \mathcal{MC}_{fiodt}}_{\text{Markups} \equiv \mu_{fiodt}} \tag{6}
$$

Define the retail cost ratio as the retail cost divided by the marginal cost multiplied by the trade cost expressed in the producer's currency as $\omega_{fiodt} \equiv \frac{\chi_i P_{dt}^N \mathcal{E}_{odt}}{\tau_{odt} \mathcal{MC}_{fiodt}}$. From the above expressions, the optimal markup adjustment is a function of changes in the exchange rate $\widehat{\mathcal{E}}_{odt}$, the marginal cost $\widehat{\mathcal{MC}}_{fiodt}$, the retail cost $\widehat{P}_{dt}^N$ and the trade cost $\widehat{\tau}_{odt}$:

$$
\widehat{\mu}_{fiodt} = \Gamma_{fiodt} \left( \widehat{\mathcal{E}}_{odt} - \widehat{\mathcal{MC}}_{fiodt} + \widehat{P}_{dt}^N \right) + (1 - \Gamma_{fiodt}) \widehat{\tau}_{odt} \tag{7}
$$

---

problem can be solved independently in each market.

[15] $P_{dt}^N$ is the price of non-tradable goods. We assume that the non-tradable goods market is monopolistically competitive and that firms selling non-tradable goods charge a constant markup.

where the markup elasticity to exchange rates is given by:

$$\Gamma_{fiodt} \equiv \frac{\omega_{fiodt}}{\rho_i + \omega_{fiodt}} = \frac{\chi_i P_{dt}^N \mathcal{E}_{odt}}{\rho_i \tau_{odt} \mathcal{MC}_{fiodt} + \chi_i P_{dt}^N \mathcal{E}_{odt}} \tag{8}$$

Equations (7) and (8) highlight the two key theoretical predictions of the model: (a) the markup elasticity to the exchange rate is *decreasing* in $\rho_i$, suggesting high differentiation goods tend to have *higher* markup adjustments relative to low differentiation goods; and (b) the markup elasticity is *increasing* in the retail cost ratio, suggesting that more productive firms—with lower marginal costs and a larger market share—tend to make *higher* markup adjustments.

## 3.2 From Theory to Data: Potential Sources of Estimation Bias

There are at least three empirical hurdles to estimating the markup elasticity to the exchange rate (8) using the observed border price defined in (6). First and foremost, product-level marginal cost is generally not observable and might reasonably be expected to vary over time. Relatedly, firms may endogenously change the set of destination markets they choose to serve, depending on the realizations of shocks to marginal costs, or other variables such as the bilateral exchange rate. This endogeneity is apparent from the market participation condition (3) and the change in the operating profits of the firm:[16]

$$\widehat{\pi}_{fiodt} = (1 - \rho_i) \left[ \frac{1}{1 + \omega_{fiodt}} \left( \widehat{\tau}_{odt} - \widehat{\mathcal{E}}_{odt} + \widehat{\mathcal{MC}}_{fiodt} \right) + \frac{\omega_{fiodt}}{1 + \omega_{fiodt}} \widehat{P}_{dt}^N \right] + \widehat{\mathcal{E}}_{odt}$$
$$+ \widehat{\alpha}_{fiodt} + (\rho_i - \eta)\widehat{P}_{idt} - \eta \widehat{P}_{dt} + \widehat{Y}_{dt} \tag{9}$$

When, conditional on current shocks, the drop in profits is sufficiently large, a firm may no longer find it optimal to export to a particular destination. Conversely, favourable shocks may raise prospective profits in a destination market enough to induce entry. Selection bias in estimating markup elasticities to exchange rates arises from variables that enter both the pricing (6) and the selection (i.e., (9) and (3)) equations. Thus, the bias arises from the unobserved marginal cost, the (destination-specific) distribution and trade costs, as well as the exchange rate.[17] This implies the

---

[16]The operating profit in levels is given by

$$\pi_{fiodt} = \frac{\alpha_{fiodt}}{\rho_i} \left( \frac{\rho_i}{\rho_i - 1} \right)^{1 - \rho_i} \left( \frac{\tau_{odt} \mathcal{MC}_{fiodt}}{\mathcal{E}_{odt}} + \chi_i P_{dt}^N \right)^{1 - \rho_i} (P_{idt})^{\rho_i - \eta} (P_{dt})^{-\eta} Y_{dt}$$

[17]Other variables in (9), such as changes in preferences $\widehat{\alpha}_{fiodt}$, the sectoral price index $\widehat{P}_{idt}$, the destination CPI $\widehat{P}_{dt}$, and the aggregate demand $\widehat{Y}_{dt}$ are important determinants of the firm's market choices, but they will not result in selection bias as the optimal price is independent from these variables in the CD model. Nonetheless, markup estimation may still suffer from omitted variable bias if any of these variables are correlated with the bilateral

markup elasticity estimate will thus be biased, even if the bilateral exchange rates are uncorrelated with the unobserved marginal cost.

A third hurdle to estimating markup elasticities is heterogeneity across firms and products. The literature addresses heterogeneity by estimating elasticities on subsamples of firms and products. For example, differences in market power are captured by grouping observations by firm size or by ownership; such as private, state-owned and foreign-invested enterprises.[18]

## 3.3 Using Trade (Selection) Patterns to Control for Omitted Variable Bias

To introduce our analysis of fixed effect estimators in the most transparent manner, we temporarily abstract from heterogeneity and simplify the firm's problem in the model—we will add all the layers of the model back in subsection 3.4. We focus on the joint pricing and market-participation decisions by *one firm producing a particular product.* For simplicity, we write the optimal price as a function of only two variables, the bilateral exchange rate and the unobserved marginal cost, positing that the same variables also drive the market participation (selection) equation. Expressing all variables in logs, let $p_{dt}$ denote the (log) border price in destination $d$ denominated in the producer currency; $e_{dt}$ the (log) bilateral exchange rate between the origin and the destination country and $m_{dt}$ the (log) marginal costs the firm incurs in producing the goods exported there. In our simplified framework the optimal price is:

$$p_{dt} = \beta_0 + \beta_1 e_{dt} + \beta_2 m_{dt} \tag{10}$$

where $\beta_1$ is the markup elasticity to exchange rates; while the selection equation is:

$$\phi_{dt} = \begin{cases} 1 \text{ (observed)} & \text{if} \quad \gamma_0 + \gamma_1 e_{dt} + \gamma_2 m_{dt} < 0 \\ 0 \text{ (missing)} & \text{if} \quad \gamma_0 + \gamma_1 e_{dt} + \gamma_2 m_{dt} \geq 0 \end{cases} \tag{11}$$

Since $m_{dt}$ is unobserved and potentially correlated with $e_{dt}$, a simple OLS regression will generally be biased:

$$E(p_{dt}|e_{dt}, \phi_{dt} = 1) = \beta_0 + \beta_1 e_{dt} + \beta_2 E(m_{dt}|e_{dt}, \phi_{dt} = 1). \tag{12}$$

Our analysis moves from a key observation stressed by Heckman (1979), that selection bias itself can be thought of as an omitted variable bias: it would not be a problem if the variables creating the bias could be controlled for explicitly or implicitly in the estimation equation, for

---

exchange rate.

[18]See Appendix D for an analysis of differences across estimators due to approximation biases and different weightings of the observations when the markup elasticities are heterogeneous across firms and products.

example, by using fixed effects.[19] Thus, in our model, if the marginal cost $m_{dt}$ only varied over time but not across destinations, then adding time fixed effects would be sufficient to eliminate both the omitted variable and the selection biases. Both destination and time fixed effects would be needed, instead, if the marginal cost variation could be broken down into an additive process of two factors each varying at a particular dimension, i.e., $m_{dt} = \psi_d + \epsilon_t$.

These very basic considerations convey a first, reassuring, message from the model: appropriately specified fixed effects can make headway in addressing these sources of bias. But of course the inclusion of destination and time fixed effects cannot be expected to eliminate the bias in general—in our simplified model, they would not work for instance when the variation in the marginal cost is driven by factors varying over both dimensions and following non-additive processes.[20] One reason why these (statistically well-established) points have not been discussed by the literature is that, in practice, estimation methods (especially those that employ iterative statistical procedures) do not facilitate the inspection of the economic underpinnings of the variation used by fixed effects to recover parameters. As shown below, a benefit from our approach is that it provides greater clarity into the economics of identification.[21]

### 3.3.1 Trade Patterns as Fixed Effects

Our main innovation is to bring forward trade patterns – the sets of destinations observed in each period, denoted $D_t$—as a new fixed effect that can be used to improve upon existing estimators in addressing the biases in estimates of the markup elasticity to exchange rates.[22] This innovation rests on recognizing that, in an unbalanced panel, trade patterns are *per se* an informative panel dimension that can be exploited econometrically. For transparency of exposition, we restrict our attention to a two ($d$ and $t$) dimensional panel—in the appendix C.2 we show that our analysis generalizes to more complex three or four dimensional panels.

In large unbalanced panels, parameters in models with fixed effects over multiple dimensions cannot be simply estimated by sequentially removing the mean of variables at each of the corresponding panel dimensions (e.g., in our two dimensional example, removing the mean across

---

[19]To wit: both the omitted variable and the selection biases could be simultaneously addressed if, trivially, $m_{dt}$ were observable. Note that $E(p_{dt}|e_{dt}, m_{dt}, \phi_{dt} = 1) = \beta_0 + \beta_1 e_{dt} + \beta_2 m_{dt}$.

[20]We cannot add destination-time fixed effects as this would absorb all the variation in the bilateral exchange rate $e_{dt}$, the variable of interest.

[21]Popular statistical programs that take an iterative approach to estimating the fixed effects include Guimaraes and Portugal (2011), Rios-Avila (2015), and Correia (2017). As stressed by Guimaraes and Portugal (2011), while these algorithms work very nicely in practice, there is no theoretical guidance on whether the estimates obtained by these approaches are consistent.

[22]In what follows, we use $D$ to denote the *set of destinations* and treat it as an additional, constructed, panel dimension. We use $D$ with subscripts to denote the realized trade patterns in a particular time period. For example, $D_t$ denotes the set of destinations at period $t$; $D_{fit}$ denotes the set of destinations for firm $f$ selling product $i$ at time $t$. For example in Figure 1, the trade pattern can be thought of as a new panel dimension $D$ in which $D_t$ takes the value A-C in periods $t = 2$ and $t = 4$.

destinations and then removing the mean across time periods). Their estimation requires the use of iterative procedures (see Guimaraes and Portugal (2011), Rios-Avila (2015), and Correia (2017)) or, equivalently, a statistical projection matrix.[23] While the complexity of either method makes it hard to inspect how the estimator achieves identification, in the following proposition we exploit the properties of projection matrices to show how fixed effects for trade patterns are inherently embedded in the implementation of estimators with $d + t$ fixed effects.[24]

**Proposition 1.** *In a two-dimensional unbalanced panel, factors varying at the $d + t$ panel dimensions can be eliminated using a two-step procedure by which, in the first step, all variables are demeaned across observed destinations within each period and, in the second step, destination $(d)$ and trade pattern $(D)$ fixed effects are applied additively, i.e., $d + D$.*

*Proof.* See Appendix C.1.1. □

To appreciate this proposition, note that, if the panel were balanced, $D$ would take the same value over time and the second step would collapse into a simple destination $d$ fixed effect. But given that the panel is unbalanced, the two step procedure in Proposition 1 highlights that trade patterns plays a crucial role in enabling the projection matrix to correctly estimate the parameters.

In light of proposition 1, the idea underlying our new estimator becomes quite intuitive. In the second step of the procedure, we combine the $d$ and $D$ fixed effects interactively instead of additively.

**Proposition 2.** *In a two-dimensional unbalanced panel, factors varying at the $dD + t$ dimensions can be eliminated in a two-step procedure in which all variables are demeaned across observed destinations within each period in the first stage and destination $(d)$ and trade pattern $(D)$ fixed effects are applied multiplicatively, i.e., $dD$, in the second stage. This procedure also eliminates all confounding factors that the $d + t$ fixed effects can address.*

*Proof.* See Appendix C.1.2. □

---

[23]See Wansbeek and Kapteyn (1989) for the projection matrix constructed for estimating two-way fixed effects.

[24]In this section, we discuss our estimator and the key propositions in the context of the firm's exporting problem. However, these results are more general and can be easily applied to other literatures. For example, the destination and time $(d+t)$ fixed effects discussed in Proposition 1 are analogue to the worker and (firm-)time fixed effects often applied in the labor literature (e.g., Abowd, Kramarz and Margolis (1999) and Card, Cardoso and Kline (2016)). A key assumption for the $(d+t)$ fixed effect approach to work is that the unobserved variables affecting the pricing and selection equations are exogenous after controlling for the two additive fixed effects (e.g., destination and time or worker and time). However, this assumption may be strong under some scenarios, e.g., in the context of our firms' exporting problem, when the unobserved time-varying marginal cost is also destination-specific, or, in the context of the labor literature, when workers systematically go to employers where they have a good match. Our Proposition 2 and subsection 3.3.2 show how a more general estimator can be constructed by explicitly exploiting the selection pattern of a particular panel dimension to relax this assumption.

To illustrate the meaning of this proposition, we find it useful to return to the stylized example in Figure 1. As noted earlier, in our example, two trade patterns (A-C and A-B-C) repeat twice, reflecting distinct time patterns of participation, in periods $\{2, 4\}$ and $\{3, 5\}$, respectively. The first step of our estimator removes the common time-varying components by differencing out the average value of a variable over all active destinations within a period. For example, for destination A, the first step calculates the price residual in period 2 as $\widetilde{p}_{A2} = p_{A2} - 1/2(p_{A2} + p_{C2})$ and in period 3 as $\widetilde{p}_{A3} = p_{A3} - 1/3(p_{A3} + p_{B3} + p_{C3})$. The second step performs a second demeaning, this time within each destination-trade pattern pair $(dD)$.[25] For destination A, again, we calculate $\ddot{p}_{A2} \equiv \widetilde{p}_{A2} - 1/2(\widetilde{p}_{A2} + \widetilde{p}_{A4})$ and $\ddot{p}_{A3} \equiv \widetilde{p}_{A3} - 1/2(\widetilde{p}_{A3} + \widetilde{p}_{A5})$, in periods 2 and 3, respectively.[26] We then run an OLS regression with the twice demeaned variables. So, in an unbalanced panel, the trade pattern fixed effect facilitates difference-in-difference identification by restricting the comparison of price residuals to sets of destinations that are part of an identical (repeated) trade pattern. Since the innovation in our approach is to use the observed trade patterns to construct an additional fixed effect to be applied in steps, we refer to our procedure as the Trade Pattern Sequential Fixed Effect Estimator (TPSFE).[27]

In light of this example, it should be clear that, relative to applying (conventional) additive $d + t$ fixed effects, combining the $d$ and $D$ panel dimensions has two main advantages. First, the novel fixed effect helps reduce the omitted variable and selection biases driven by factors that vary simultaneously at *both* the destination and time panel dimensions. Second, it enhances interpretability of identification in economic terms.

To the extent that the underlying time-varying unobservables take similar values across identical time patterns, the new fixed effect limits the variation of the unobserved factors that may confound the estimator.[28] In practice, the reduction of variability of unobserved confounding factors takes

---

[25]See table B1 in Appendix B.1.1 for an example of the construction of destination-trade pattern $(dD)$ fixed effects and see equations (C12)-(C15) in Appendix C.1.2 for formal definitions.

[26]In unbalanced panels, conventional fixed effects estimators do not achieve identification through a difference-in-difference comparison. Consider the $d+t$ fixed effects, implemented in the two-step estimator in Proposition 1. While the second step contains $D$ in additive form, i.e., $(d+D)$, these fixed effects can only be eliminated through iterative statistical procedures. These procedures prevent any clear economic interpretation of the underlying variation used for identification. As an alternative estimator, consider the possibility of simply ignoring the unbalanced structure of the data, and including only the destination fixed effects in the second step of Proposition 1. From each variable, the second step would remove a mean calculated *for a particular destination across all time periods*. For the price in destination A, for instance, it would calculate $\ddot{p}_{A2} \equiv \widetilde{p}_{A2} - 1/5(\widetilde{p}_{A1} + \widetilde{p}_{A2} + \widetilde{p}_{A3} + \widetilde{p}_{A4} + \widetilde{p}_{A5})$ and $\ddot{p}_{A3} \equiv \widetilde{p}_{A3} - 1/5(\widetilde{p}_{A1} + \widetilde{p}_{A2} + \widetilde{p}_{A3} + \widetilde{p}_{A4} + \widetilde{p}_{A5})$ in period 2 and 3, respectively. Such de-meaning is at least analytically interpretable. However, it is apparent that the demeaning using the $d$ fixed effect ignores the fact that averages of prices systematically vary across trade-patterns—the price residuals thus confound changes in prices with changes in means. Heuristically, they mix "apples" with "oranges".

[27]Statistically, as shown in Proposition 2, there is no difference in the parameters estimated through our TPSFE procedure and a statistical iterative procedure adding $dD + t$ fixed effects. An important reason to prefer our TPSFE procedure is that it has a clear difference-in-difference interpretation, and allows a transparent inspection of the variation used to identify the parameter of interest.

[28]In Appendix B, we provide a formal discussion on how our estimator can be thought of as a control function

---

17

place through two margins: (a) an extensive margin, where the observations associated with the non-repeated trade patterns are dropped and (b) an intensive margin where a different and more stringent fixed effect is applied to those observations associated with the repeated trade patterns. We discuss how these two channels work in the following subsection.

### 3.3.2 The Informative Content of Trade (Selection) Patterns

To show how our approach and estimator improves upon conventional fixed effect estimators, we consider the case in which the bilateral exchange rate is in the selection equation (11), and allow for interactive terms in the unobserved marginal cost as specified hereafter:

$$m_{dt} = \psi_d + \epsilon_t + \xi_d * \upsilon_t \tag{13}$$

Given the data generating process described by (10), (11) and (13), destination and time fixed effects can control for $\psi_d$ and $\epsilon_t$, but not for the last term $\xi_d * \upsilon_t$—therefore conventional fixed effect estimates will generally be biased. To the extent that the trade pattern reflects variation in $\upsilon_t$, adding destination-trade pattern fixed effects will help to reduce the bias caused by $\xi_d * \upsilon_t$.

This point is illustrated by the two graphs in Figure 2, drawn based on numerical simulations using our simplified model.[29] In these graphs, calendar time (in years) is on the $y$-axis; on the $x$-axis, there are possible values of the unobserved time-varying factors, either in the additive component $\epsilon_t$ (the graph to the left) or the destination-specific component $\upsilon_t$ (the graph to the right). For each year and realization of the unobserved factor, we plot the trade patterns chosen by the firm, distinguishing distinct patterns with different colors— the detailed set of destinations for each trade pattern is in the legend. We also distinguish between repeated and non-repeated trade patterns using, respectively, squares and circles—both shapes include the number of destinations reached by the firm in that year.

---

approach (e.g., Heckman (1979) and Kyriazidou (1997)). In subsection B.1.1, we discuss a numerical example on the indicative value of trade patterns and how it helps to reduce the selection bias.

[29]We simulate the data generating process described by (10), (11), and (13) for 10 destinations and 15 periods. We draw $\psi_d$ $\epsilon_t$, $\xi_d$ and $\upsilon_t$ from a standard normal distribution. To generate an exchange rate that is correlated with marginal cost, we set $e_{dt} = \sigma_e(m_{dt} + u_{dt})$ where $u_{dt}$ is drawn from a standard normal distribution. We set $\sigma_e$ to be 0.5 such that the bilateral exchange rate shocks are slightly less volatile than idiosyncratic marginal cost shocks. We set $\beta_1 = \beta_2 = 1$ such that an exchange rate appreciation of the home currency and a positive marginal cost shock increase the border price denominated in the home currency. This also implies a positive omitted variable bias. We set $\gamma_1 = -0.1$ and $\gamma_2 = 1$ such that the selection bias is also positive. The magnitude of $\gamma_1$ is set to be smaller than that of $\gamma_2$ to reflect the fact that the aggregate shocks (such as bilateral exchange rates) are less detrimental to the firm's entry decisions compared to idiosyncratic factors (such as the unobserved marginal cost). We set $\gamma_0$ such that around 70% observations (destination-year pairs) are dropped. In the cases where $\epsilon_t$ or $\upsilon_t$ is restricted to take a particular number of values (e.g., $x$), we rank all values of the variable ($\epsilon_t$ or $\upsilon_t$) generated, put them into $x$ bins and set the new value of the variable to be the average of the originally simulated values in each bin.

Figure 2: Example of the Indicative Value of Trade Patterns
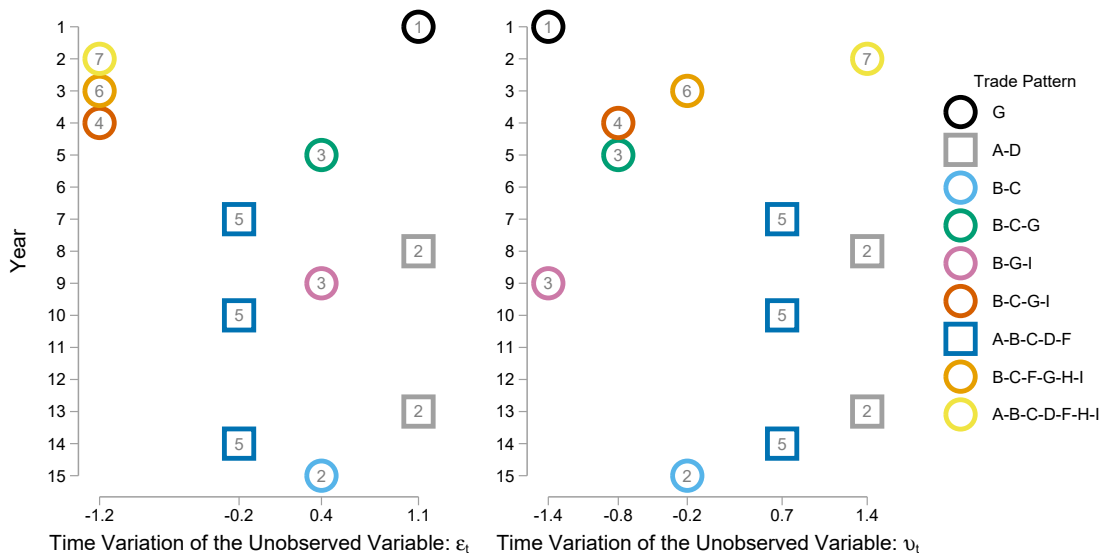
Note: These two figures plot the key time variation of the unobserved marginal cost $m_{dt}$ for a firm randomly generated according to the data generating process described in footnote 29. The left and right figures show the relationship between the formulated trade patterns and the values of $\epsilon_t$ and $\upsilon_t$ respectively. The trade pattern in each year is indicated by a different color with the legend on the right hand side showing the detailed set of destinations. The empty squares indicate the trade pattern is observed more than once in the sample. The empty circles indicate the non-repeated trade patterns. The number in the circle/square indicates the number of destinations associated with the trade pattern.

Trade patterns can vary over time in our simple model for three reasons: (a) changes in the unobserved time-varying factor $\epsilon_t$, (b) changes in the time-varying component of destination-specific shocks $v_t$ and (c) changes in the value of bilateral exchange rates conditional on the values of $\epsilon_t$ and $v_t$.[30] The graph to the left reports the trade patterns that the firm chooses over time for different realizations of the time-varying component of the firm's marginal costs $\epsilon_t$. As expected, a lower value of $\epsilon_t$ is associated with larger number of destinations—but note that the set of destinations is more informative than the number of destinations. Both the trade patterns A-D in year 8 and 13 and B-C in year 15 contain two destination markets—yet the underlying value of $\epsilon_t$ is very different. The graph to the right shows how the choice of a trade pattern changes with the underlying values of $v_t$, which drives the destination-specific time-varying component of marginal costs. Since the effect of $v_t$ on trade patterns depends of the realized value $\xi_d$, there is no monotonic relationship between the number of markets and the values of $v_t$. Nonetheless, trade patterns remain informative because, for a given realization of the destination factors $\{\xi_A, ..., \xi_d, ..., \xi_J\}$, an increase (decrease) in $v_t$ makes the firm more likely to enter and less likely to exit those markets with a negative (positive) $\xi_d$, forming distinct trade patterns.[31]

In table 3 we show the results from estimating $\beta_1$ from equation (10) on simulated data from the same model used for Figure 2—where the value of the relevant elasticity is set parametrically to be equal to 1 in all exercises. We employ the two fixed effect models, $d+t$ and $dD+t$, described in propositions 1 and 2, alongside OLS as a reference benchmark.[32] The two panels in the table, (a) and (b), differ in that the estimation sample in (b) is restricted to observations whose trade pattern repeats at least once.[33]

The table highlights three key properties of these estimators. First, comparing columns (3) and (4) over each row of the table, the inclusion of $dD$ fixed effects significantly reduces, without necessarily eliminating completely, the bias. In general, the underlying time-varying factors that drive selection of any specific market are unlikely to be perfectly correlated with the trade patterns.

---

[30]For part (c), you can think of the bilateral exchange rate as driven by two components: (1) a component that is related to the unobserved marginal cost and thus $\epsilon_t$ and $v_t$ and (2) a random component that is uncorrelated with the unobserved marginal cost.

[31]A positive $\xi_d$ suggests that an increase in $v_t$ leads to an increase in the marginal cost specific to destination $d$ and reduces the probability of the destination being observed. Similarly, a negative $\xi_d$ suggests that a rise in $v_t$ reduces the marginal cost and increases the probability of the destination being observed.

[32]The $dD + t$ fixed effects are estimated using our TPSFE estimator. The estimates are exactly the same if we use iterative statistical procedures such as the Stata `reghdfe` program to estimate $dD + t$ fixed effects.

[33]Our identification approach requires the same trade pattern to be repeated at least twice for the same firm. In our simplified example with only one firm, the probability of a trade pattern ever being repeated may be small. Therefore, we run 1000 simulations and compare the estimates only when the $FE(dD + t)$ specification can be estimated in the resulting sample. The share of simulations where we observe at least one repetitive trade pattern remains well above 90%. We should stress that the availability of the firm dimension in large datasets significantly improves identification; we find that allowing for 10 firms of a similar kind is in general enough to guarantee a sufficient number of repetitive trade patterns to apply our estimator.

Table 3: Comparison of Estimators when $\beta_1 = 1$

| Distinct Values of $\epsilon_t$ and $\upsilon_t$ | OLS | FE $(d+t)$ | TPSFE $(dD)$ | Distinct TPs | Reduction in Time Variation of $m_{dt}$ |
|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) |
| *Panel (a): Keeping non-repetitive trade patterns* | | | | | |
| 2 | 1.66 | 1.24 | 1.01 | 4.76 | 0.96 |
|   | (0.13) | (0.12) | (0.03) | | |
| 5 | 1.81 | 1.42 | 1.12 | 8.01 | 0.60 |
|   | (0.14) | (0.16) | (0.20) | | |
| 10 | 1.85 | 1.44 | 1.17 | 8.15 | 0.52 |
|   | (0.14) | (0.17) | (0.29) | | |
| 15 | 1.86 | 1.45 | 1.15 | 8.23 | 0.53 |
|   | (0.14) | (0.17) | (0.29) | | |
| *Panel (b): Dropping non-repetitive trade patterns* | | | | | |
| 2 | 1.66 | 1.21 | 1.01 | 3.47 | 0.94 |
|   | (0.15) | (0.12) | (0.03) | | |
| 5 | 1.73 | 1.23 | 1.12 | 2.96 | 0.45 |
|   | (0.23) | (0.22) | (0.20) | | |
| 10 | 1.74 | 1.23 | 1.17 | 2.81 | 0.33 |
|   | (0.25) | (0.30) | (0.29) | | |
| 15 | 1.72 | 1.21 | 1.15 | 2.68 | 0.29 |
|   | (0.26) | (0.30) | (0.29) | | |

Note: Estimates and standard errors are calculated from the average of 1000 simulations according to the data generating process described in footnote 29. The true value of the markup elasticity to exchange rates is set to 1. Panels (a) and (b) show the estimation results keeping and dropping non-repetitive trade patterns respectively. Column 1 shows the maximum number of distinct values that the time varying factor (i.e., $\epsilon_t$ or $\upsilon_t$) can take in the 15 periods of simulation. Column 2 gives OLS estimates obtained by regressing prices (in logs) on bilateral exchange rates (in logs). Column 3 shows the estimates after adding time and destination fixed effects. Column 4 shows the estimates after adding time and destination-trade pattern fixed effects. Column 5 shows the number of distinct trade patterns (averaged across all simulations). Column 6 shows the reduction in the time variation of the unobserved variable due to the trade pattern fixed effects.

The reason why the inclusion of $dD$ as a control variable improves estimation is that it *reduces* the variation of the unobserved variable in the time dimension (reflecting the correlation, if only weak, between the unobservable variable and the observed trade pattern). This is shown in column 6, where we include a measure of the reduction in the time variation of $m_{dt}$ due to the destination-specific trade pattern fixed effects—calculated as the time variation of $m_{dt}$ in a $d + t$ fixed effect specification minus the time variation of $m_{dt}$ in $dD + t$ fixed effect specification divided by the time variation of $m_{dt}$ in a $d + t$ fixed effect specification. This measure ranges between 50 and 96 per cent in panel (a).

Second, comparing estimates across different rows, the $TPSFE(dD)$ estimator performs better (i.e., the bias in column (4) is smaller) when the number of distinct values that the time-varying factors, $\epsilon_t$ and $\upsilon_t$, can take is fewer.[34] As the number of distinct values falls, the reduction in time variation of $m_{dt}$ coming from the inclusion of $dD$ fixed effects is larger (see column (6)). This means that the observed trade patterns can better distinguish the underlying values of $\epsilon_t$ and $\upsilon_t$ and therefore the $TPSFE(dD)$ approach yields estimates that are closer to the true theoretical value.

Third, comparing panels (a) and (b) suggests that a significant share of the bias in the $d + t$ fixed effects estimator in column (3) is due to (the confounding effects of) including observations with non-repetitive trade patterns in the sample.[35] This is a remarkable result, illustrating how the aforementioned extensive margin channel of our TPSFE estimator is at work. Subject to a sample selection caveat discussed in the next subsection, it suggests that dropping those trade patterns that only appear once in the sample (e.g., in Figure 2, these are the trade patterns G and A-D-F-G-H-J), effectively helps control for, by reducing, the potential values that $\upsilon_t$ can take. This purposeful reduction of the estimation sample by the removal of observations with singleton trade patterns improves the estimation bias even when conventional fixed effect estimators are used. Intuitively, the product prices observed in a singleton trade pattern may correspond to a very different realization of the marginal cost compared to its values in the repeated trade patterns. To illustrate this point, in the appendix B.1.1 we construct a numerical example and discuss in detail how conditioning on a repeated trade pattern pins down the range of variation that unobservable factors can take.

---

[34]It is worth noting that, given that we simulate the firm for 15 periods, the maximum number of distinct values that the time varying factors (i.e., $\epsilon_t$ and $\upsilon_t$) can take is 15. Therefore, the last row of each panel in Table 3 resembles the case where the time varying factors (i.e., $\epsilon_t$ and $\upsilon_t$) are drawn from a continuous distribution.

[35]Since we condition on the same set of observations, the difference in the estimates of columns (3) and (4) in panel (b) only reflects the distinctions in the "within variations" controlled for by the two sets of fixed effects (i.e., $FE(d + t)$ versus $TPSFE(dD)$). While the difference in the estimates of columns (3) and (4) of panel (b) are relatively small in our setting where the time span is quite short (15 periods), it can be very large in other settings where the time span with available information is sufficiently large (e.g., 100+ periods).

## 3.4 Omitted Variable Bias and Heterogeneous Markup Responses

We are now ready to evaluate the performance of fixed effect estimators, with and without the use of the trade pattern fixed effects, using simulated data from the model introduced at the beginning of this section. In contrast to the previous subsection, issues of selectivity and omitted variable bias will be compounded by problems raised by the heterogeneity in the markup elasticities of firms, for example, due to differences in the substitutability of the products they sell and/or their size.

We compare our estimators with a variety of estimators and approaches to identification employed in leading contributions to the pricing-to-market and exchange rate pass through literature.[36] In doing so, we call attention to key methodological differences in commonly implemented estimation strategies and highlight lessons that can be drawn from comparing results across estimators.

### 3.4.1 Model Simulations

We simulate our model with 30 *ex ante* symmetric countries. In each country, there are 200 industries. To introduce a realistic degree of heterogeneity across product markets, we model half of these industries selling goods with a low elasticity of substitution ($\rho_i = 3$) and the other half selling goods with a high elasticity ($\rho_i = 12$).[37] Within each industry, the productivity of firms is drawn from a Pareto distribution with the dispersion parameter equal to 2.5. We draw 10 firms for every industry in each country.[38] The firms make separate market decisions in all of the 30 markets depending on productivity, exchange rates, trade costs and other aggregate variables. We calibrate the distribution cost such that the mean distribution margin is about 50%.

We solve the model in partial equilibrium, where the aggregate variables such as the total national output, the CPI and the bilateral exchange rates are determined exogenously. In each period, the firms receive two correlated exogenous shocks, namely a marginal cost shock that can be destination-specific and an exchange rate shock.[39] Under the non-arbitrage condition of bilateral exchange rates, the values of the exchange rate shocks cannot be bilaterally independent. For a model of 30 countries, only 30 (rather than 30*30) shocks can be realized. We refer to these 30 shocks as base shocks and calculate the bilateral exchange rates based on the realization of the

---

[36]For example, Berman, Martin and Mayer (2012) and Chatterjee, Dix-Carneiro and Vichyanond (2013) apply $fid + t$ fixed effects; Fitzgerald, Haller and Yedid-Levi (2016) apply $fit + d$ fixed effects and Amiti, Itskhoki and Konings (2014) apply $fit + d$ fixed effects with time differenced variables.

[37]We set the cross-industry elasticity of substitution to 1.1.

[38]Given we have 30 countries, the maximum number of firms for an industry in a specific market is 300.

[39]We assume a linear production function where labour is the only input, i.e., $Y = AL$. We use the nominal wage in each country as the numeraire and set them to 1. Therefore, the productivity shock maps one-to-one to the marginal cost shock, i.e., $\mathcal{MC} = 1/A$.

base shocks:

$$\mathcal{E}_{ct}^{base} = \exp(\sigma_e \xi_c \upsilon_t + u_{ct})\mathcal{E}_{ct-1}^{base} \quad \forall c \in H \tag{14}$$

$$\mathcal{E}_{odt} \equiv \mathcal{E}_{ot}^{base}/\mathcal{E}_{dt}^{base} = \exp[\sigma_e(\xi_o - \xi_d)\upsilon_t + u_{ot} - u_{dt}]\mathcal{E}_{odt-1} \quad \forall o, d \in H \tag{15}$$

where $\xi_c, \upsilon_t$ and $u_{ct}$ are generated from a standard normal distribution and $c$ represents one of the 30 countries; $\sigma_e$ governs the size the systemic shock $\upsilon_t$. We set the marginal cost shocks to be partly correlated with bilateral exchange rate shocks:

$$\mathcal{MC}_{fiodt} = \exp[\sigma_{mc}\psi_{fiot} - \sigma_{e,mc}(\xi_o - \xi_d)\upsilon_t]\mathcal{MC}_{fiodt-1} \tag{16}$$

where $\psi_{fiot}$ is drawn from a standard normal distribution; $\sigma_{mc}$ governs the size of non-destination specific marginal cost shocks; $(\xi_o - \xi_d)$ captures the origin-destination-specific effect of the common shock $\upsilon_t$ and $\sigma_{e,mc}$ governs the correlation between the marginal cost and the bilateral exchange rates. Finally, to reflect the data structure of a typical customs database, we drop domestic sales of firms in each country and run the estimation equations separately for each exporting country.[40]

To keep the environment as tractable as possible, in the experiment below we abstract from imported inputs, as these will not affect the results as long as imported input costs are not destination specific.[41] We present a selection of simulation results from our extensive robustness checks, using a variety of alternative calibrations of the parameters and the shocks, in the Online Appendix.

### 3.4.2 What Can We Learn from Comparing Fixed Effect Estimators?

Based on the simulated data from the model, Table 4 shows the estimation results of a range of standard fixed effects estimators in the literature, together with our new estimator.[42] In panel (a), we simulate the model setting $\sigma_{e,mc} = 0$, so that marginal cost shocks are firm-product-time specific, whereas in the simulations underlying panel (b) shocks to marginal costs are firm-product-*destination*-time specific. In each panel, we show estimates of the markup elasticities for the whole sample, but we also highlight the role of heterogeneity by reporting results for sub-samples of goods with low and high within-industry elasticities of substitution – which, for comparison with

---

[40]In practice, we do not observe a firm that exports to all of the destinations in our sample. Therefore, in our simulations, we drop firms exporting to all of the other 29 countries.

[41]A destination-specific imported input shock is similar to a destination-specific productivity shock.

[42]We focus on the unobserved marginal cost changes as the main confounding factor of the markup elasticity estimation in all of our simulations. In principle, in the model, the omitted variable and selection biases can also arise from the unobserved destination-specific trade and distribution costs. Qualitatively, shocks to trade and distribution costs result in a similar bias compared to the shocks to marginal costs because the optimal border price and the firm's profit are affected in the same way by trade and distribution cost shocks as marginal cost shocks: a rise in the trade cost or the distribution costs increases the optimal border price (see equation (7)) and reduces the profitability of the firm (see equation (9)).

our empirical analysis, we dub high and low differentiation goods. We should stress that, while the exact estimates in this table are quite sensitive to our assumptions, the main takeaways we spell out below are robust to different simulation settings.[43]

As a reference benchmark, the first three columns of Table 4 show the estimates from running OLS. Column (1) shows a native OLS estimate regressing logged prices on logged bilateral exchange rates. In column (2) the regression model includes information that can be (at least in principle) obtained via the productivity and marginal cost estimation approach as in De Loecker, Goldberg, Khandelwal and Pavcnik (2016)—it adds the average of a firm's marginal cost of a product across destinations as an additional control variable. Column (3) shows the performance of a linear estimator relying on (hypothetical) accurate information on key unobserved variables—that is, we include the firm-product-destination-specific marginal cost and preference shifts of demand as additional controls. Column (4) refines (3) by adding product fixed effects—defining the best linear estimator. In the last column, (10) we include a measure often used in the literature, taking an unweighted average of the markup elasticities defined in expression (7).

Comparing estimates of the first three columns in panel (a), we find that adding the marginal cost as a control significantly reduces the bias in OLS estimation relative to benchmarks in columns (4) and (10), provided high-differentiation (HD) and low-differentiation (LD) goods are studied separately. In panel (a), the OLS specification in column (2) already precisely recovers the theoretical mean for both types of goods, whereas in panel (b), the same specification suffers from a small downward bias in the sample of HD goods. As it is clear from the richer model specification in column (3), the downward bias in HD goods would be corrected if one could observe the true destination-specific marginal costs. A notable and perhaps surprising finding from the "All" row in panel (b), is that the estimation bias can be high when the heterogeneity created by different type of products is ignored. As shown in column (3), this is the case even when the complete information of marginal costs and preference are available.

Columns 5-9 of Table 4 show the estimates from commonly applied fixed effects approaches in the literature and our proposed TPSFE approach. Column (5) includes the $d+t$ fixed effects applied by Knetter (1989).[44] Columns (6) and (7) add the firm-product fixed effects to the destination and time panel dimensions respectively. Relative to column (5), the estimator in column (6) also accounts for the differences in prices, marginal costs and preferences that are firm-product-destination specific, whereas the estimator in column (7) also absorbs firm-product-time specific

---

[43]See Online Appendix OA8 for robustness checks on alternative calibrations and shock structures.

[44]As the bilateral exchange rate only varies at the destination and time panel dimensions and is naturally independent from the unobserved variables varying at firm and product dimensions, the Knetter (1989)'s $FE(d+t)$ specification is sufficient to control for firm-product-time varying unobserved marginal costs and will give unbiased estimates if (a) the panel is balanced and (b) the markup elasticity is homogeneous in the estimation sample. However, the firm and product panel dimensions are relevant for identifying the markup elasticity if either of the two conditions fail, in which case the Knetter (1989) specification will be in general biased.

Table 4: Comparison of Estimators (based on simulated data from the model)

| Sample | OLS | OLS with $\overline{MC}_{fit}$ | OLS with $MC_{fidt}$ & $\alpha_{fidt}$ | OLS with $MC_{fidt}$ & $\alpha_{fidt}$ & $i-FE$ | FE $(d+t)$ | FE $(fid+t)$ | FE $(fit+d)$ | FE $(fit+fid)$ | TPSFE $(fidD)$ | Un-weighted Theo. Mean |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| **Panel (a): firm-product-time specific marginal cost shock** | | | | | | | | | | |
| All | 0.66 | 0.90 | 0.90 | 0.09 | 0.34 | 0.22 | 0.10 | 0.10 | 0.09 | 0.10 |
| | (0.03) | (0.03) | (0.03) | (0.00) | (0.05) | (0.01) | (0.00) | (0.00) | (0.00) | |
| HD Goods | 0.74 | 0.25 | 0.25 | 0.25 | 0.38 | 0.33 | 0.25 | 0.25 | 0.25 | 0.25 |
| | (0.02) | (0.00) | (0.00) | (0.00) | (0.03) | (0.01) | (0.00) | (0.00) | (0.00) | |
| LD Goods | 0.56 | 0.06 | 0.06 | 0.06 | 0.34 | 0.20 | 0.06 | 0.06 | 0.06 | 0.06 |
| | (0.02) | (0.00) | (0.00) | (0.00) | (0.03) | (0.01) | (0.00) | (0.00) | (0.00) | |
| **Panel (b): firm-product-*destination*-time specific marginal cost shock** | | | | | | | | | | |
| All | 0.65 | 0.89 | 0.88 | 0.09 | 0.30 | 0.19 | 0.05 | 0.06 | 0.08 | 0.10 |
| | (0.03) | (0.03) | (0.03) | (0.00) | (0.05) | (0.01) | (0.00) | (0.00) | (0.00) | |
| HD Goods | 0.70 | 0.22 | 0.25 | 0.25 | 0.23 | 0.20 | 0.09 | 0.13 | 0.20 | 0.25 |
| | (0.02) | (0.00) | (0.00) | (0.00) | (0.03) | (0.01) | (0.00) | (0.01) | (0.00) | |
| LD Goods | 0.56 | 0.06 | 0.06 | 0.06 | 0.33 | 0.18 | 0.04 | 0.04 | 0.06 | 0.06 |
| | (0.02) | (0.00) | (0.00) | (0.00) | (0.03) | (0.01) | (0.00) | (0.00) | (0.00) | |

Note: Estimates based on a randomly generated sample of 30 countries, 200 industries, maximum 300 firms per industry-country pair and 15 time periods. We calibrate $\rho_{HD} = 3$ and $\rho_{LD} = 12$. We set $\sigma_{mc} = 0.01$ and $\sigma_e = 0.01$. The entry cost of high differentiation goods is set to be 200 times higher compared to that of low differentiation goods to obtain a similar entry rate. We set the $\sigma_{e,mc}^{HD} = 0.5$ and $\sigma_{e,mc}^{LD} = 0.05$. We run the estimation equations separately for each of the 30 exporting countries and the table reports the average of the estimates.

26

shocks. In columns (8) and (9), we add firm-product fixed effects to the estimators discussed in subsection 3.3.[45] Column (8) reports estimates from the richest linear additive fixed effect model, where we add the firm-product fixed effects to both the destination and time fixed effects. Column (9) includes our proposed estimator with firm-product-destination-specific trade patterns.[46]

A first important takeaway from Table 4 is that the results of fixed effect estimators in column (8) or (9) are either close to (in panel (a)), or not too far from (in panel (b)) the results in column (4), obtained from rich OLS models relying on estimating marginal costs at either the firm-product-time or the firm-product-destination-time level. This suggests that fixed effect estimators do provide a good alternative to the estimation of markup elasticities when sufficiently accurate information on productivity and marginal cost is not available.

A second important takeaway from the table is that, from panel (a), the $(fit + fid)$ fixed effects and the $TPSFE(fidD)$ estimators produce similar results when marginal cost shocks are not destination-specific. As established in Proposition 2 and Corollary 2, in this case the coefficients estimated by the two models should be exactly the same when markup elasticities are homogeneous.

A final takeaway is that, from panel (b), the $TPSFE(fidD)$ estimator outperforms $(fit + fid)$ fixed effects (yielding estimates that are closer to the theoretical mean) if marginal cost shocks are destination-specific. A divergence across estimates can thus be informative about the structure of the underlying shocks. Altogether, these findings suggest that running the $TPSFE(fidD)$ estimator in conjunction with conventional fixed effects provides a useful diagnostic test for potential issues arising from firm-product-destination-time specific confounding factors.

# 4    Product Differentiation as a Proxy for Market Power: a New Classification

Market power is an important factor underpinning pricing to market. In our model, this varies not only with a firm size, with larger and more productive firms enjoying more market power, but also with the type of product: exporters of a product whose substitutability with other goods is limited, or can be limited by the inclusion of specific attributes, are better able to segment markets in order to exploit (local) market power. In this section we propose a new product classification that, relative to existing ones, better captures the degree of product differentiation.

Relative to the enormously popular classification of Rauch (1999), our new classification splits Rauch's large class of differentiated goods into two groups, high- and low-differentiation goods. The qualifying feature of the Corsetti-Crowley-Han-Song (CCHS) classification is that it exploits linguistics-based information uniquely available in Chinese customs data. This information allows

---

[45]We provide a formal discussion on the differences in Appendix C.2.

[46]We provide a formal description of our approach in Appendix A.

us to create a general, finely defined, and comprehensive system which is applicable internationally to all datasets that use the Harmonized System.

## 4.1    A Comprehensive Classification Based on Chinese Linguistics

The core principle underlying our classification is a simple one: traded goods which are discrete items are more differentiated than traded goods which are continuous. The main value added of our classification consists of the way it identifies discrete versus continuous goods. We rely on a feature of Chinese linguistics present in Chinese customs reporting – the use of indigenous Chinese measure words to record quantity for specific HS08 products. In the Chinese Customs Database, we find quantity reported in 36 different measures, many of which exist only in Chinese.[47] Linguists categorize Chinese measure words as count/discrete or mass/continuous classifiers; we operationalize this linguistic distinction to categorize each Harmonized System product as highly differentiated (i.e., for discrete goods) or less differentiated (i.e., for continuous goods).[48]

The key advantage to using Chinese linguistics to identify if a good is discrete or continuous arises from the facts that (a) all Chinese nouns have an associated measure word that inherently reflects the noun's *physical attributes* and (b) the Chinese Customs Authority mandates the reporting of quantity for Chinese HS08 products in these measure words. The first fact means that identifying discrete products from Chinese "count classifiers" is arguably more accurate and systematic than alternatives. Specifically, Chinese measure words are more distinctive and more precisely tied to specific nouns by Chinese grammar rules than the eleven units of measure recommended by the World Customs Organization (WCO) are linked to nouns in languages such as English or German.[49] Moreover, because the choice of the *measure word* used to record a product's quantity is predetermined by Chinese grammar and linguistics, we can set aside concerns that the choice of a quantity measure could be endogenous.[50]

---

[47]Notably, the linguistic structure of other East Asian languages also requires the use of measure words. In Appendix G.5 we explain how Japanese customs declarations integrate indigenous Japanese measure words into the World Customs Organization quantity measurement framework.

[48]See Cheng and Sybesma (1998, 1999) for a discussion of mass classifiers and count classifiers in Chinese. Cheng and Sybesma (1998) explain: "while massifiers [mass classifiers] *create* a measure for counting, count-classifiers simply *name the unit* in which the entity denoted by the noun it precedes naturally presents itself. This acknowledges the cognitive fact that some things in the world present themselves in such discrete units, while others don't. In languages like English, the cognitive mass-count distinction is grammatically encoded at the level of the noun..., in Chinese the distinction seems to be grammatically encoded at the level of the classifier" (emphasis added).

[49] See Fang, Jiquing and Connelly, Michael (2008), *The Cheng and Tsui Chinese Measure Word Dictionary*, Boston: Cheng and Tsui Publishers, Inc. for a mapping of Chinese nouns to their associated measure words. In Appendix G.5 we provide examples of how measure words are used in Chinese grammar.

[50]Since 2011, the WCO has recommended that *net weight* be reported for *all transactions* and supplementary units, such as number of items, be reported for 21.3% of Harmonized System products. However these recommendations are *non-binding*; the adoption and enforcement of this recommendation by a country might be endogenously determined by the value or volume of trade in a product, with high-value products subject to stricter enforcement that counts be reported. The sophistication of a country's border operations and tax authority could also play a

To illustrate the variety of measures used in the Chinese Customs Dataset, table 5 reports a selection of the most commonly used measure words, the types of goods that use the measure word, and the percent of export value that is associated with products described by each measure word. In this table, qiān kè (千克) and mǐ, (米) are mass/continuous classifiers; the remaining measure words are count/discrete classifiers. The main point to be drawn from the table is that the nature of the Chinese language means that the reporting of differentiated goods, for example, automobiles, spark plugs and engines, takes place by reporting a number of items and the count classifier that is linguistically-associated with that type of good. All products within an HS08 code use the same measure word. See appendix G.5 for an example of the different Chinese measures words used to quantify closely-related products in our dataset.

Table 5: Measure word use in Chinese customs data for exports, 2008

| Quantity Measure | Meaning | Types of goods | Percent of export value |
|---|---|---|---|
| qiān kè, 千克 | kilogram | grains, chemicals | 40.5 |
| tái, 台 | machines | engines, pumps, fans | 24.7 |
| gè, 个 | small items | golf balls, batteries, spark plugs | 12.8 |
| jiàn, 件 | articles of clothing | shirts, jackets | 6.6 |
| shuāng, 双 | paired sets | shoes, gloves, snow-skis | 2.6 |
| tiáo, 条 | tube-like, long items | rubber tyres, trousers | 2.5 |
| mǐ, 米 | meters | camera film, fabric | 2.1 |
| tào, 套 | sets | suits of clothes, sets of knives | 1.8 |
| liàng, 辆 | wheeled vehicles | cars, tractors, bicycles | 1.4 |
| sōu, 艘 | boats | tankers, cruise ships, sail-boats | 1.3 |
| kuài, 块 | chunky items | multi-layer circuit boards | 0.7 |

The second fact, that quantity must be reported on Chinese Customs forms in indigenous count units for discrete objects, means that the Chinese Custom system will likely be quite accurate in accounting for discrete items, relative to what can be inferred from the quantity measures actually reported in other customs systems. By way of example, in Egyptian customs records over 2005-2016, a mere 0.006% of export observations report the discrete unit "pieces" as the unit of quantity. In comparison, the share of Chinese export data that uses a count/discrete measure for reporting quantity is 40.9% of observation-weighted HS08 data and 52.8% of value-weighted HS08 data (see the last rows of panels (a) and (b) in table 6).[51]

---

role in which measures are reported. See United Nations Statistics Division (2010).

[51] Authors' calculations from EID-Exports-2005-2016 obtained from http://erfdataportal.com. Egypt is a useful comparator in that it had a similar per capital income to China during the midpoint of our sample, 2007, $1667 (Egypt) versus $2693 (China), and it used a similarly large variety of quantity measures, 32, in its export statistics over 2005-2016. See Appendix G.5.2 for a discussion of quantity reporting in other customs systems.

Table 6: Classification of goods: Integrating the insights from CCHS with Rauch

(a) Share of goods by classification: observation weighted

| | Corsetti-Crowley-Han-Song (CCHS) | | |
| | Low Differentiation / (Mass nouns) | High Differentiation / (Count nouns) | |
|---|---|---|---|
| **Rauch (Liberal Version)** | | | |
| Differentiated Products | 41.1 | 38.8 | 79.8 |
| Reference Priced | 6.9 | 0.7 | 7.6 |
| Organized Exchange | 0.6 | 0.0 | 0.6 |
| Unclassified[†] | 10.5 | 1.5 | 12.0 |
| | 59.1 | 40.9 | 100.0 |

(b) Share of goods by classification: value weighted

| | Corsetti-Crowley-Han-Song (CCHS) | | |
| | Low Differentiation / (Mass nouns) | High Differentiation / (Count nouns) | |
|---|---|---|---|
| **Rauch (Liberal Version)** | | | |
| Differentiated Products | 24.2 | 47.1 | 71.3 |
| Reference Priced | 9.1 | 2.8 | 11.9 |
| Organized Exchange | 2.0 | 0.0 | 2.0 |
| Unclassified[†] | 11.9 | 2.9 | 14.8 |
| | 47.2 | 52.8 | 100.0 |

Notes: Share measures are calculated based on Chinese exports to all countries including Hong Kong and the United States during periods 2000-2014. [†] "Unclassified" refers to HS08 products that do not uniquely map to differentiated, referenced priced, or organized exchange under the SITC Rev. 2-based classification of Rauch.

## 4.2 Improvements Relative to the Rauch (1999) Industry Classification

Table 6 demonstrates the value added of our classification system in relation to the leading industry classification set forth by Rauch (1999). According to Rauch (1999), a product is differentiated if it does not trade on organized exchanges and/or its price is not regularly published in industry sales catalogues. While this principle is quite powerful in identifying commodities, a drawback is that the vast majority of manufactured goods end up being classified as differentiated.

In Table 6, we integrate our classification of high versus low differentiation goods with that obtained by mapping HS08 product codes from the Chinese Customs Data to Rauch's original 4 digit SITC Rev. 2 classification of differentiated, reference priced, and organized exchange traded goods. There are two crucial improvements from our approach. First, our classification refines the class of differentiated goods in Rauch into two categories—high and low differentiation. From table 6 panel (a), we observe that 79.8 percent of observations in the Chinese Customs Database at the firm-HS08 product level are classified by Rauch as differentiated. Of these, only 48.6 percent (38.8/79.8) use count classifiers and are categorized as high differentiation under the CCHS approach. The picture is similar in panel (b), where observations are value weighted: of the 71.3 percent of the export value classified by Rauch as differentiated, 66.1 percent (47.1/71.3) use count classifiers. Further, table 6 confirms that every good that Rauch categorizes as a commodity (i.e., an organized-exchange traded good) is reported in the Chinese Customs Database with a mass classifier. This conforms with our prior that mass nouns are low differentiation goods and serves as a useful reality check on our approach.

The second improvement from our classification system is that we are able to provide a CCHS classification for *all* HS08 (and HS06) products, including those that cannot be classified under Rauch's system due to issues with the mapping from HS06 to SITC Rev. 2. This enables us to expand our analysis of market power to include the 12% percent of observations (table 6 panel (a)) and 14.8% of export value (table 6 panel (b)) in the Chinese Customs Database in HS08 products that do not uniquely map to a single Rauch category.[52] In concluding this section, we would like to emphasize that the CCHS linguistics-based product classification can be applied to the universal 6-digit Harmonized System used by all countries by categorizing as high (low) differentiation those

---

[52]To be clear, Rauch provides a classification for each SITC Rev. 2 industry as differentiated, reference priced or organized exchange, but the SITC Rev. 2 industries in his classification are more aggregated than HS06 products. Because the concordance of disaggregated HS06 product codes to (more aggregated) SITC Rev. 2 involves one-to-many or many-to-many mappings for 81 percent of concordance lines, we are only able to classify HS06 products (and even finer HS08 products) into one of the three Rauch groupings if *all* SITC Rev. 2 industries associated with an HS06 product are "differentiated," etc. under Rauch. This one-to-many and many-to-many concordance issue implies that no unique mapping into Rauch's three categories is possible for 12% of observations in the Chinese Customs Database.

HS06 categories in which all HS08 products use a count/discrete (mass countinuous) classifier.[53]

# 5 Empirical Results

In this section, we present our empirical estimates of pricing to market. To make our results comparable with leading studies in the literature on exchange rate pass through, we apply all estimators *conditional on a price change* in line with the methodology of Gopinath, Itskhoki and Rigobon (2010).[54] Our sample period includes an important change in the exchange rate regime pursued by China. In the years 2000-2005, China pursued a fixed exchange rate regime; after that, it switched to a managed float regime. We will show evidence that exporters' pricing behavior differs across the corresponding subsample periods. Throughout our analysis, to ensure comparability of our estimates across policy regimes, we exclude exports to the US and Hong Kong, and treat eurozone countries as a single economic entity, integrating their trade flows into a single economic region.[55]

In the next subsection, we present the markup elasticities from our TPSFE estimator and CCHS product classification. We then proceed to a thorough comparison against standard fixed effect estimators used by leading papers in the literature. The final subsection examines and discusses the role of quantity adjustments in relation to pricing to market.

---

[53]See Appendix G.5.4 for examples of closely-related HS08 products and the types of measure words they use. In most cases, all HS08 products within an HS06 category use the same measure word.

[54]Specifically, we estimate all parameters *after* applying a data filter to the Chinese export data: for each product-firm-destination combination, we filter out absolute *price changes in dollars* smaller than 5 percent. To be clear, while we condition on price changes in dollars, we regress unit values denominated in renminbi on the bilateral renminbi/local currency exchange rate. We provide an example on how the price change filter is constructed and how trade patterns are subsequently formulated based on the price-change-filtered database in appendix G.9. The main conclusion of our analysis remains roughly similar if we apply our estimator without conditioning on price changes (see Online Appendix OA1) as well as if we filter out absolute price changes in renminbi smaller than 5 percent (see Online Appendix OA2).

[55]Qualitatively, results do not change if we include exports to the United States and Hong Kong. We aggregate the export quantity and value at the firm-product-year level for 17 eurozone countries including Austria, Belgium, Cyprus, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Malta, Netherlands, Portugal, Slovakia, Slovenia and Spain. Latvia and Lithuania joined the eurozone in 2014 and 2015, respectively. We treat them as separate countries throughout our analysis. Our results are robust to the inclusion and exclusion of small countries that adopted the euro in the later period of our sample. We performed two robustness checks. One excludes Slovenia, Cyprus, Malta, Slovakia and Estonia from the eurozone group and treats them as separate individual countries, resulting in an estimation sample of 157 destinations. Another excludes Slovenia, Cyprus, Malta, Slovakia and Estonia from the eurozone group and drops these five countries from our estimation sample, resulting in an estimation sample of 152 destinations. These two alternative estimation samples yield results very similar to our primary estimation sample (152 destinations) which integrates the 17 eurozone countries together.

## 5.1 Markup elasticities By Product Differentiation

We start our analysis by applying our TPSFE estimator to estimate markup elasticities to exchange rates. Specifically, we run the following OLS regression on the twice demeaned variables:[56]

$$\ddot{p}_{fidt} = \beta_0 + \beta_1 \ddot{e}_{fidt} + \ddot{\boldsymbol{x}}'_{fidt}\boldsymbol{\beta}_2 + \ddot{v}_{fidt} \tag{17}$$

where $p$ denotes the log export price denominated in the producer's currency (i.e., in RMB), $e$ is the log bilateral nominal exchange rate defined as units of RMB per destination currency and $\boldsymbol{x}$ includes destination-specific control variables including the CPI, real GDP and the import-to-GDP ratio of the destination market. For each variable $y \in \{p, e, x\}$, $\ddot{y}_{fidt}$ denotes twice demeaned variables, with the first demean taken at the firm-product-year level and the second demean taken at the firm-product-destination-trade pattern level.[57] Finally, $\beta_1$ is the markup elasticity to bilateral exchange rates.

Two points are worth stressing upfront. First, the estimated markup elasticity would be zero if exporters set the same dollar (or RMB) price for their product in all destinations—irrespective of whether these prices are sticky or move across time, and whether they are set in RMB or dollars. Second, our estimation procedure is robust to the choice of bilateral exchange rates. For example, we get the same estimates from using either the dollar-destination currency or the RMB-destination currency exchange rate as the independent variable. This is because the RMB-dollar exchange rate movement is common across destinations and thus differenced out from our procedure.

### 5.1.1 Baseline Results

In table 7, we show results across exchange rate regimes and according to the degree of product differentiation. On average, we estimate an average markup elasticity to exchange rates of 5% during the dollar peg period (2000-2005) and of 11% during the later managed floating period (2006-2014). The finding that the markup elasticity is rising over time indicates that exporters from China engaged more extensively in price discrimination in the later period, after China abandoned its strict peg to the US dollar.[58]

In both periods, nonetheless, our econometric model detects significant differences in markup

---

[56]We discuss the detail procedures in Appendix A. We adjust the degrees of freedom when calculating the standard errors of the estimates taking into account that dependent and independent variables have been demeaned twice (see e.g.,Wansbeek and Kapteyn (1989), p. 346). Practically, we use the algorithm written by Abowd, Creecy and Kramarz (2002) and adapted in `reghdfe` to calculate the correct degree of freedom of the demeaned data.

[57]See Appendix A for more details of variable construction.

[58]Plausibly, the differences in markup elasticities we detect across the two time periods reflect more than just the policy reform of switching from a dollar peg to a managed float in China. They may stem from structural changes at the firm and market level, as well as from changes in the frequency and importance of cyclical (policy and technology) shocks at the national and global level that have occurred between the two time periods.

Table 7: Markup elasticities to the exchange rate

|              | All       | HD Goods  | LD Goods  | n. of obs    |
|--------------|-----------|-----------|-----------|--------------|
| $2000-2005$  | 0.05**    | 0.10***   | 0.01      | 4,279,808    |
|              | (0.02)    | (0.03)    | (0.03)    | [1,073,300]  |
| $2006-2014$  | 0.11***   | 0.20***   | 0.06***   | 19,272,657   |
|              | (0.01)    | (0.02)    | (0.01)    | [4,839,333]  |

Note: Estimates based on specification (17) and the sample of multi-destination trade flows at the firm-product-time level to 152 destinations excluding Hong Kong and the United States. The bilateral exchange rate is defined as RMBs per unit of destination currency; an increase means an appreciation of the destination currency. Robust standard errors are reported in parentheses. The actual number of observations used for identification is reported in the brackets of the last column. Statistical significance at the 1, 5 and 10 percent level is indicated by ***, **, and *.

elasticities across high and low differentiation goods—validating the usefulness of our linguistics-inspired product classification as a proxy for market power. Starting with the first row, the low average estimate of the markup elasticity during the dollar-peg period conceals important differences across types of goods. For CCHS high differentiation exports, the markup elasticity is 10%, while for low differentiation goods it is zero. In the period of the managed float of the renminbi (second row), markup elasticities are considerably higher. For high differentiation goods, the markup elasticity rises from 10 to 20%. For low differentiation goods, the markup elasticity is smaller but becomes significantly positive, at 6%. For these low differentiation goods, pricing-to-market appears to play only a moderate role after the strict peg is abandoned. It is important to keep in mind that, all else equal, a larger markup adjustment measured in producer's currency implies a smaller change in import prices *measured in the currency of the destination market.* This means that firms exporting more highly differentiated goods keep their prices in local currency more stable against bilateral currency movements relative to firms exporting low differentiation goods.

The last column of table 7 (and all tables in this section) reports the size of the whole estimation sample (in the same row as the parameter estimates), and the size of the sample that provides identification to the TPSFE estimator (in square brackets [·] in the same row as the standard errors). This smaller identification sample excludes observations corresponding to non-repetitive trade patterns. Because the TPSFE procedure yields identical parameter estimates when applied to either sample,[59] it is important to verify that the (sometimes considerably) smaller identification subsample is representative of the whole estimation sample. For this table and all remaining

---

[59]This occurs because, for non-repetitive trade patterns, the demeaning procedure creates entries of zeros (for both dependent and independent variables) for those observations associated with singleton trade patterns. These entries of zeros do not affect the point estimates of an OLS regression but may generate incorrect standard errors if one fails to correct the true degrees of freedom. Fixed effect estimators typically correct the degrees of freedom when estimating the standard errors (see e.g.,Wansbeek and Kapteyn (1989), p. 346). Thus, the standard errors we report are based on the size of the *identification sample* rather than the full estimation sample.

tables, we check that the identification sample is representative. We fail to detect any noticeable differences in samples, with the exception of the sample partitioned by firm size.[60]

### 5.1.2 Combining the CCHS Classification with Firm and Product Characteristics

**CCHS with Firm Ownership.** The Chinese economy is widely understood to be a hybrid in which competitive, market-oriented private firms operate alongside large, state-owned enterprises (SOEs).[61] Looking at exports, the picture is actually more complex. Quantitatively, export activity is dominated by firms that are wholly foreign owned or are Sino-foreign joint enterprises—the leading types in a group that we label foreign-invested enterprises (FIEs).[62]

A firm's ownership type likely reflects a host of differences including cost structures, available technologies, and the types of products made. First, SOEs and FIEs are believed to have relatively easy access to capital, but are likely to differ in the extent to which they rely on imported intermediates in production. Conversely, private firms are widely seen as facing tighter financing constraints and, relative to FIEs, a lower level of integration with global supply chains. Second, the average size of a firm also differs across these groups; private enterprises are smaller on average, which likely reflects a high rate of entry by young firms. Third, being more integrated in supply chains, FIEs may engage in transfer pricing. In light of these considerations, we might expect SOEs, FIEs and private firms to endogenously end up producing different products, using different production processes, and possibly targeting different markets. This prompts us to ask whether a firm's registration type contributes to explaining observable differences in markup elasticities.

Evidence on markup elasticities by firm type is presented in table 8, where we focus on the period 2006-2014. Private enterprises stand out for their extremely low markup elasticity of 4% (column 1, row 3). This suggests that these firms follow a pricing strategy that is indistinguishable from setting a single dollar price for their output across destinations. The picture is totally different for state-owned and foreign-invested enterprises. For these firms, more than 20% of any bilateral exchange rate change is absorbed into markup changes—22% for SOEs and 23% for FIEs. This is evidence that both FIEs and SOEs hold a high degree of market power which enables them to exploit market segmentation and strategically price-to-market. Although these results may in part capture transfer pricing motivated by profit shfiting practices, at a broad level, the pricing strategies of SOEs and FIEs appear to be totally different from those of private enterprises. A large divide across Chinese firms is apparent in our results.

---

[60]We report a wide range of checks on the distributions of variables in the full estimation sample versus the identification sample in Online Appendix OA5. We only detect a small sample difference for medium and large firms in table 9 below.

[61]See Hsieh and Song (2015) and Wu (2016) for analyses of the inter-relations of firms and the state in the Chinese economy and Hale and Long (2012) on the importance of inward FDI into China.

[62]Over 2000-2014, about one-half of Chinese export value originated from FIEs. See appendix G.2 for details.

Table 8: Markup elasticities by firm registration types $(2006 - 2014)$

| Category | All | HD Goods | LD Goods | n. of obs |
|---|---|---|---|---|
| State-owned Enterprises | 0.21*** | 0.41*** | 0.09** | 3,526,943 |
|  | (0.03) | (0.05) | (0.04) | [646,352] |
| Foreign Invested Enterprises | 0.24*** | 0.33*** | 0.20*** | 4,990,504 |
|  | (0.03) | (0.04) | (0.04) | [1,042,481] |
| Private Enterprises | 0.04*** | 0.10*** | 0.01 | 9,897,091 |
|  | (0.01) | (0.02) | (0.02) | [2,996,133] |

Note: Estimates based on specification (17) and the sample of multi-destination trade flows at the firm-product-time level to 152 destinations excluding Hong Kong and the United States. The bilateral exchange rate is defined as RMBs per unit of destination currency; an increase means an appreciation of the destination currency. Robust standard errors are reported in parentheses. The actual number of observations used for identification is reported in the brackets of the last column. Statistical significance at the 1, 5 and 10 percent level is indicated by ***, **, and *.

Yet, product differentiation plays a non-secondary role in explaining differences across firm types. The estimated markup elasticity for highly differentiated products sold by SOEs is as high as 41% while the markup elasticity for high differentiation goods sold by FIEs is 33%. This suggests that these firms are better able to segment markets and more actively pursue local currency price stability than other groups of firms, although is is not incomprehensible that these large markup adjustments are driven by tax considerations, especially for FIEs. Remarkably, even when SOEs and FIEs export less differentiated goods, estimated elasticities are sizeable, 9% for SOEs and 20% for FIEs. Product differentiation is also relevant for private firms with a sizeable elasticity of 10% for high differentiation goods.

**CCHS with Firm Size.** Our results from table 8 show that market power is best captured by a combination of product *and* firm type. We now consider a measure of firm size; a firm's product-level global export revenue.[63] For a given firm-product-year triplet, we calculate the firm's global export revenue, summed over all active destinations in that year. We then rank firms within products and years by product-level export revenue, and place them into three equally-sized bins, labelled small, medium and large.[64]

Exporters' markup elasticities to the bilateral exchange rate increase systematically with their product-level export revenues (table 9 column 1). Regardless of the degree of product differentiation, large exporters appear to command more market power and adjust their markups in

---

[63]This definition of size differs from that in papers such as Berman, Martin and Mayer (2012) and Amiti, Itskhoki and Konings (2014) which measure firm size as total domestic and foreign revenues for *all* products. The categorization we employ emphasizes that a firm's market power could vary across distinct products.

[64]Our definition of firm-size categories is at the product-year level. That is, all the firms selling the same product in the year are placed in bins containing the same number of observations. When the number of firms cannot be divided by three, we place more firms in the lower ranked bins. For example, say we have 5 firms selling to 2 destinations each. We put two firms in the "Small" bin, two firms in the "Medium" bin and one firm in the "Large" bin. This is why, in table 9, the number of observations in the "Small" and "Medium" categories is slightly higher compared to the "Large" category.

Table 9: Pricing-to-market by exporters' product-level global revenues $(2006 - 2014)$

| Category | All | HD Goods | LD Goods | n. of obs |
|----------|-----|----------|----------|-----------|
| Small Exporters | 0.05*** | 0.11*** | 0.02 | 6,639,830 |
| | (0.02) | (0.03) | (0.02) | [2,646,437] |
| Medium Exporters | 0.11*** | 0.24*** | 0.05** | 6,519,743 |
| | (0.02) | (0.04) | (0.03) | [1,448,368] |
| Large Exporters | 0.25*** | 0.37*** | 0.20*** | 6,113,084 |
| | (0.03) | (0.05) | (0.04) | [744,528] |

Note: Estimates based on specification (17) and the sample of multi-destination trade flows at the firm-product-time level to 152 destinations excluding Hong Kong and the United States. The bilateral exchange rate is defined as RMBs per unit of destination currency; an increase means an appreciation of the destination currency. Robust standard errors are reported in parentheses. The actual number of observations used for identification is reported in the brackets of the last column. Statistical significance at the 1, 5 and 10 percent level is indicated by ***, **, and *.

response to bilateral exchange rate movements by 25% on average. In contrast, small exporters adjust markups by a mere 5%, suggesting that their pricing strategies are close to setting a single global price across all destinations.

Further segmenting the sample according to the degree of product differentiation reveals striking heterogeneity in pricing. In response to bilateral exchange rate movements, large firms adjust markups substantially, 37% when exporting highly differentiated products. These firms appear to command a high level of market power even when they sell low differentiation products, with an estimated elasticity of 20%. Thus, significant price stability in local currency terms can be partially understood as a reflection of the fact that large firms let their markups (measured in exporter's currency) absorb between one fifth and one third of a bilateral exchange rate movement between the origin and destination. But even small exporters can wield market power; if they export high differentiations goods, their markup elasticity of 10% is far from negligible.

**CCHS with UN end-use categories.** Firms selling directly to consumers typically engage in branding and advertising campaigns to a much larger extent than firms selling intermediate products. Insofar as producers of consumption goods are successful in making their products less substitutable with other products or product varieties, markets for consumption goods should be less competitive than markets for intermediates. Thus, we may expect destination specific markup elasticities to be higher for consumption goods than for intermediates.

In table 10, we partition our data into four categories by integrating our CCHS classification with the classification of consumption goods and intermediates under the UN's Broad Economic Categories (BEC).[65] We find a clear difference in the markup elasticities for consumption versus

---

[65]The UN's BEC classifies all internationally traded goods according to their end-use. The most disaggregated classification available in BEC Rev. 4 maps HS06 products into end-use categories of consumption goods, intermediate inputs, and capital equipment. For our analysis, all HS08 products into the Chinese Customs Database are

Table 10: Markup Elasticities by BEC Classification $(2006 - 2014)$

| Category | All | HD Goods | LD Goods | n. of obs |
|---|---|---|---|---|
| Consumption | 0.21*** | 0.33*** | 0.09*** | 6,133,394 |
| | (0.02) | (0.02) | (0.02) | [1,759,243] |
| Intermediate | 0.05*** | 0.12* | 0.04** | 6,288,252 |
| | (0.02) | (0.07) | (0.02) | [1,579,220] |

Note: Estimates based on specification (17) and the sample of multi-destination trade flows at the firm-product-time level to 152 destinations excluding Hong Kong and the United States. The bilateral exchange rate is defined as RMBs per unit of destination currency; an increase means an appreciation of the destination currency. Robust standard errors are reported in parentheses. The actual number of observations used for identification is reported in the brackets of the last column. Statistical significance at the 1, 5 and 10 percent level is indicated by ***, **, and *.

intermediate goods; the elasticities of exporters selling consumption goods (0.21) are roughly 4 times larger than those of exporters of intermediates (0.05). When we further refine consumption goods into our CCHS product categories, the elasticity of high-differentiation consumption goods becomes strikingly large (0.33).

## 5.2   Comparing Fixed Effect Estimators

In line with our theoretical analysis, we now turn to a comparative analysis of different fixed effect estimators. In the first four columns of table 11, we report the estimates from our benchmark TPSFE estimator (column 1) alongside estimates from three alternative fixed effects that are commonly used in the literature:[66] specification (2) controls for firm-product-destination and time varying factors; specification (3) controls for firm-product-time and destination varying factors, such as the average marginal cost of the firm's product in the year; and specification (4), the most stringent fixed effects among the three, controls for both firm-product-destination and firm-product-time varying factors.[67]  In the last two columns of the table, we empirically verify Proposition 1 and Corollary 1, providing some empirical insight into the relevance of trade pattern fixed effects.[68]  In column (5) we implement the $FE(fit + fid)$ in two steps: in the first step, we demean the variables across destinations at the firm-product-year-and, inherently in an unbalanced

---

assigned the end-use of their corresponding HS06 code.

[66]Each estimating equation in table 11 includes four variables to proxy for changes in local market conditions, namely bilateral exchange rates, local CPI, real GDP in the destination, and the import-to-GDP ratio. We report the coefficient in front of bilateral exchanges rates in tables 11 and 12 and the coefficients in front of other control variables in Online Appendix OA4.

[67]These estimates are obtained iteratively using the Stata *reghdfe* program written by Correia (2017). As we discussed in subsection 3.3, while it is clear which unobserved variables can be controlled for by these fixed effects, the variation used to identify the parameter is, generally, not clear. In most cases, the iterative demeaning procedures and the statistical projection matrix of these approaches do not have meaningful economic interpretations.

[68]Corollary 1 restates the results of Proposition 1 in a panel which includes additional firm and product dimensions. See Appendix C.2.

panel, the trade pattern level (the first step of the TPSFE procedure); then, in a second step, we apply $fid + fiD$ fixed effects iteratively using the *reghdfe* program by Correia (2017). In column (6) we repeat the same two-step procedure as in column (5), but in the second step we remove the firm-product-trade pattern ($fiD$) fixed effects.

A first important takeaway from table 11 is that the $FE(fit + fid)$ estimator in column (4) and the two-step implementation of $FE(fit + fid)$ in column (5) yield exactly the same estimates, validating our Corollary 1 in our four dimensional empirical panel. Note that, intuitively, removing the trade pattern fixed effects in the second step of (5) is tantamount to ignoring selection entirely. When we do so (i.e., when we remove the trade pattern from the second step of the estimator used in column (5)) to perform the exercise in column (6), the results are strikingly different from those in all other columns.[69] The point estimates in column (6) are centered around values between 0.01 and 0.02 in all samples, failing to to detect any differences in elasticities over time and across types of goods. We take this as evidence of the importance of controlling for trade patterns.

Turning to the first four columns of table 11, our results highlight a remarkable difference in the relative performance of estimators across high- and low-differentiation goods. For low-differentiation goods, the estimated markup elasticities are similar across all specifications. For high-differentiation goods, instead, the differences in estimates across columns are sizeable. The elasticity estimated by the $TSPFE$ estimator is twice as large as that of the $FE(fit + fid)$ estimators in the 2000-2005 sample; 35% higher in the 2006-2014 sample.

A primary reason for observing a difference between estimators is that there might be unobserved variables which vary along dimensions that are not controlled for by a particular fixed effect specification. For example, the $FE(fid + t)$ estimator does not control for the average marginal cost of a firm's product in a year; the $FE(fid + fit)$ estimator cannot control for marginal cost shocks that are destination-specific. These observations, corroborated by our model-based comparative analysis of estimators in our theoretical section, offer a key for interpreting the results in the table.

For low differentiation goods, the lack of significant differences in the estimators' performances suggest that the frequent market changes observed in the data may be driven by unobserved variables that either are not directly related to exchange rates, or do not directly affect firms' markup decisions.[70] If the firm-product-time varying marginal cost for these products were highly correlated with exchange rates, we would expect to see a sizeable difference between the $FE(fit + fid)$ estimates from column (4) and the $FE(fid + t)$ estimates from columns (2). Similarly, if the marginal cost shocks were destination-specific, we would find a large difference between the

---

[69]After removing the trade pattern $D$ fixed effects, the $fiD$ fixed effects collapse to $fi$, which is a subset of $fid$ fixed effects.

[70]Recall that from section 3, we know if a variable, such as a preference shock, only enters the selection equation but not the pricing equation, then there is no selection bias.

$FE(fit + fid)$ and the $TPSFE$ estimates.

In contrast, for high differentiation goods, our analysis in subsection 3.4 suggests that the $TSPFE$ and the $FE(fit + fid)$ estimators would give very similar results if the underlying shock were not destination-specific. The significant difference in estimates points to a potentially non-trivial incidence of firm-product-destination-time specific shocks.[71] We will see below that a breakdown of the analysis distinguishing firms by size and ownership lends further support to this interpretation.

In comparing estimators, attention should be paid to an issue that may arise because, to facilitate identification, a stringent fixed effect approach may ignore variation from some observations in the analysis—this could introduce a sample selection bias. Specifically, our TPSFE requires the same firm-product-destination-trade pattern to appear at least twice, in order to exploit the time variation of price residuals.[72] If the TPSFE estimator somehow ends up utilizing variation from a sample with a higher proportion of large firms than is in the entire sample, the estimated markup elasticity may be higher than those obtained from the other (less stringent) FE estimators. This observation underscores the need to check systematically for differences in the distribution of key variables in the subsamples used for identification by $TPSFE$ and $FE(fid + fit)$ estimators. Encouragingly, when we do so in our data, in all but one case we fail to find significant differences, which leads us to conclude that there is little evidence that the TPSFE estimator obtains identification form a non-representative subsample in our study.[73]

Table 12 shows a breakdown of the sample by firm and product characteristics, focusing on the 2006-2014 sample. The upper and lower panels show estimates for high- and low-differentiation goods, respectively. In each panel, the first three rows give a breakdown by the global export sales of the exporter; the next three rows give a breakdown by the registration type of the firm. The last two rows show a breakdown of products from an integration of the CCHS and BEC classifications.

Consider first the upper panel of table 12 for high differentiation goods. Comparing the $FE(fit + fid)$ and the $TPSFE$ estimators across size and type of firms shows that most of the difference detected in table 11 is driven by FIEs and SOEs. These are typically large firms, integrated in complex supply chains, that likely employ complicated production processes relying on a large share of destination-specific components; hence they may more exposed to destination-

---

[71]We should note that we calibrated the simulated model in panel (b) of table 4 so that we would obtain estimates of the same magnitude and ranking as the ones estimated in the data. Conditional on the model, this allows us to have an idea of the relative incidence of the shocks we use in the model calibration.

[72]As discussed in subsection 3.3.2, removing the singleton firm-product-destination-trade pattern is a key step to control for unobserved confounding variables moving at the firm-product-destination-time dimensions.

[73]See Online Appendix OA5. We should mention here another model-based result potentially relevant to the comparison. If the markup elasticities are heterogeneous across firms, products, destinations and time, each estimator may weight the markup elasticity slightly differently due to the different statistical procedures applied to identify the parameters of interest. See Appendix D for a thorough discussion. However, our model simulations suggest that the difference in estimation driven by weighting tends to be relatively small.

Table 11: Comparison across Estimators

| | (1) | (2) | (3) | (4) | (5) | (6) | |
|---|---|---|---|---|---|---|---|
| | | | | | | (5) in Two Steps without Trade Pattern $D$ in the second step | |
| | TPSFE | FE $(fid+t)$ | FE $(fit+d)$ | FE $(fit+fid)$ | FE $(fit+fid)$ in Two Steps | | n. of obs |
| Destination Demean Fixed Effects | Yes $fidD$ | No $fid+t$ | No $fit+d$ | No $fid+fit$ | Yes $fid+fiD$ | Yes $fid$ | |
| **2000-2014** | | | | | | | |
| All Products | 0.10*** | 0.09*** | 0.06*** | 0.08*** | 0.08*** | 0.02*** | 23,552,465 |
| | (0.01) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | |
| High Differentiation | 0.18*** | 0.14*** | 0.09*** | 0.12*** | 0.12*** | 0.01*** | 10,185,981 |
| | (0.02) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | |
| Low Differentiation | 0.06*** | 0.05*** | 0.04*** | 0.05*** | 0.05*** | 0.02*** | 13,366,484 |
| | (0.01) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | |
| **2000-2005** | | | | | | | |
| All Products | 0.05** | 0.02*** | 0.04*** | 0.02*** | 0.02*** | 0.02*** | 4,279,808 |
| | (0.02) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| High Differentiation | 0.10*** | 0.04*** | 0.07*** | 0.05*** | 0.05*** | 0.02*** | 1,922,739 |
| | (0.03) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| Low Differentiation | 0.01 | 0.00 | 0.01** | 0.00 | 0.00 | 0.02*** | 2,357,069 |
| | (0.03) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| **2006-2014** | | | | | | | |
| All Products | 0.11*** | 0.11*** | 0.08*** | 0.09*** | 0.09*** | 0.02*** | 19,272,657 |
| | (0.01) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | |
| High Differentiation | 0.20*** | 0.16*** | 0.11*** | 0.13*** | 0.13*** | 0.01*** | 8,263,242 |
| | (0.02) | (0.01) | (0.00) | (0.01) | (0.01) | (0.00) | |
| Low Differentiation | 0.06*** | 0.07*** | 0.06*** | 0.07*** | 0.07*** | 0.02*** | 11,009,415 |
| | (0.01) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | |

 Note: Each cell reports the estimated markup elasticity from the estimation method specified on top of each column. Columns (5) and (6) empirically verify Proposition 1 and Corollary 1. Column (5) implements the $FE(fit+fid)$ in two steps: in the first step, we demean the variables across destinations at the firm-product-year-and, inherently in an unbalanced panel, the trade pattern level (the first step of the TPSFE procedure); then, in a second step, we apply $fid+fiD$ fixed effects iteratively using the *reghdfe* program by Correia (2017). Column (6) repeats the same two-step procedure as in column (5), but in the second step we remove the firm-product-trade pattern ($fiD$) fixed effects. Each row indicates a different subsample. Within in a row, all methods are applied based on the same sample. The number of observations in the last column corresponds to Stage 7 of the data cleaning procedure specified in appendix G.9. Robust standard errors are reported in parentheses. Statistical significance at the 1, 5 and 10 percent level is indicated by ***, **, and *.

Table 12: Comparison across Estimators: Further Breakdown

| | (1) TPSFE | (2) FE ($fid+t$) | (3) FE ($fit+d$) | (4) FE ($fit+fid$) | (5) FE ($fit+fid$) in Two Steps | (6) (5) in Two Steps without $D$ in the second step | n. of obs |
|---|---|---|---|---|---|---|---|
| Destination Demean | Yes | No | No | No | Yes | Yes | |
| Fixed Effects | $fidD$ | $fid+t$ | $fit+d$ | $fid+fit$ | $fid+fiD$ | $fid$ | |
| **2006-2014, High Differentiation** | | | | | | | |
| SOEs | 0.41*** | 0.25*** | 0.20*** | 0.22*** | 0.22*** | 0.01*** | 1,617,483 |
| | (0.05) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| FIEs | 0.33*** | 0.18*** | 0.13*** | 0.15*** | 0.15*** | 0.02*** | 2,267,880 |
| | (0.04) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| PEs | 0.10*** | 0.09*** | 0.06*** | 0.07*** | 0.07*** | 0.01*** | 3,988,833 |
| | (0.02) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| Small Exporters | 0.09*** | 0.10*** | 0.08*** | 0.09*** | 0.09*** | 0.00*** | 2,839,781 |
| | (0.03) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| Medium Exporters | 0.26*** | 0.17*** | 0.11*** | 0.13*** | 0.13*** | 0.01*** | 2,795,992 |
| | (0.04) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| Large Exporters | 0.47*** | 0.20*** | 0.14*** | 0.17*** | 0.17*** | 0.03*** | 2,632,132 |
| | (0.06) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| Intermediate | 0.12* | 0.13*** | 0.06*** | 0.08*** | 0.08*** | 0.02*** | 580,037 |
| | (0.07) | (0.03) | (0.02) | (0.02) | (0.02) | (0.00) | |
| Consumption | 0.33*** | 0.26*** | 0.16*** | 0.22*** | 0.22*** | 0.02*** | 3,581,291 |
| | (0.02) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| **2006-2014, Low Differentiation** | | | | | | | |
| SOEs | 0.09** | 0.04*** | 0.05*** | 0.06*** | 0.06*** | 0.02*** | 1,909,460 |
| | (0.04) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| FIEs | 0.20*** | 0.17*** | 0.13*** | 0.15*** | 0.15*** | 0.01*** | 2,722,624 |
| | (0.04) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| PEs | 0.01 | 0.02*** | 0.03*** | 0.03*** | 0.03*** | 0.02*** | 5,908,258 |
| | (0.02) | (0.01) | (0.00) | (0.01) | (0.01) | (0.00) | |
| Small Exporters | 0.03 | 0.02** | 0.03*** | 0.04*** | 0.04*** | 0.02*** | 3,807,296 |
| | (0.02) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| Medium Exporters | 0.09*** | 0.04*** | 0.04*** | 0.05*** | 0.05*** | 0.02*** | 3,723,759 |
| | (0.03) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| Large Exporters | 0.27*** | 0.14*** | 0.09*** | 0.12*** | 0.12*** | 0.03*** | 3,480,952 |
| | (0.04) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |
| Intermediate | 0.04** | 0.08*** | 0.06*** | 0.06*** | 0.06*** | 0.02*** | 5,712,115 |
| | (0.02) | (0.01) | (0.00) | (0.01) | (0.01) | (0.00) | |
| Consumption | 0.09*** | 0.09*** | 0.05*** | 0.08*** | 0.08*** | 0.02*** | 2,553,583 |
| | (0.02) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | |

Note: See the footnote of table 11 and the description in the text for an explanation of the columns.

specific marginal costs shocks.

Similar considerations may help rationalize the large differences in estimates observed in the last two rows of the upper panel; the differences across estimators are more pronounced for consumption goods, i.e., the product which might contain a destination-specific component.

Looking at the bottom panel of table 12, which reports estimates for low-differentiation goods, we find that the differences between the $FE(fit + fid)$ and the $TPSFE$ estimates are generally much smaller, in line with table 11.[74] The notable exceptions are, again, for products supplied by FIEs and large firms.[75]

## 5.3 Quantities and Markups

Thus far, we have presented evidence that some groups of firms exporting from China, particularly larger firms selling highly differentiated goods, discriminate across countries when changing their prices in response to bilateral exchange rate changes. Consistent with theory, we may expect them to systematically charge higher markups where, relative to other destinations, market (i.e., demand) conditions are more favorable. In this section we show how to use our framework to shed light on this point.

Our key observation is that, from the vantage point of a firm, holding production costs constant, changes in the exchange rates act as demand shifters. Thus, to the extent that our TPSFE estimator controls for cost-side factors, the predicted values from a projection of prices on exchange rates using (17) can be interpreted as changes in *relative markups* in response to changes in *relative demand across destinations driven by currency movements*. Under this interpretation, an increase in the relative markup charged in a market, raising the revenue per sale accruing to the firm, should be systematically associated with an increase in the relative quantity sold in that market.

In table 13, we report the results from running the following regression:

$$\ddot{q}_{fidt} = \gamma_0 + \gamma_1 \widehat{\ddot{p}}_{fidt} + \ddot{\boldsymbol{x}}'_{fidt} \boldsymbol{\gamma}_2 + \ddot{u}_{fidt}, \tag{18}$$

---

[74]A theoretical result from section 3 may help to refine the lessons from comparing estimators. The theoretical model is calibrated with a negative correlation between the exchange rate and marginal cost which generates a downward omitted variable bias; this omitted variable bias results in a positive difference between the estimates from the $TPSFE$ and $FE(fit+fid)$ estimators. In light of this model-based result, the positive difference between the empirical estimates from the $TPSFE$ and $FE(fit+fid)$ estimators suggests the presence of an omitted variable bias that the $TPSFE$ can reduce, but not necessarily entirely eliminate. Hence, the true markup elasticities might be even higher than estimates suggest.

[75]As a caveat, we should note that for large firms exporting low-differentiation goods, the estimate of the markup elasticity may in part reflect the fact that the mean firm size in the sample used by the TPSFE for identification is larger than the mean firm size in the estimation sample. This is the only case in which we are able to detect evidence that may suggest a bias from obtaining identification from a non-representative subsample with the TPSFE estimator.

where $\ddot{q}_{fidt}$ is the residual obtained by twice demeaning quantities using our *TPSFE* procedure, applying firm-product-time and firm-product-destination-trade pattern fixed effects, and $\widehat{\ddot{p}}_{fidt}$ are the predicted values of relative markups from (17).[76] We will refer to the coefficient of interest, $\gamma_1$, as the Quantity/Markup Elasticity (QME).[77] Table 13 also includes the results from a useful comparison model (labelled $Cor(\ddot{q}, \ddot{p})$). This is a regression of relative quantities directly on relative prices in which both variables have been demeaned twice as in the *TPSFE* procedure; that is, in this model the relative prices are *not projected on the exchange rate*.[78]

Table 13: Quantity-Markup Elasticity QME by CCHS Classification

| | All | | High Differentiation | | Low Differentiation | | |
|---|---|---|---|---|---|---|---|
| | (1) $Cor(\ddot{q}, \ddot{p})$ | (2) QME | (3) $Cor(\ddot{q}, \ddot{p})$ | (4) QME | (5) $Cor(\ddot{q}, \ddot{p})$ | (6) QME | n. of obs |
| $2000-2005$ | -0.71*** | 6.18*† | -0.75*** | 4.07** | -0.68*** | 19.72† | 4,279,808 |
| | (0.01) | (3.18) | (0.01) | (1.72) | (0.01) | (55.14) | [1,073,300] |
| $2006-2014$ | -0.70*** | 1.53*** | -0.72*** | 0.72*** | -0.69*** | 2.72*** | 19,272,657 |
| | (0.00) | (0.28) | (0.00) | (0.20) | (0.00) | (0.80) | [4,839,333] |

Note: Estimates based on the sample of multi-destination trade flows at the firm-product-time level to 152 destinations excluding Hong Kong and the United States. The "$Cor(\ddot{q}, \ddot{p})$" column is estimated using specification (A7) in the Appendix. The QME column is estimated based on equations (A5) and (18). Robust standard errors are reported in parentheses. Statistical significance at the 1, 5 and 10 percent level is indicated by ***, **, and *. † indicates that the t-statistic of the bilateral exchange rate in the first stage is smaller than 2.58.

To appreciate the contribution of our QME, note that, in columns (1), (3) and (5), the sign of the regression coefficient from the model including relative prices is consistently negative. For example, in column (1), a 1% increase in relative prices is statistically associated with a 0.7% *decline* in relative quantities. When twice demeaned prices are not projected on bilateral exchange rate movements, the coefficient seems to simply reflect that firms export relatively less in markets where they set prices relatively high.

In contrast, our estimates of QMEs all have positive signs, consistent with our idea that projections of prices on the exchange rate can be used to obtain a statistical measure of how relative quantities move with shifts in the demand for a firm and product driven by currency movements.

In the managed float regime in the 2006-2014 period (table 13, row 3), our estimated QME is positive and equal to 1.53 (row 3, column (2)): a one percent increase in the relative markup (driven by the exchange rate) is associated with 1.53 percent change in the relative quantity

---

[76]Because $\widehat{\ddot{p}}_{fidt}$ is an estimate, we correct the standard errors as in a standard two-step least squares approach.

[77]As is well known, holding the relative supply curve fixed, a shift in relative demand induces a movement along the relative supply curve. Heuristically, $\gamma_1$ is related to the slope of the relative supply curve; it is a statistical measure of the extent to which firms change their export supply in relation to markup adjustments to destination-specific demand changes. In the appendix, we explain how our approach relates to the quantity elasticities analyzed in the literature, e.g., by Berman, Martin and Mayer (2012).

[78]Appendix A describes in more detail the demeaning of variables for the *TPSFE* procedure.

across destinations. Table 13 further documents sharp differences in estimates across high and low differentiation goods. Over the same 2006-2014 period, the QME estimate is very low for high-differentiation goods, 0.72 (row 3, column 4): a one percent increase in the markup charged in a market is associated with a mere 0.72% increase in the export quantities supplied to that market. The estimated QME for low-differentiation goods is instead quite high, 2.72%. Recall that high- and low-differentiation goods feature, respectively, a high and a low markup elasticity; there is more pricing to market in high-differentiation exports. Our evidence thus lends empirical support to the view that firms with market power, such as those exporting high-differentiation products, respond to destination-specific exchange rate movements by adjusting markups substantially while keeping the relative quantity supplied across destinations relatively stable.

Our results pick up an interesting evolution of Chinese exporters over time and across exchange rate regimes. We have seen above that Chinese exporters' engagement in pricing-to-market was modest during the years of the fixed exchange rate regime (with the notable exception of exporters of high differentiation goods). Correspondingly, the QME estimates shown in the table for the period of the fixed exchange rate regime are quite high, ranging from 4.07 to 19.72 for high- and low-differentiation goods. Altogether, these results may suggest that, during the strict peg period, those firms that responded to bilateral exchange rate movements with modest markup adjustments were pursuing more aggressively any openings for expanding their market shares abroad.

The pattern highlighted in table 13, that goods and firms for which we estimate a higher relative markup adjustment tend to display a lower QME, is confirmed by table 14. From this table, once again, the divide between private firms, on the one hand, and FIEs and SOEs, on the other, is apparent. For private firms, a one percent increase in the relative markup in a market is associated with a 5.23 percent increase in the relative quantity sold in that destination (2.59 for exporters of high differentiation goods, 10.57 for exporters of low-differentiation goods). This is evidence that, on average, private Chinese firms keep their relative markups in check in response to currency movements; they price-to-market less and let relative export quantities move with demand conditions (possibly to gain market share). Relative to private firms, the opposite pattern emerges for SOEs and FIEs. Corresponding to their much higher markup elasticities, the estimated QMEs are very small and not significantly different from zero (0.34 for SOEs and 0.28 for FIEs).

The results in the table underscore the extent and importance of international market segmentation and market power. At one extreme we have SOEs, FIEs and exporters of highly differentiated consumption goods: the low estimate of quantity substitution across destinations (statistically indistinguishable from zero) suggests that the markets served by these firms and exporters of these goods are highly segmented. At the other extreme, for exporters of low-differentiation intermediates, quantity substitution is quite high (3.84) and markets appear quite integrated.

Table 14: Quantity-Markup Elasticity QME by Product and Firm Types $(2006 - 2014)$

| Category | All | | High Differentiation | | Low Differentiation | | |
|---|---|---|---|---|---|---|---|
| | (1) $Cor(\ddot{q},\ddot{p})$ | (2) QME | (3) $Cor(\ddot{q},\ddot{p})$ | (4) QME | (5) $Cor(\ddot{q},\ddot{p})$ | (6) QME | n. of obs |
| State-owned Enterprises | -0.70*** | 0.46 | -0.67*** | 0.11 | -0.71*** | 1.26† | 3,526,943 |
| | (0.01) | (0.31) | (0.01) | (0.23) | (0.01) | (1.19) | [646,352] |
| Foreign Invested Enterprises | -0.70*** | 0.19 | -0.70*** | 0.31 | -0.70*** | -0.11 | 4,990,504 |
| | (0.00) | (0.21) | (0.01) | (0.24) | (0.01) | (0.32) | [1,042,481] |
| Private Enterprises | -0.70*** | 5.23*** | -0.75*** | 1.99*** | -0.67*** | 16.87† | 9,897,091 |
| | (0.00) | (1.88) | (0.00) | (0.72) | (0.00) | (19.58) | [2,996,133] |
| Small Exporters | -0.65*** | 3.48** | -0.69*** | 1.85** | -0.63*** | 9.93† | 6,639,830 |
| | (0.00) | (1.56) | (0.01) | (0.81) | (0.00) | (14.35) | [2,646,437] |
| Medium Exporters | -0.72*** | 1.58*** | -0.74*** | 0.51 | -0.71*** | 3.53† | 6,519,743 |
| | (0.00) | (0.60) | (0.01) | (0.35) | (0.01) | (2.33) | [1,448,368] |
| Large Exporters | -0.77*** | 0.44* | -0.77*** | 0.06 | -0.77*** | 0.77* | 6,113,084 |
| | (0.01) | (0.26) | (0.01) | (0.26) | (0.01) | (0.44) | [744,528] |
| Consumption | -0.71*** | 0.47** | -0.77*** | 0.16 | -0.63*** | 1.68** | 6,133,394 |
| | (0.00) | (0.19) | (0.00) | (0.16) | (0.01) | (0.79) | [1,759,243] |
| Intermediate | -0.71*** | 3.34** | -0.73*** | 1.04† | -0.71*** | 3.84*† | 6,288,252 |
| | (0.00) | (1.55) | (0.01) | (1.39) | (0.00) | (1.98) | [1,579,220] |

Note: Estimates based on the sample of multi-destination trade flows at the firm-product-time level to 152 destinations excluding Hong Kong and the United States. The "$Cor(\ddot{q},\ddot{p})$" column is estimated using specification (A7) in the Appendix. The QME column is estimated based on equations (A5) and (18). Robust standard errors are reported in parentheses. Statistical significance at the 1, 5 and 10 percent level is indicated by ***, **, and *. † indicates that the t-statistic of the bilateral exchange rate in the first stage is smaller than 2.58.

## 5.4 Discussion

We have seen that the trade pattern fixed effect helps improve estimates of pricing to market by obtaining identification from observations which are associated with repeated trade patterns. We conclude our empirical analysis with an exercise, in which we run conventional estimators on the subsample of observations that identify the TPSFE estimator (see Appendix G.7), i.e., a dataset that excludes singleton trade patterns. In our dataset, the results from these estimators become close and in some case almost indistinguishable from TPSFE estimates. This finding provides valuable insights into how our estimator works in the Chinese Customs data.

When applied to a dataset including only observations with repeated trade patterns, conventional fixed effect and the TPSFE estimators are bound to yield similar results if the number of distinct trade patterns per product is small. Intuitively, at the limit, only one repeated trade pattern per product would iron out the difference in the variation used by the TPSFE estimator and the alternative (because destination and trade pattern variation would be equivalent in this case).[79] In line with this insight, in our Chinese Customs data, we find that the vast majority of

---

[79]For any given distribution of the unobserved fundamental shocks driving a firm's behavior, the number of

observations contain only one repeated trade pattern. This feature of our data implies that most of the bias reduction from applying the TPSFE estimator comes from eliminating the confounding variation of singleton trade patterns.

A natural question is whether the reduction in bias associated with the removal of confounding variation (associated with singleton trade patterns) outweighs the potential cost of information loss from excluding these observations.[80] In this respect, we should note that much can be learned from a systematic comparison of estimators through the lens of a model. If the systematic behavior in these singleton trade patterns is explained by factors that move additively along the dimensions of the panel, both the TPSFE and the $FE(fit+fid)$ estimators should yield very similar estimates (regardless whether observations associated with singleton trade patterns are included in the sample). In our empirical findings, similar estimates across estimators are obtained for the cases of exporters of low differentiation goods and private firms. In contrast, if the systematic behavior in trade patterns is explained by factors varying in complex ways, the observations associated with singleton trade patterns would confound conventional fixed effects estimators, as shown in table 3. The TPSFE and the $FE(fit+fid)$ estimators would yield very different estimates in this case. In our empirical analysis, we find large differences in estimates for the exporters of high differentiation goods, SOEs and FIEs. We take these results as evidence of an additional benefit of using the TPSFE estimator in conjunction with other fixed effect estimators; together, these different estimators can act as a tool to gain insight into the variation of unobservable factors driving firms' exporting and pricing choices.

# 6    Concluding Remarks

We close our study with two observations highlighting the significance of our contributions on methodological and policy grounds. Methodologically, a notable conclusion of our study is that, at fine levels of disaggregation (i.e., firm-product-destination), appropriately specified fixed effect estimators may actually perform quite well in relation to alternative, powerful methods that rely on the direct estimation of productivity and (unobservable) marginal costs at the firm level. The development of productivity- and cost-estimation methods has clearly enabled firm-level studies to break important new ground, shedding light on the level and time variation in firms' markups. Yet, applying these methods to our question of interest, concerning the time variation of markups at the

---

(different) repeated trade patterns can be expected to grow in the length of the sample (over time) and the average lifespan of traded products – as similar combinations of shocks would have a higher chance of reappearing over time and generating identical repeated trade patterns. Thus, variation associated with destinations within trade patterns would increase, leading the difference between the two estimators to grow.

[80]Specifically, singleton trade patterns could capture circumstances/shocks which, although rare in the sense that they do not translate into repeated trade patterns in the sample, contain useful information to ascertain elasticities.

product-destination level, gives rise to a key issue. Even if one could obtain the required data for the universe of firms in our sample, information on production inputs would generally be available only at the firm level, not at the firm-product level.[81] Estimates of marginal cost at the firm-product-destination level could still be obtained under some assumptions on how inputs are allocated across products and destinations; e.g., by positing that the production functions of single-product, single-destination firms are representative of those of multi-product multi-destination firms. However, in this paper we have shown that, under the identification assumptions of De Loecker, Goldberg, Khandelwal and Pavcnik (2016), well-defined fixed effect estimators would also give unbiased estimates of the markup elasticity to exchange rates. Future research may integrate these different approaches as complementary tools, with application to a wide range of topics including the effects of taxes and tariffs at the international and regional levels.

Concerning policy, the rising importance of China as a global exporter has spawned research into how enhanced competitive pressures worldwide have influenced corporates' decisions to upgrade their product mix (Bernard, Jensen and Schott (2006)), innovate (Bloom, Draca and Van Reenen (2016)), lay off workers (Autor, Dorn and Hanson (2013), Pierce and Schott (2016)), and outsource to lower wage countries (Pierce and Schott (2016)). Business people and economists routinely speak of the problem of "the China price," the low price of Chinese merchandise that exporters from other markets and domestic import-competing firms must match if they want to survive. Our contribution is to offer a more detailed and refined account of the nature of competitive pressures originating in China, one that cautions against overplaying the role of exchange rates in the policy debate. Our estimated markup elasticities imply that, for roughly 50% of the value of exports from China, a renminbi appreciation would not yield a uniform impact on Chinese prices. Because of the strategic response of Chinese firms that hold market power, the impact would vary considerably in different destinations and product markets. The effectiveness of a renminbi appreciation in reducing China's competitive pressure globally is far from certain.

---

[81]In addition, as it is well understood, the mapping between customs databases and industrial-survey data is incomplete in most countries, raising issues of the representativeness of data in the matched sample. Additionally, balance sheet data is typically only available at annual frequencies.

# References

**Abowd, John M, Francis Kramarz, and David N Margolis.** 1999. "High Wage Workers and High Wage Firms." *Econometrica*, 67(2): 251–333.

**Abowd, John, Robert H. Creecy, and Francis Kramarz.** 2002. "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data." Center for Economic Studies, U.S. Census Bureau Longitudinal Employer-Household Dynamics Technical Papers.

**Alessandria, George, Joseph P. Kaboski, and Virgiliu Midrigan.** 2010. "Inventories, Lumpy Trade, and Large Devaluations." *The American Economic Review*, 100(5): 2304–2339.

**Amiti, Mary, Oleg Itskhoki, and Jozef Konings.** 2014. "Importers, Exporters, and Exchange Rate Disconnect." *The American Economic Review*, 104(7): 1942–1978.

**Amiti, Mary, Oleg Itskhoki, and Jozef Konings.** 2019. "International shocks, variable markups, and domestic prices." *The Review of Economic Studies*, 86(6): 2356–2402.

**Amiti, Mary, Oleg Itskhoki, and Jozef Konings.** 2020. "Dominant Currencies: How Firms Choose Currency Invoicing and Why It Matters." NBER Working Paper 27926.

**Atkeson, Andrew, and Ariel Burstein.** 2008. "Pricing-to-Market, Trade Costs, and International Relative Prices." *The American Economic Review*, 98(5): 1998–2031.

**Autor, David H., David Dorn, and Gordon H. Hanson.** 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *The American Economic Review*, 103(6): 2121–68.

**Bas, Maria, Thierry Mayer, and Mathias Thoenig.** 2017. "From Micro to Macro: Demand, Supply, and Heterogeneity in the Trade Elasticity." *Journal of International Economics*, 108: 1–19.

**Bergin, Paul R., and Robert C. Feenstra.** 2001. "Pricing-to-Market, Staggered Contracts, and Real Exchange Rate Persistence." *Journal of International Economics*, 54(2): 333–359.

**Berman, Nicolas, Philippe Martin, and Thierry Mayer.** 2012. "How Do Different Exporters React to Exchange Rate Changes?" *The Quarterly Journal of Economics*, 127(1): 437–492.

**Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott.** 2006. "Survival of the Best Fit: Exposure to Low-wage Countries and the (uneven) Growth of U.S. Manufacturing Plants." *Journal of International Economics*, 68(1): 219–237.

**Bloom, Nicholas, Mirko Draca, and John Van Reenen.** 2016. "Trade Induced Technical Change? the Impact of Chinese Imports on Innovation, IT and Productivity." *The Review of Economic Studies*, 83(1): 87–117.

**Burstein, Ariel, and Gita Gopinath.** 2014. "International Prices and Exchange Rates." *Handbook of International Economics*, 4: 391–451.

**Card, David, Ana Rute Cardoso, and Patrick Kline.** 2016. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." *The Quarterly Journal of Economics*, 131(2): 633–686.

**Chatterjee, Arpita, Rafael Dix-Carneiro, and Jade Vichyanond.** 2013. "Multi-product Firms and Exchange Rate Fluctuations." *American Economic Journal: Economic Policy*, 5(2): 77–110.

**Cheng, L. Lai-Shen, and Rint Sybesma.** 1998. "Yi-wang Tang, Yi-ge Tang: Classifiers and Massifiers." *Tsing-hua Journal of Chinese Studies, New Series*, XXVIII(3): 385–412.

**Cheng, L. Lai-Shen, and Rint Sybesma.** 1999. "Bare and Not-so-bare Nouns and the Structure of NP." *Linguistic Inquiry*, 30(4): 509–542.

**Correia, Sergio.** 2017. "Linear Models with High-dimensional Fixed Effects: An Efficient and Feasible Estimator." *Working Paper*.

**Corsetti, Giancarlo, and Luca Dedola.** 2005. "A Macroeconomic Model of International Price Discrimination." *Journal of International Economics*, 67(1): 129–155.

**Corsetti, Giancarlo, Luca Dedola, and Sylvain Leduc.** 2008. "High Exchange-Rate Volatility and Low Pass-Through." *Journal of Monetary Economics*, 55(6): 1113–1128.

**Crowley, Meredith, Ning Meng, and Huasheng Song.** 2018. "Tariff Scares: Trade Policy Uncertainty and Foreign Market Entry by Chinese Firms." *Journal of International Economics*, 114: 96–115.

**Dai, Mi, and Jianwei Xu.** 2017. "Firm-specific Exchange Rate Shocks and Employment Adjustment: Evidence from China." *Journal of International Economics*, 108: 54–66.

**De Loecker, Jan, and Frederic Warzynski.** 2012. "Markups and Firm-level Export Status." *The American Economic Review*, 102(6): 2437–71.

**De Loecker, Jan, Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik.** 2016. "Prices, Markups, and Trade Reform." *Econometrica*, 84(2): 445–510.

**Dornbusch, Rudiger.** 1987. "Exchange Rates and Prices." *The American Economic Review,* 77(1): 93–106.

**Fitzgerald, Doireann, and Stefanie Haller.** 2014. "Pricing-to-Market: Evidence from Plant-Level Prices." *The Review of Economic Studies,* 81(2): 761–786.

**Fitzgerald, Doireann, and Stefanie Haller.** 2018. "Exporters and Shocks." *Journal of International Economics,* 113: 154–171.

**Fitzgerald, Doireann, Stefanie Haller, and Yaniv Yedid-Levi.** 2016. "How Exporters Grow." National Bureau of Economic Research Working Paper 21935.

**Goldberg, Pinelopi K., and Michael M. Knetter.** 1997. "Goods Prices and Exchange Rates: What Have We Learned?" *Journal of Economic Literature,* 35(3): 1243–1272.

**Gopinath, Gita.** 2015. "The International Price System." National Bureau of Economic Research Working Paper 21646.

**Gopinath, Gita, Emine Boz, Camila Casas, Federico J Díez, Pierre-Olivier Gourinchas, and Mikkel Plagborg-Møller.** 2020. "Dominant Currency Paradigm." *American Economic Review,* 110(3): 677–719.

**Gopinath, Gita, Oleg Itskhoki, and Roberto Rigobon.** 2010. "Currency Choice and Exchange Rate Pass-Through." *The American Economic Review,* 100(1): 304–336.

**Guimaraes, Paulo, and Pedro Portugal.** 2011. "A Simple Feasible Procedure to Fit Models with High-dimensional Fixed Effects." *The Stata Journal,* 10(4): 628–649.

**Hale, Galina, and Cheryl Long.** 2012. *Foreign Direct Investment in China: Winners and Losers.* World Scientific: Singapore.

**Han, Lu.** 2018. "The Mutable Geography of Firms' International Trade: Evidence and Macroeconomic Implications." *University of Cambridge mimeo.*

**Heckman, James J.** 1979. "Sample Selection Bias As a Specification Error." *Econometrica: Journal of the Econometric Society,* 153–161.

**Helpman, Elhanan, Marc Melitz, and Yona Rubinstein.** 2008. "Estimating Trade Flows: Trading Partners and Trading Volumes." *The quarterly journal of economics,* 123(2): 441–487.

**Hsieh, Chang-Tai, and Zheng (Michael) Song.** 2015. "Grasp the Large, Let Go of the Small: The Transformation of the State Sector in China." *Brookings Papers on Economic Activity,* 295–366.

**Khandelwal, Amit K., Peter K. Schott, and Shang-Jin Wei.** 2013. "Trade Liberalization and Embedded Institutional Reform: Evidence from Chinese Exporters." *American Economic Review*, 103(6): 2169–95.

**Knetter, Michael M.** 1989. "Price Discrimination by US and German Exporters." *The American Economic Review*, 79(1): 198–210.

**Kropf, Andreas, and Philip Sauré.** 2014. "Fixed costs per shipment." *Journal of International Economics*, 92(1): 166–184.

**Krugman, Paul.** 1986. "Pricing to Market When the Exchange Rate Changes." National Bureau of Economic Research Working Paper 1926.

**Kyriazidou, Ekaterini.** 1997. "Estimation of a Panel Data Sample Selection Model." *Econometrica: Journal of the Econometric Society*, 1335–1364.

**Manova, Kalina, and Zhiwei Zhang.** 2012. "Export Prices across Firms and Destinations." *The Quarterly Journal of Economics*, 127(1): 379–436.

**Mayer, Thierry, Marc J. Melitz, and Gianmarco I. Ottaviano.** 2014. "Market Size, Competition, and the Product Mix of Exporters." *The American Economic Review*, 104(2): 495–536.

**Melitz, Marc J., and Stephen J. Redding.** 2014. "Chapter 1 - Heterogeneous Firms and Trade." In *Handbook of International Economics*. Vol. 4 of *Handbook of International Economics*, ed. Gita Gopinath, Elhanan Helpman and Kenneth Rogoff, 1–54. Elsevier.

**Nakamura, Emi, and Jón Steinsson.** 2012. "Lost in Transit: Product Replacement Bias and Pricing to Market." *The American Economic Review*, 102(7): 3277–3316.

**Pierce, Justin R., and Peter K. Schott.** 2016. "The Surprisingly Swift Decline of US Manufacturing Employment." *The American Economic Review*, 106(7): 1632–62.

**Rauch, James E.** 1999. "Networks Versus Markets in International Trade." *Journal of International Economics*, 48(1): 7–35.

**Rios-Avila, Fernando.** 2015. "Feasible Fitting of Linear Models with N Fixed Effects." *The Stata Journal*, 15(3): 881–898.

**United Nations Statistics Division.** 2010. "Chapter 15 Quantity Measurement." In *International Merchandise Trade Statistics: Compiler's Manual, Revision 1.* https://unstats.un.org/wiki/display/I2CG/A.+++An+overview+of+the+World+Customs+Organization+standard+units+of+quantity.

**Wansbeek, Tom, and Arie Kapteyn.** 1989. "Estimation of the Error-components Model with Incomplete Panels." *Journal of Econometrics*, 41(3): 341–361.

**Wu, Mark.** 2016. "The 'China, Inc.' Challenge to Global Trade Governance." *Harvard International Law Journal*, 57(2): 261–324.

# Appendix for
# "Markets and Markup: A New Empirical Framework and Evidence on Exporters from China"

Giancarlo Corsetti

Cambridge, Cambridge-INET and CEPR

Meredith Crowley

Cambridge, Cambridge-INET and CEPR

Lu Han

Liverpool and CEPR

Huasheng Song

Zhejiang University

28 May 2021

# Contents

# A The TPSFE Estimator: A Three-step Procedure for Application to Customs Databases

As discussed in subsection 3.3, our proposed TPSFE approach consists of sequentially applying controls that reduce the variation of unobserved confounding variables.[1] To elaborate on our approach, in the first step of the sequential procedure, for every product in every firm, we strip out the component of the price that is common across the collection of foreign destinations reached in period $t$. We calculate the destination residual of each dependent and independent variable by subtracting the mean value of each variable (across destinations) over all active destinations for a firm's product in a period:

$$\widetilde{x}_{fidt} \equiv x - \frac{1}{n_{fit}^D} \sum_{d \in D_{fit}} x \quad \forall x \in \{p_{fidt}, q_{fidt}, e_{dt}, \boldsymbol{x}_{dt}\} \tag{A1}$$

where $n_{fit}^D$ is the number of active foreign destinations of firm $f$ selling product $i$ in year $t$ and $D_{fit}$ denotes the set of destinations of this firm-product pair in year $t$; $e_{dt}$ is the bilateral exchange rate defined as the units of RMB per units of destination market currency and $\boldsymbol{x}_{dt}$ is a vector of destination-specific macro variables including local CPI, real GDP and import-to-GDP ratio.

Our second step applies firm-product-destination-trade pattern ($fidD$) fixed effects to the residual prices, exchange rates, and other explanatory variables obtained in the first step. That is, we subtract the mean of the $\widetilde{x}_{fidt}$ variables for all time periods associated with the firm-product-destination-trade pattern $fidD$, i.e., $t \in T_{fidD}$:

$$\ddot{x}_{fidt} \equiv \widetilde{x}_{fidt} - \frac{1}{n_{fidD}^T} \sum_{t \in T_{fidD}} \widetilde{x}_{fidt} \quad \forall x \in \{p_{fidt}, p_{fidt}, e_{dt}, \boldsymbol{x}_{dt}\} \tag{A2}$$

where $\ddot{x}_{fidt}$ are the twice-differenced variables. Note that the aggregate variables which normally vary along only two dimensions $d$ and $t$ may "become" firm and product specific, i.e., $\ddot{e}_{fidt}$ and $\ddot{\boldsymbol{x}}_{fidt}$ due to the unbalancedness of the panel.

Using these twice-differenced variables, in the final step, we run an OLS regression that identifies how markups respond to the bilateral exchange rate; this approach exploits cross-destination variation in prices within a firm-product's trade pattern as well as intertemporal variation in

---

[1]The key contribution of our approach is to use trade patterns to construct better fixed effects. As long as the correct fixed effect specification is applied, the proposed method can be estimated in various ways, i.e., not necessarily using the sequential procedure above. For example, an equivalent approach is to directly specify two set of fixed effects, i.e., the firm-product-time $fit$ and the firm-product-destination-trade pattern $fidD$ fixed effects, and use standard iterative estimation procedures such as `reghdfe` to directly estimate the coefficients of interest. A key advantage of our sequential approach compared to alternative statistically iterative procedures is that each step gives a well defined economic object to interpret, which enables us to clearly state the underlying variation of the variables used for identification.

prices within a *time pattern of export participation* at the destination and trade pattern level for a firm-product pair:

$$\ddot{p}_{fidt} = \beta_0 + \beta_1 \ddot{e}_{fidt} + \ddot{\boldsymbol{x}}'_{fidt} \boldsymbol{\beta}_2 + \ddot{v}_{fidt}. \tag{A3}$$

We refer to the above procedure as the *trade pattern sequential fixed effects* (TPSFE) estimator. $\beta_1$ is the markup elasticity to bilateral exchange rates.

**Quantity Elasticity.** The same estimator can be applied to study destination specific adjustment in quantities, corresponding to adjustment in markups. Substituting the dependent variable described in the procedures above with the quantity sold, we obtain an estimator of the quantity elasticity to bilateral exchange rates:

$$\ddot{q}_{fidt} = \zeta_0 + \zeta_1 \ddot{e}_{fidt} + \ddot{\boldsymbol{x}}'_{fidt} \boldsymbol{\zeta}_2 + \ddot{v}_{fidt} \tag{A4}$$

in which $\ddot{q}_{fidt}$ is the residual quantity sold, that is, demeaned according to equations (A1) and (A2). This approach is comparable to the estimation strategy of Berman, Martin and Mayer (2012). It captures the *overall* effects of changing bilateral exchange rates on quantities—direct and indirect, i.e., via markup adjustment.

**Quantity-Markup Elasticity: A Two-Step Approach.** Based on our approach, however, we can also build a two-step statistical procedure to estimate the quantity responses specifically driven by markup adjustments to exchange rate movements. This consists of deriving the predicted prices, $\widehat{\ddot{p}}_{fidt}$ from the TPSEF estimator (A3) in the first step:

$$\widehat{\ddot{p}}_{fidt} = \widehat{\beta}_0 + \widehat{\beta}_1 \ddot{e}_{fidt} + \ddot{\boldsymbol{x}}'_{fidt} \widehat{\boldsymbol{\beta}}_2. \tag{A5}$$

In the second step, the predicted prices replace the exchange rate among the explanatory variables in the *relative quantity* regression:

$$\ddot{q}_{fidt} = \gamma_0 + \gamma_1 \widehat{\ddot{p}}_{fidt} + \ddot{\boldsymbol{x}}'_{fidt} \boldsymbol{\gamma}_2 + \ddot{u}_{fidt}. \tag{A6}$$

The coefficient $\gamma_1$ is a projection of changes in relative quantities on changes in relative prices driven by changes in the relative market condition measure, $\ddot{e}_{fidt}$, while controlling for other aggregate variables. Statistically, as in all two-step least square approaches, the relationship among the three elasticities (markup elasticity, quantity elasticity and quantity-markup elasticity) is given by $\widehat{\gamma}_1 \approx \widehat{\zeta}_1 / \widehat{\beta}_1$.

**The $Cor(\ddot{q}, \ddot{p})$ estimator.** For comparison purposes, we estimate the general relationship between the cross-market adjustments of prices and quantities by regressing the twice-differenced quantities directly on twice-differenced prices (labelled $Cor(\ddot{q}, \ddot{p})$) without projecting the price

changes on bilateral exchange rates, i.e.,

$$\ddot{q}_{fidt} = \lambda_0 + \lambda_1 \ddot{p}_{fidt} + \ddot{\boldsymbol{x}}'_{fidt} \boldsymbol{\lambda}_2 + \ddot{u}_{fidt} \tag{A7}$$

where $\lambda_1$ captures the general correlation between the relative quantity changes and the relative price changes across markets.

# B   The TPSFE Estimator in View of the Control Function Approach

As stated in the text, the TPSFE estimator is designed to address unobserved factors in an environment where firms endogenously select the set of destination markets for their product, i.e., firms' participation in a given market is endogenous. The idea is that the realized selection of markets (the trade patterns) conveys useful information about the unobservable factors that drive the selection process. By controlling for these patterns, we restrict the variation of unobservables in the selection equation. In this appendix, we discuss how our approach relates to the classical control function approach (e.g., Heckman (1979)) and the first difference approach pursued by Kyriazidou (1997).[2]

We start by rewriting the problem addressed by Heckman (1979) in his seminal work on selection in cross-sectional data. In what follows, think of $p_t$ as the price of a product, as a function of a set of controls $\boldsymbol{x}'_t$, observed if the firm decides to enter the market:

$$p_t = \boldsymbol{x}'_t \boldsymbol{\beta} + \varepsilon_t \tag{B1}$$
$$= \boldsymbol{x}'_t \boldsymbol{\beta} + E(\varepsilon_t | \boldsymbol{x}_t, s_t) + \nu_t$$
$$s_t = \mathbb{1}\{\boldsymbol{w}'_t \boldsymbol{\gamma} + u_t\} \tag{B2}$$

where $s_t$ is an indicator variable that equals one if $p_t$ is observed; $E(\varepsilon_t | \boldsymbol{x}_t, s_t)$ is the selection bias and $\nu_t \equiv [\varepsilon_t - E(\varepsilon_t | \boldsymbol{x}_t, s_t)]$ is an error term that is uncorrelated with the vector of observed vari-

---

[2]Our estimation approach is related to three strands of the panel data literature. The first strand focuses on estimating the parameter of interest in a panel data model with selection. Existing discussions are restricted to selection equations with one dimensional fixed effects or those that can be combined into one dimensional fixed effects (see recent handbook chapters Verbeek and Nijman (1996), Honoré, Vella and Verbeek (2008) and Matyas (2017) for a complete literature review). The second strand constructs methods of estimating selection equations with unobserved heterogeneity along two dimensions (e.g., Fernández-Val and Weidner (2016) and Charbonneau (2017)). Our approach differs from theirs in that we do not need to estimate the selection equation, but instead, we rely on the realized patterns and a third panel dimension to address the selection problem. A few papers have examined multi-dimensional fixed effects in unbalanced panels (e.g., Wansbeek and Kapteyn (1989) and Balazsi, Matyas and Wansbeek (2018)). However, they do not discuss the type of selection rules for which their fixed effect estimators are consistent.

ables $\boldsymbol{x_t}$ and the selection bias. $\boldsymbol{w}_t$ is a vector of observed variables in the selection equation which can overlap with the elements in $\boldsymbol{x}_t$. As is well known, selection is a problem if $E(\varepsilon_t|\boldsymbol{x}_t, s_t) \neq 0$. The solution of Heckman (1979) is to estimate the function of $E(\varepsilon_t|\boldsymbol{x}_t, s_t)$ under some parametric assumptions and then add the predicted value $\widehat{E(\varepsilon_t|\boldsymbol{x}_t, s_t)}$ as a control variable in the main estimating equation. The essence of this approach is to estimate the parameter of interest conditional on the probability of an observation being observed.

Closer to our problem where the firm chooses among possible export destination markets, Kyriazidou (1997) studies selection in a two dimensional panel with one fixed effect:

$$p_{dt} = \boldsymbol{x}'_{dt}\boldsymbol{\beta} + \mathcal{M}_d + \varepsilon_{dt} \tag{B3}$$
$$= \boldsymbol{x}'_{dt}\boldsymbol{\beta} + \mathcal{M}_d + E(\mathcal{M}_d|\boldsymbol{x}_{dt}, s_{dt}) + E(\varepsilon_{dt}|\boldsymbol{x}_{dt}, s_{dt}) + \nu_{dt}$$
$$s_{dt} = \mathbb{1}\{\boldsymbol{w}'_{dt}\boldsymbol{\gamma} + \mathcal{W}_d + u_{dt}\} \tag{B4}$$

where $\mathcal{M}_d$ and $\mathcal{W}_d$ are unobserved variables varying along the destination $d$ dimension (i.e. destination fixed effects). $E(\mathcal{M}_d|\boldsymbol{x}_{dt}, s_{dt})$ and $E(\varepsilon_{dt}|\boldsymbol{x}_{dt}, s_{dt})$ represent the selection biases caused by the unobserved destination-specific heterogeneity and other omitted variables, respectively. $\nu_{dt} \equiv [\varepsilon_{dt} - E(\varepsilon_{dt}|\boldsymbol{x}_{dt}, s_{dt}) - E(\mathcal{M}_d|\boldsymbol{x}_{dt}, s_{dt})]$ is an error term that is uncorrelated with the observed explanatory variables and the selection biases. $p_{dt}$ denotes the price and $s_{dt}$ is an indicator variable that takes a value of one if the firm exports to destination $d$ in period $t$ and zero otherwise.[3] Kyriazidou (1997) notes that $E(\mathcal{M}_d|\boldsymbol{x}_{dt}, s_{dt})$ and $E(\varepsilon_{dt}|\boldsymbol{x}_{dt}, s_{dt})$ no longer vary along the time dimension when $\boldsymbol{w}'_{d1}\boldsymbol{\gamma} = \boldsymbol{w}'_{d2}\boldsymbol{\gamma}$ under the following *conditional exchangeability* condition:

$$F(\varepsilon_{d1}, \varepsilon_{d2}, u_{d1}, u_{d2}|\boldsymbol{\psi}_d) = F(\varepsilon_{d2}, \varepsilon_{d1}, u_{d2}, u_{d1}|\boldsymbol{\psi}_d) \tag{B5}$$

where $\boldsymbol{\psi}_d \equiv (\boldsymbol{x}_{d1}, \boldsymbol{x}_{d2}, \boldsymbol{w}_{d1}, \boldsymbol{w}_{d2}, \mathcal{W}_d, \mathcal{M}_d)$ is a destination specific vector containing information on observed and unobserved variables. Condition (B5) states that $(\varepsilon_{d1}, \varepsilon_{d2}, u_{d1}, u_{d2})$ and $(\varepsilon_{d2}, \varepsilon_{d1}, u_{d2}, u_{d1})$ are identically distributed conditional on $\boldsymbol{\psi}_d$. As noted by Kyriazidou (1997), the main term causing the selection bias, $E(\varepsilon_{dt}|\boldsymbol{x}_{dt}, s_{dt})$, is no longer time-varying when $\boldsymbol{w}'_{d1}\boldsymbol{\gamma} = \boldsymbol{w}'_{d2}\boldsymbol{\gamma}$ under condition

---

[3]Kyriazidou (1997) discusses a case in which the number of time periods is small ($n^T = 2$). Therefore, a Heckman (1979) style estimator cannot be applied as it will suffer from the incidental parameters problem due to the limited time dimension.

4

(B5):

$$E(\varepsilon_{d1}|s_{d1} = 1, s_{d2} = 1|\boldsymbol{\psi}_d)$$
$$\equiv E(\varepsilon_{d1}|u_{d1} < \boldsymbol{w}'_{d1}\boldsymbol{\gamma} + \mathcal{W}_d, u_{d2} < \boldsymbol{w}'_{d2}\boldsymbol{\gamma} + \mathcal{W}_d, \boldsymbol{\psi}_d)$$
$$= E(\varepsilon_{d1}|u_{d1} < \boldsymbol{w}'_{d2}\boldsymbol{\gamma} + \mathcal{W}_d, u_{d2} < \boldsymbol{w}'_{d1}\boldsymbol{\gamma} + \mathcal{W}_d, \boldsymbol{\psi}_d) \tag{B6}$$
$$= E(\varepsilon_{d2}|u_{d2} < \boldsymbol{w}'_{d2}\boldsymbol{\gamma} + \mathcal{W}_d, u_{d1} < \boldsymbol{w}'_{d1}\boldsymbol{\gamma} + \mathcal{W}_d, \boldsymbol{\psi}_d) \tag{B7}$$
$$\equiv E(\varepsilon_{d2}|s_{d2} = 1, s_{d1} = 1|\boldsymbol{\psi}_d)$$

where the first equality (B6) holds because $\boldsymbol{w}'_{d1}\boldsymbol{\gamma} = \boldsymbol{w}'_{d2}\boldsymbol{\gamma}$ and the second equality (B7) holds because of the *conditional exchangeability* condition (B5). Since the selection bias is no longer time varying, i.e., $E(\varepsilon_{d1}|s_{d1} = 1, s_{d2} = 1|\boldsymbol{\psi}_d) = E(\varepsilon_{d2}|s_{d2} = 1, s_{d1} = 1|\boldsymbol{\psi}_d)$, it can be absorbed by destination fixed effects. Kyriazidou (1997) proposes a two-step estimator: the first step consistently estimates $\widehat{\boldsymbol{\gamma}}$ and the second step differences out the fixed effect and the selection terms conditional on destinations for which $\boldsymbol{w}'_{d1}\widehat{\boldsymbol{\gamma}} = \boldsymbol{w}'_{d2}\widehat{\boldsymbol{\gamma}}$.

Our problem is specified in (B8) and (B9) below:

$$p_{fidt} = \boldsymbol{x}'_{dt}\boldsymbol{\beta} + \mathcal{M}_{fid} + \mathcal{C}_{fit} + \varepsilon_{fidt} \tag{B8}$$
$$s_{fidt} = \mathbb{1}\{\boldsymbol{w}'_{dt}\boldsymbol{\gamma} + \mathcal{W}_{fid} + \mathcal{Q}_{fit} + u_{fidt}\} \tag{B9}$$

This problem differs from Kyriazidou (1997)'s in two crucial respects. On the one hand, our problem adds unobserved firm-product-time-varying variables $\mathcal{C}_{fit}$ to equation (B3) and $\mathcal{Q}_{fit}$ to equation (B4). In the presence of these time-varying unobserved factors, the *conditional exchangeablitiy* condition no longer holds. On the other hand, many aggregate-level economic indicators of interest in our study—e.g., exchange rates—vary along the destination and time dimensions, but not at the firm or product dimensions. This is actually helpful. As discussed below, the fact that key variables vary along dimensions that are a subset of the dimensions of the dependent variable facilitates the control of selection biases.

While the method we propose to address the above problem is conceptually close to Kyriazidou (1997), the approach we take is fundamentally different. Specifically, if we were to follow Kyriazidou (1997)'s approach, we would require all variables driving $\mathcal{Q}_{fit}$ to be observed and controlled for. For our purposes, however, this condition cannot be satisfied—if only because the marginal cost is unobserved and cannot be generally estimated at product-firm level. Rather, we need to rely on a method that avoids direct estimation of the selection equation and works in a multi-dimensional panel where more than one fixed effect is present in both the structural equation and the selection equation. Our main innovation is to use the realized selection pattern in a panel dimension, instead of the observed variables in the selection equation, to control for selection biases.

5

Before analyzing how our method addresses the general problem characterized in equations (B8) and (B9), we find it useful to provide insight by focusing on a two-dimensional panel, tracking the choices of a single firm selling one product across a set of endogenous destinations.

## B.1   A two dimensional panel case

Consider the following for a firms' destination choices with two panel dimensions, destination $d$ and time $t$:

$$p_{dt} = \boldsymbol{x}'_{dt}\boldsymbol{\beta} + \mathcal{M}_d + \mathcal{C}_t + \varepsilon_{dt} \tag{B10}$$

$$s_{dt} = \mathbb{1}\{u_{dt}\} \tag{B11}$$

where $\mathcal{M}_d$ and $\mathcal{C}_t$ are unobserved destination and time specific factors, respectively, which are potentially correlated with the explanatory variables contained in the vector $\boldsymbol{x}_{dt}$. The price $p_{dt}$ is observed only if $s_{dt}$ equals one or equivalently, if $u_{dt} > 0$.

The first two steps in our approach involve transforming the variables in (B10) to eliminate the unobserved destination and time specific factors. Specifically, in the first step, we demean variables at the time $(t)$ dimension. In the second step, we demean variables at the destination-trade pattern $(dD)$ dimension. After applying these two transformations,

$$\ddot{p}_{dt} = \ddot{\boldsymbol{x}}'_{dt}\boldsymbol{\beta} + \ddot{\varepsilon}_{dt} \tag{B12}$$

where

$$\ddot{\boldsymbol{x}}_{dt} = \boldsymbol{x}_{dt} - \frac{1}{n_t^D}\sum_{d\in D_t}\boldsymbol{x}_{dt} - \frac{1}{n_{dD}^T}\sum_{t\in T_{dD}}\boldsymbol{x}_{dt} + \frac{1}{n_{dD}^T}\sum_{t\in T_{dD}}\frac{1}{n_t^D}\sum_{d\in D_t}\boldsymbol{x}_{dt} \tag{B13}$$

$$\ddot{\varepsilon}_{dt} = \varepsilon_{dt} - \frac{1}{n_t^D}\sum_{d\in D_t}\varepsilon_{dt} - \frac{1}{n_{dD}^T}\sum_{t\in T_{dD}}\varepsilon_{dt} + \frac{1}{n_{dD}^T}\sum_{t\in T_{dD}}\frac{1}{n_t^D}\sum_{d\in D_t}\varepsilon_{dt}, \tag{B14}$$

$D_t$ is the set of destinations the firm serves at time $t$; and $n_t^D \equiv |D_t|$ the number of export destinations at time $t$. Similarly, $T_{dD}$ denotes the set of time periods in which a destination-specific trade pattern $dD$ is observed, and $n_{dD}^T$ represents the corresponding number of time periods in which the destination-specific trade pattern emerges. For our proposed approach to work in a two

6

dimensional panel, we need[4]

$$F(\varepsilon_{dD1}, \varepsilon_{dD2}, u_{dD1}, u_{dD2} | \boldsymbol{\psi}_{dD}) = F(\varepsilon_{dD2}, \varepsilon_{dD1}, u_{dD2}, u_{dD1} | \boldsymbol{\psi}_{dD}), \tag{B16}$$

where we use $\varepsilon_{dD1}$ to indicate the first error within the destination-specific trade pattern $dD$. We provide an example of the formulation of $dD$ in table B1. Given (B16), it is straightforward to see that the selection bias can be differenced out over two time periods within a destination-specific trade pattern $dD$, since the following relationship holds:

$$E\left(\varepsilon_{dDt} | u_{dD1} > 0, u_{dD2} > 0, \boldsymbol{\psi}_{dD}\right) = E\left(\varepsilon_{dD\tau} | u_{dD1} > 0, u_{dD2} > 0, \boldsymbol{\psi}_{dD}\right) \quad \forall \tau \in T_{dD} \tag{B17}$$

Condition (B16) can be viewed as the analog of the *conditional exchangeability* assumption imposed by Kyriazidou (1997). Instead of controlling for the relationship among the observed variables in the selection process (i.e., $\boldsymbol{w}'_{d1}\boldsymbol{\gamma} = \boldsymbol{w}'_{d2}\boldsymbol{\gamma}$), we control for the realised patterns of selection in a panel dimension (i.e., the pattern of $d$ conditional on $t$). That is, as long as the distribution of errors is the same for all time periods satisfying a destination-specific trade pattern $dD$, our approach produces unbiased and consistent estimates.[5]

### B.1.1 The indicative value of trade patterns: a numerical example

As an illustration of how conditioning on the realized trade patterns $D$ reduces selection bias, consider the numerical example in Table B1. This table reports the realization of trade flows to four destinations ($d = A, B, C, K$) over four time periods ($t = 1, 2, 3, 4$) with $u_{dt} \equiv \mathcal{W}_d + \mathcal{Q}_t$, where $\mathcal{W}_d$ and $\mathcal{Q}_t$ are destination-specific and time-specific variables in the selection process, respectively. The third and fourth columns in each panel in table B1 show the realized values of $\mathcal{W}_d$ and $\mathcal{Q}_t$ and the fifth column shows the corresponding outcomes, $s_{dt} = 1$ if the firm exports to destination $d$ at time $t$, $s_{dt} = 0$ otherwise. Only the dimensional indicator $d$ and the selection outcome $s_{dt}$ are observable to the researcher, whereas $\mathcal{W}_d$ and $\mathcal{Q}_t$ are unobservable.

---

[4]Note that Kyriazidou (1997)'s original conditions (and proofs) only cover the case when the number of time periods is equal to two. For a more general case with more than two time periods, we impose a condition as follows

$$E\left(\varepsilon_{dDt} | u_{dD1} > 0, ..., u_{dDn_{dD}^T} > 0, \psi_{dD}\right) = E\left(\varepsilon_{dD\tau} | u_{dD1} > 0, ..., u_{dDn_{dD}^T} > 0, \psi_{dD}\right) \quad \forall \tau \in T_{dD} \tag{B15}$$

As will be discussed later, our estimator works under a much weaker condition than (B15) if another panel dimension is available.

[5]It is straightforward to see the condition for consistency, i.e., $E(s_{dt}\ddot{\boldsymbol{x}}_{dt}\ddot{\varepsilon}_{dt}) = 0$, is satisfied under (B15). First, note that $\frac{1}{n_t^D}\sum_{d \in D_t} \varepsilon_{dt} - \frac{1}{n_{dD}^T}\sum_{t \in T_{dD}} \frac{1}{n_t^D}\sum_{d \in D_t} \varepsilon_{dt} = 0$. This is because the expression $\frac{1}{n_t^D}\sum_{d \in D_t} \varepsilon_{dt}$ is moving at the $dD$ dimension only. As there is no variation left after conditioning on the $dD$ dimension, the demeaning process naturally gives zero. Second, demeaning conditional on the same trade pattern is zero under assumption (B15), i.e., $E\left(\varepsilon_{dt} - \frac{1}{n_{dD}^T}\sum_{t \in T_{dD}} \varepsilon_{dt} \Big| s_{dD1}, s_{dD2}, s_{dD3}, ..., \psi_{dD}\right) = 0$.

Table B1: A Numerical Example of the Indicative Value of Trade Patterns

| $d$ | $t$ | $\mathcal{W}_d$ | $\mathcal{Q}_t$ | $s_{dt}$ | $D_t$ | | $dD$ | $t$ | $\mathcal{W}_d$ | $\mathcal{Q}_t$ | $s_{dt}$ | | $dD$ | $t$ | $\mathcal{W}_d$ | $\mathcal{Q}_t$ | $s_{dDt}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 1 | 1.5 | 1 | 1 | A-B-C | | A-**A-B-C** | 1 | 1.5 | **1** | 1 | | 1 | 1 | 1.5 | 1 | 1 |
| B | 1 | 0.5 | 1 | 1 | A-B-C | | B-**A-B-C** | 1 | 0.5 | **1** | 1 | | 2 | 1 | 0.5 | 1 | 1 |
| C | 1 | -0.5 | 1 | 1 | A-B-C | | C-**A-B-C** | 1 | -0.5 | **1** | 1 | | 3 | 1 | -0.5 | 1 | 1 |
| K | 1 | -1.5 | 1 | 0 | A-B-C | | K-**A-B-C** | 1 | -1.5 | **1** | 0 | | 4 | 1 | -1.5 | 1 | 0 |
| A | 2 | 1.5 | 0 | 1 | A-B | | A-**A-B** | 2 | 1.5 | **0** | 1 | | 5 | 2 | 1.5 | 0 | 1 |
| B | 2 | 0.5 | 0 | 1 | A-B | | B-**A-B** | 2 | 0.5 | **0** | 1 | | 6 | 2 | 0.5 | 0 | 1 |
| C | 2 | -0.5 | 0 | 0 | A-B | | C-**A-B** | 2 | -0.5 | **0** | 0 | | 7 | 2 | -0.5 | 0 | 0 |
| K | 2 | -1.5 | 0 | 0 | A-B | $\Rightarrow$ | K-**A-B** | 2 | -1.5 | **0** | 0 | $\Rightarrow$ | 8 | 2 | -1.5 | 0 | 0 |
| A | 3 | 1.5 | 1 | 1 | A-B-C | | A-**A-B-C** | 3 | 1.5 | **1** | 1 | | 1 | 3 | 1.5 | 1 | 1 |
| B | 3 | 0.5 | 1 | 1 | A-B-C | | B-**A-B-C** | 3 | 0.5 | **1** | 1 | | 2 | 3 | 0.5 | 1 | 1 |
| C | 3 | -0.5 | 1 | 1 | A-B-C | | C-**A-B-C** | 3 | -0.5 | **1** | 1 | | 3 | 3 | -0.5 | 1 | 1 |
| K | 3 | -1.5 | 1 | 0 | A-B-C | | K-**A-B-C** | 3 | -1.5 | **1** | 0 | | 4 | 3 | -1.5 | 1 | 0 |
| A | 4 | 1.5 | -1 | 1 | A | | A-**A** | 4 | 1.5 | **-1** | 1 | | 9 | 4 | 1.5 | -1 | 1 |
| B | 4 | 0.5 | -1 | 0 | A | | B-**A** | 4 | 0.5 | **-1** | 0 | | 10 | 4 | 0.5 | -1 | 0 |
| C | 4 | -0.5 | -1 | 0 | A | | C-**A** | 4 | -0.5 | **-1** | 0 | | 11 | 4 | -0.5 | -1 | 0 |
| K | 4 | -1.5 | -1 | 0 | A | | K-**A** | 4 | -1.5 | **-1** | 0 | | 12 | 4 | -1.5 | -1 | 0 |

Original     Transformed based on the realized trade pattern     Relabelled

The trade pattern $D_t$ in each period can be constructed based on realized values of the observed selection indicator in each destination, i.e., $\{s_{At}, s_{Bt}, s_{Ct}, s_{Kt}\}$. The last column of the first panel ("original") in table B1 shows the constructed trade patterns. In this specific example, there are three unique realized trade patterns, i.e., $D_1 = D_3 = \{1, 1, 1, 0\} \equiv$ "A-B-C", $D_2 = \{1, 1, 0, 0\} \equiv$ "A-B", and $D_4 = \{1, 0, 0, 0\} \equiv$ "A".

As illustrated in the example, the trade pattern formulated in each period is reflective of the values of $\mathcal{Q}_t$. Intuitively, in an estimation procedure without conditioning on the trade pattern, the comparisons (or the within variation) are made for each destination in all time periods. For example, the price of destination $A$ in period 1 is compared to that of periods 2, 3 and 4. After controlling for the trade pattern, the comparisons are made within the same trade pattern so that the price of destination $A$ in period 1 is only compared to that of period 3 where the underlying value of $\mathcal{Q}_t$ is the same.[6] Therefore, by separating observations by their corresponding trade patterns, we effectively control for the variation of unobserved variables. As part of this process, we drop those observations with no repeated trade pattern, i.e., observations in periods 2 and 4. Empirically, given that we have a very limited number of time periods (i.e., 15 years), the number of repeated trade patterns is quite small for each firm. Therefore, most of the bias reduction is

---

[6]More precisely, our approach exploits variation in the dependent and independent variables in the $dD$ and $t$ panel dimensions, rather than in the $d$ and $t$ panel dimensions.

done by dropping those observations associated with non-repetitive patterns.

In this numerical example, by construction, the unobserved time-varying factors are exactly the same for the identical trade pattern in periods 1 and 3, i.e., $\mathcal{Q}_1 = \mathcal{Q}_3$. Therefore, taking time differences after conditioning on the trade pattern *completely* eliminates the selection bias, i.e., condition (B17) is satisfied. In general, however, the underlying time-varying factors in the selection equation need not be restricted to be exactly the same for identical trade patterns. To see why, observe that $\mathcal{Q}_t$ take very similar values in those periods where the same pattern emerges:

$$
\begin{aligned}
\mathcal{W}_d + \mathcal{Q}_t > 0 \qquad &\text{if} \quad s_{dt} = 1 \\
\mathcal{W}_d + \mathcal{Q}_t \leq 0 \qquad &\text{if} \quad s_{dt} = 0
\end{aligned}
\tag{B18}
$$

In other words, given a trade pattern, the range of values that $\mathcal{Q}_t$ can take is limited. Conditioning on a trade pattern is therefore useful as it pins down the range of variation in $\mathcal{Q}_t$. By way of example, given the realized values of $\mathcal{W}_d$ specified in table B1, the conditions that $\mathcal{Q}_t$ needs to satisfy to be in the pattern $\{1,1,1,0\}$ are:

$$
1.5 + \mathcal{Q}_t > 0, \quad 0.5 + \mathcal{Q}_t > 0, \quad -0.5 + \mathcal{Q}_t > 0, \quad -1.5 + \mathcal{Q}_t \leq 0
\tag{B19}
$$

Since the equations in (B19) must be simultaneously satisfied, the range of values $\mathcal{Q}_t$ can take is $0.5 < \mathcal{Q}_t \leq 1.5$. Similarly, we can derive the condition for being in the patterns of $\{1,1,0,0\}$ and $\{1,0,0,0\}$ to be $-0.5 < \mathcal{Q}_t \leq 0.5$ and $\mathcal{Q}_t \leq -0.5$.

Since conditioning on the realized trade pattern restricts the variability of the unobserved $\mathcal{Q}_t$, our approach in general reduces the selection bias relative to conventional fixed effect approaches.

## B.2  General Setting

We now discuss the general multi-dimensional setting specified in (B8) and (B9).[7] With an additional dimension,[8] we can write the condition for identification as follows:

$$
E\left[E\left(\varepsilon_{fidDt}|\mathbf{s}_{fidD},\boldsymbol{\psi}_{fidD}\right)\Big|dt\right] = E\left[E\left(\varepsilon_{fidD\tau}|\mathbf{s}_{fidD},\boldsymbol{\psi}_{fidD}\right)\Big|dt\right] \quad \forall\,\tau \in T_{fidD}
\tag{B20}
$$

where $\mathbf{s}_{fidD} \equiv (\boldsymbol{w}'_{d1}\boldsymbol{\gamma} + \mathcal{W}_{fid} + \mathcal{Q}_{if1} + u_{fidD1} > 0, ..., \boldsymbol{w}'_{dn^T_{fidD}}\boldsymbol{\gamma} + \mathcal{W}_{fid} + \mathcal{Q}_{ifn^T_{fidD}} + u_{fidDn^T_{fidD}} > 0)$, $\boldsymbol{\psi}_{fidD} \equiv (\boldsymbol{x}_{dD1}, ..., \boldsymbol{x}_{dDn^T_{fidD}}, \boldsymbol{w}_{dD1}, ..., \boldsymbol{w}_{dDn^T_{fidD}}, \mathcal{W}_{fid}, \mathcal{M}_{fid})$ and $E(.|dt)$ means taking the expectation over the firm ($f$) and product ($i$) panel dimensions while keeping the destination and time panel dimensions fixed.

---

[7]Note that table B1 represents the observed trade pattern of a particular firm selling a particular product. In the customs data, we observe realized trade patterns of many firm-product pairs.

[8]In the following discussions, we consider firm and product as one combined panel dimension $fi$.

As can be seen from (B20), we no longer need the error to be zero conditional on the observed pattern ($E\left(\varepsilon_{fidDt} - \varepsilon_{fidD\tau}|\mathbf{s}_{fidD}, \boldsymbol{\psi}_{fidD}\right) = 0$) as in the two dimensional case. Instead, it is sufficient to have the expectation of $E\left(\varepsilon_{fidDt} - \varepsilon_{fidD\tau}|\mathbf{s}_{fidD}, \boldsymbol{\psi}_{fidD}\right)$ to be zero, once it is aggregated at the firm and product dimension. For example, if $E\left(\varepsilon_{fidDt} - \varepsilon_{fidD\tau}|\mathbf{s}_{fidD}, \boldsymbol{\psi}_{fidD}\right)$ consists of random errors for each firm and product, the mean of these random errors converges to zero when the number of firm-product pairs increases.

We now show that our proposed approach gives unbiased estimates under condition (B20). Let $v_{fidt} \equiv \mathcal{M}_{fid} + \mathcal{C}_{fit} + \varepsilon_{fidt}$. The underlying independent variables and the error term under our estimation approach can be written as

$$\ddot{\boldsymbol{x}}_{fidt} = \boldsymbol{x}_{dt} - \frac{1}{n_{fit}^D} \sum_{d \in D_{fit}} \boldsymbol{x}_{dt} - \frac{1}{n_{fidD}^T} \sum_{t \in T_{fidD}} \boldsymbol{x}_{dt} + \frac{1}{n_{fidD}^T} \sum_{t \in T_{fidD}} \frac{1}{n_{fit}^D} \sum_{d \in D_{fit}} \boldsymbol{x}_{dt} \tag{B21}$$

$$\ddot{v}_{fidt} = v_{fidt} - \frac{1}{n_{fit}^D} \sum_{d \in D_{fit}} v_{fidt} - \frac{1}{n_{fidD}^T} \sum_{t \in T_{fidD}} v_{fidt} + \frac{1}{n_{fidD}^T} \sum_{t \in T_{fidD}} \frac{1}{n_{fit}^D} \sum_{d \in D_{fit}} v_{fidt}. \tag{B22}$$

The independent variable of interest now varies along four dimensions because it embodies selection that varies across firms and products, even if the variable is specified for only two dimensions, i.e., $\boldsymbol{x}_{dt}$ or $e_{dt}$.

First, it is straightforward to verify that our estimator controls for firm-product-destination and firm-product-time fixed effects in the main estimation equation.

$$\ddot{v}_{fidt} = \mathcal{M}_{fid} + \mathcal{C}_{fit} + \varepsilon_{fidt} - \frac{1}{n_{fit}^D} \sum_{d \in D_{fit}} \left(\mathcal{M}_{fid} + \mathcal{C}_{fit} + \varepsilon_{fidt}\right)$$

$$- \frac{1}{n_{fidD}^T} \sum_{t \in T_{fidD}} \left(\mathcal{M}_{fid} + \mathcal{C}_{fit} + \varepsilon_{fidt}\right) + \frac{1}{n_{fidD}^T} \sum_{t \in T_{fidD}} \frac{1}{n_{fit}^D} \sum_{d \in D_{fit}} \left(\mathcal{M}_{fid} + \mathcal{C}_{fit} + \varepsilon_{fidt}\right)$$

$$= \varepsilon_{fidt} - \frac{1}{n_{fit}^D} \sum_{d \in D_{fit}} \left(\mathcal{M}_{fid} + \varepsilon_{fidt}\right) - \frac{1}{n_{fidD}^T} \sum_{t \in T_{fidD}} \left(\mathcal{C}_{fit} + \varepsilon_{fidt}\right)$$

$$+ \frac{1}{n_{fidD}^T} \sum_{t \in T_{fidD}} \frac{1}{n_{fit}^D} \sum_{d \in D_{fit}} \left(\mathcal{M}_{fid} + \mathcal{C}_{fit} + \varepsilon_{fidt}\right)$$

$$= \varepsilon_{fidt} - \frac{1}{n_{fit}^D} \sum_{d \in D_{fit}} \varepsilon_{fidt} - \frac{1}{n_{fidD}^T} \sum_{t \in T_{fidD}} \varepsilon_{fidt} + \frac{1}{n_{fidD}^T} \sum_{t \in T_{fidD}} \frac{1}{n_{fit}^D} \sum_{d \in D_{fit}} \varepsilon_{fidt}$$

$$= \ddot{\varepsilon}_{fidt}$$

Second, note that the exchange rate depends on the firm and product dimensions only through trade and time patterns. To see this, it is useful to rewrite the variables in expressions (B21) and

(B22) in terms of their corresponding variability:

$$\ddddot{\boldsymbol{x}}_{fidt} = \boldsymbol{x}_{dt} - \boldsymbol{x}_{Dt} - \boldsymbol{x}_{dT} + \boldsymbol{x}_{DT}$$

$$\ddddot{v}_{fidt} = v_{fidt} - v_{fiDt} - v_{fidT} + v_{fiDT}$$

$$= \varepsilon_{fidt} - \varepsilon_{fiDt} - \varepsilon_{fidT} + \varepsilon_{fiDT}$$

$$= \ddddot{\varepsilon}_{fidt}.$$

Rearranging these expressions, we can show that our main variables of interest $\boldsymbol{x}$ (including exchange rates) in the following expression no longer depend on firm and product dimensions:

$$\frac{1}{n^{FIDT}} \sum_{fidt} \ddddot{\varepsilon}_{fidt} \ddddot{\boldsymbol{x}}_{fidt} = \frac{1}{n^{FIDT}} \sum_{fidt} (\varepsilon_{fidt} - \varepsilon_{fiDt} - \varepsilon_{fidT} + \varepsilon_{fiDT}) \boldsymbol{x}_{dt} \tag{B23}$$

$$= \frac{1}{n^{FIDT}} \sum_{fidt} (\varepsilon_{fidt} - \varepsilon_{fidT}) \boldsymbol{x}_{dt}. \tag{B24}$$

As a result, the identification condition, $E(\ddddot{\varepsilon}_{fidt} \ddddot{\boldsymbol{x}}_{fidt} s_{fidt}) = 0$, can be rewritten as

$$E(\ddddot{\varepsilon}_{fidt} \ddddot{\boldsymbol{x}}_{fidt} s_{fidt})$$

$$= E\left[(\varepsilon_{fidt} - \varepsilon_{fidT}) \boldsymbol{x}_{dt} s_{fidt}\right]$$

$$= E\left\{\boldsymbol{x}_{dt} E\left[E\left(\varepsilon_{fidt} - \varepsilon_{fidT} | \mathbf{s}_{fidD}, \boldsymbol{\psi}_{fidD}\right) \Big| dt\right]\right\}$$

$$= E\left\{\boldsymbol{x}_{dt} E\left[E\left(\varepsilon_{fidDt} - \frac{1}{n^T_{fidD}} \sum_{\tau \in T_{fidD}} \varepsilon_{fidD\tau} | \mathbf{s}_{fidD}, \boldsymbol{\psi}_{fidD}\right) \Big| dt\right]\right\}$$

$$= 0 \tag{B25}$$

where the first equality follows from using (B24) under our proposed "within transformation"; the second equality from applying the law of iterated expectations; and the last equality from using condition (B20).

Two remarks are in order to clarify the implications of our identification condition and place our approach in the literature. First, note that the condition (B20) is trivially satisfied if $\varepsilon$ is always zero. For example, if goods sold to different destinations by the same firm under the same product category are identical, the marginal cost is only firm-product-time specific and therefore absorbed by $\mathcal{C}_{fit}$, leaving no additional residual term. It is worth stressing that the maintained assumption that marginal costs are non-destination-specific is implicit in studies aimed at estimating productivity (as these do not try to distinguish the marginal cost at the destination level)—see, e.g., Olley and Pakes (1996), Levinsohn and Petrin (2003), Wooldridge (2009) and

De Loecker, Goldberg, Khandelwal and Pavcnik (2016).[9]

Second, an important instance in which condition (B20) is satisfied is when the distribution of the destination-specific component does not change over time, e.g., when the composition of shipments is such that high quality varieties of a product are consistently sold to high-income destinations. From this perspective, the condition clarifies that the existence of destination-specific marginal cost components in $\varepsilon$ does not automatically lead to a violation of identification.

In the sections to follows, we reinterpret the identification implied by (B20) through the lens of two strands of the literature. In Section C we compare our estimator to alternative fixed effect estimators. In Section E, we give a structural interpretation of the identification condition. In Section F we show the theoretical relationship among the markup and quantity elasticities.

# C  The TPSFE and Fixed Effect Estimators: a Comparative Analysis of Identification

In this section, we carry out a comparative analysis of identification across estimators of markup elasticities in unbalanced panels. Subsection C.1 focuses on a two-dimensional panel to discuss how fixed effect and trade-pattern augmented fixed effect estimators can help address selection and/or omitted variable biases. In Subsection C.2 we extend the discussion to a three dimensional panel.

---

[9]Olley and Pakes (1996), Levinsohn and Petrin (2003) and Wooldridge (2009) estimate firm-level productivity and thus can infer the average marginal cost over all products and destinations at the firm level. De Loecker, Goldberg, Khandelwal and Pavcnik (2016) estimate the average marginal cost over destinations at the firm-product level. As an exercise, in appendix E, we explore an extension of De Loecker, Goldberg, Khandelwal and Pavcnik (2016) in which we add a destination dimension to production costs. We discuss the assumptions that would be required in a structural framework for (B20) to be satisfied. Specifically, we allow the functional form of the production function to be firm-product specific with a log-additive productivity term that is firm-product-destination specific. Note that De Loecker, Goldberg, Khandelwal and Pavcnik (2016) would not be identifiable under these assumptions as their identification strategy requires some degree of separability in the functional form in which they have assumed the production function to be product-specific and the Hicks-neutral productivity to be firm-specific. In this extended framework, we show that our identification strategy recovers an unbiased estimate of the markup elasticity even when the marginal cost at the firm-product level varies across destinations, but only if the production function is constant returns to scale. It is only when changes in relative demand across destinations lead to relative changes in quantities (which are associated with changes in destination-specific marginal cost) that condition (B20) will be violated. This is only the case if the production function is destination-specific. Under the standard assumptions of De Loecker, Goldberg, Khandelwal and Pavcnik (2016) where the production function is not destination-specific, our estimator yields unbiased estimates with constant returns to scale (CRS), increasing returns to scale (IRS) and decreasing returns to scale (DRS) production functions.

## C.1 Identifying the Markup Elasticity in a Two-dimensional Unbalanced Panel

In this subsection, we discuss the identification of the markup elasticity in a two-dimensional unbalanced panel. We offer the proofs of Propositions 1 and 2, showing, first, that an estimator adding destination and time fixed effects tackles the selection problem by implicitly applying trade pattern fixed effects; second, compared to an estimator with destination and time fixed effects, the scope of bias that can be effectively addressed by a fixed effect approach is extended by our TPSFE estimator.

As a reference, we write the following data generating process:

$$p_{dt} = \boldsymbol{x}'_{dt}\boldsymbol{\beta} + \mathcal{M}_d + \mathcal{C}_t + \varepsilon_{dt} \tag{C1}$$

$$s_{dt} = \mathbb{1}\{\boldsymbol{w}'_{dt}\boldsymbol{\gamma} + \mathcal{W}_d + \mathcal{Q}_t + u_{dt}\} \tag{C2}$$

where $d$ $(d = 1, \ldots, n^D)$ denotes the destinations of an exporter, and $t$ $(t = 1, \ldots, n^T)$, the years; $\boldsymbol{x}_{dt}$ and $\boldsymbol{w}_{dt}$ are vectors of observed variables including bilateral exchange rates; and $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are vectors of unknown parameters. $\mathcal{M}_d$ and $\mathcal{W}_d$, and $\mathcal{W}_t$ and $\mathcal{Q}_t$ are unobserved confounding components varying at the destination and the time dimension, respectively, that are correlated among themselves and with $\boldsymbol{x}_{dt}$ and $\boldsymbol{w}_{dt}$; $u_{dt}$ and $\varepsilon_{dt}$ are unobserved variables that vary at both destination and time dimensions.

An important observation is that, given the above process, an estimator adding both destination and time fixed effects gives unbiased estimates if $\varepsilon_{dt}$ and $u_{dt}$ are not correlated.[10] That is, if the unobserved variables (e.g., marginal costs) can be broken into additive variation in the destination and time panel dimensions, adding fixed effects simultaneously addresses both the omitted variable bias and the selection bias as $E(\varepsilon_{dt}|\boldsymbol{x}_{dt}, \boldsymbol{w}_{dt}, \mathcal{M}_d, \mathcal{C}_t, \mathcal{W}_d, \mathcal{Q}_t, \boldsymbol{w}'_{dt}\boldsymbol{\gamma} + \mathcal{W}_d + \mathcal{Q}_t + u_{dt} > 0) = 0$.

As is well known, a fixed effect estimator can be implemented in various ways. One way is to include dummy variables for each destination and time and run an OLS estimation with all of the dummy variables and the independent variables $x_{dt}$. However, this approach becomes infeasible when the set of dummies is very large. An alternative and commonly-used way is to avoid directly estimating the fixed effects by demeaning variables.[11]

We are now ready to consider the proofs of our propositions.

---

[10]Note that our TPSFE estimator allows for a less stringent condition than the uncorrelatedness between $\varepsilon_{dt}$ and $u_{dt}$. For example, if the correlation between $\varepsilon_{dt}$ and $u_{dt}$ is caused by unobserved factors that are destination-trade pattern specific, our TPSFE estimator also gives unbiased estimates. We compare our proposed estimator with a conventional fixed effect estimator adding destination and time fixed effects in Subsection C.1.2. A numerical example is given in Subsection C.1.3. The exact identification condition of our estimator is given by (B20).

[11]There are several interactive algorithms based on the demeaning logic, such as Guimaraes and Portugal (2011), Rios-Avila (2015), and Correia (2017).

### C.1.1 Proof of Proposition 1

The proof proceeds with two steps. In the first step, we construct a demeaned fixed effect estimator following Wansbeek and Kapteyn (1989). In the second step, we show that the constructed estimator implicitly applies trade pattern fixed effects.

**Step 1:** Let $n_t^D \left( n_t^D \leq n^D \right)$ be the number of observed destinations for year $t$. Let $n^{DT} \equiv \sum_t n_t^D$. Let $A_t$ be the $\left( n_t^D \times n^D \right)$ matrix obtained from the $(n^D \times n^D)$ identity matrix from which rows corresponding to destinations not observed in year $t$ have been omitted, and consider

$$Z \equiv \begin{pmatrix} Z_1, & Z_2 \\ n^{DT} \times n^D & n^{DT} \times n^T \end{pmatrix} \equiv \begin{bmatrix} A_1 & A_1 \iota_{n^D} \\ \vdots & & \ddots \\ A_{n^T} & & A_{n^T} \iota_{n^D} \end{bmatrix} \tag{C3}$$

where $\iota_x$ is a vector of ones with length $x$, e.g., $\iota_{n^D}$ is a vector of ones with length $n^D$. The matrix $Z$ gives the dummy-variable structure for the incomplete-data model. (For complete data, $Z_1 = \iota_{n^T} \otimes I_{n^D}, Z_2 = I_{n^T} \otimes \iota_{n^D}$.) Define

$$P_2 \equiv I_{n^{DT}} - Z_2 \left( Z_2' Z_2 \right)^{-1} Z_2' \tag{C4}$$

$$\bar{Z} \equiv P_2 Z_1. \tag{C5}$$

Wansbeek and Kapteyn (1989) show $P$ is the projection matrix onto the null-space of $Z$:

$$P \equiv P_2 - \bar{Z}(\bar{Z}'\bar{Z})^- \bar{Z}' \tag{C6}$$

where '$-$' stands for a generalized inverse. It follows that the model described in (C1) and (C2) can be estimated using OLS with the demeaned data obtained by pre-multiplying the data matrix $(Y, X)$ by the projection matrix $P$.

**Step 2:** We now show the projection matrix $P$ can be decomposed into two projection matrices with the second projection matrix applying destination and trade pattern fixed effects in additive terms. We begin by noting that the following relationship holds:

$$P \equiv P_2 - \bar{Z}(\bar{Z}'\bar{Z})^- \bar{Z}' = (I_{n^{DT}} - \bar{Z}(\bar{Z}'\bar{Z})^- \bar{Z}')P_2 \equiv P_1 P_2 \tag{C7}$$

where $P_1 \equiv I_{n^{DT}} - \bar{Z}(\bar{Z}'\bar{Z})^- \bar{Z}'$ and the equality of (C7) uses the fact that $P_2$ is idempotent (i.e., $P_2 Z_1 = P_2 P_2 Z_1 = P_2 \bar{Z}$). Therefore, applying the projection matrix $P$ to the data matrix $(Y, X)$ is equivalent to first pre-multiplying $(Y, X)$ by the projection matrix $P_2$, and then pre-multiplying $(P_2 Y, P_2 X)$ by the projection matrix $P_1$. The projection $P_2$ applied in the first step is essentially

a destination-demean process (the same first step as our TPSFE estimator).[12] The projection $P_1$ applied in the second step is, by definition, a "demeaning" process at the $\bar{Z}$ level. To see the exact dummy structure based on which the second "demeaning" process is applied, note that $\bar{Z}$ can be rewritten as

$$\bar{Z} = P_2 Z_1 = Z_1 - Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' Z_1 \tag{C8}$$

where $Z_1$ is a set of destination dummies as defined in (C3) and $Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' Z_1$ is a set of trade pattern dummies.

To see that $Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' Z_1$ follows a trade pattern structure, note that $Z_2 \left(Z_2' Z_2\right)^{-1} Z_2'$ is a block diagonal matrix with its diagonal blocks equal to a matrix of ones multiplied by (the inverse of) the number of destinations in each period, i.e.,

$$
\begin{aligned}
Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' &= \operatorname{diag}\left(\frac{1}{n_1^D} A_1 \iota_{n^D} \iota_{n^D}' A_1', ..., \frac{1}{n_{n^T}^D} A_{n^D} \iota_{n^D} \iota_{n^D}' A_{n^D}'\right) \\
&= \operatorname{diag}\left(\frac{1}{n_1^D} \iota_{n_1^D} \iota_{n_1^D}', ..., \frac{1}{n_{n^T}^D} \iota_{n_{n^T}^D} \iota_{n_{n^T}^D}'\right)
\end{aligned} \tag{C9}
$$

where the first equality holds by the definition of $Z_2$ in (C3) and by virtue of the fact that $\left(Z_2' Z_2\right)^{-1}$ is a diagonal matrix, with its elements indicating (the inverse of) the number of observed destinations in each period, i.e.,

$$\left(Z_2' Z_2\right)^{-1} = \operatorname{diag}\left(\frac{1}{n_1^D}, \frac{1}{n_2^D}, ..., \frac{1}{n_{n^T}^D}\right); \tag{C10}$$

and the second equality in (C8) holds by the definition of the $A$ matrices in (C3). Pre-multiplying to $Z_1$ by $Z_2 \left(Z_2' Z_2\right)^{-1} Z_2'$ and using the definition of $Z_1$, we have

$$Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' Z_1 = \begin{bmatrix} \frac{1}{n_1^D} \iota_{n_1^D} \iota_{n_1^D}' A_1 \\ \vdots \\ \frac{1}{n_{n^T}^D} \iota_{n_{n^T}^D} \iota_{n_{n^T}^D}' A_{n^D} \end{bmatrix} \tag{C11}$$

where $\iota_{n_t^D}' A_t$ gives the trade pattern in year $t$ and pre-multiplying it by $\iota_{n_t^D}$ repeats the same trade pattern $n_t^D$ times—resulting in the trade pattern matrix for all destinations in period $t$.[13]

Therefore, the second "demeaning" projection matrix $P_1 \equiv I_{n^{DT}} - \bar{Z}(\bar{Z}'\bar{Z})^- \bar{Z}'$ is applied on $\bar{Z}$ that consists of two *additive* parts: (a) the destination dummies $Z_1$ and (b) the trade pattern dummies $Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' Z_1$. QED.

---

[12]See the numerical example in subsection C.1.3.

[13]See Appendix C.1.3 for an numerical example of the matrices.

### C.1.2 Proof of Proposition 2

The key difference between our proposed TPSFE estimator and a conventional fixed effect estimator adding destination and time fixed effects lies in the way the trade patterns are applied in the second step. While the conventional approach applies the destination and trade pattern fixed effects additively (as can be seen from (C8) and (C11)), our estimator applies the trade pattern fixed effect multiplicatively.

We start our proof introducing notation and definitions. Denote the set of exporting destinations in year $t$ as $D_t$.[14] Let $\mathcal{TP}$ be the set of unique trade patterns in all years, i.e.,

$$\mathcal{TP} \equiv \{D_1, ..., D_{n^T}\}_{\neq} \tag{C12}$$

and $n^{\mathcal{TP}} \equiv |\mathcal{TP}|$ be the number of unique trade patterns. Let $\mathcal{TP}_x$ denote the $x$'th element of $\mathcal{TP}$. We create destination-specific trade patterns by combining the destinations in a trade pattern with the trade pattern itself, i.e., $\{(d, \mathcal{TP}_x) : d \in \mathcal{TP}_x\}$. Let $\mathcal{DTP}$ be the set of destination-specific trade patterns, i.e.,

$$\mathcal{DTP} \equiv \{(d, \mathcal{TP}_1) : d \in \mathcal{TP}_1, ..., (d, \mathcal{TP}_{n^{\mathcal{TP}}}) : d \in \mathcal{TP}_{n^{\mathcal{TP}}}\}. \tag{C13}$$

Let $n^{\mathcal{DTP}} \equiv |\mathcal{DTP}|$ be the number of unique destination-trade pattern pairs observed in the data.

The dummy structure of destination-specific trade patterns is given by the following $(n^{DT} \times n^{\mathcal{DTP}})$ matrix:

$$Z_3 \equiv \begin{bmatrix} B_1 \\ \vdots \\ B_{n^T} \end{bmatrix} \equiv \begin{bmatrix} K_{11} & \cdots & K_{1n^{\mathcal{TP}}} \\ \vdots & \ddots & \vdots \\ K_{n^T 1} & \cdots & K_{n^T n^{\mathcal{TP}}} \end{bmatrix} \tag{C14}$$

where $B_t$ is a $n_t^D \times n^{\mathcal{DTP}}$ matrix indicating the destination-specific trade patterns in period $t$. Each $B_t$ can be decomposed into $n^{\mathcal{TP}}$ block matrices with its $y$'th block being equal to an identity matrix if the trade pattern of period $t$, $D_t$, is the same as the $y$'th trade pattern, $\mathcal{TP}_y$, and a matrix of zeros otherwise. That is, $\forall x \in \{1, ..., n^T\}, y \in \{1, ..., n^{\mathcal{TP}}\}$,

$$K_{xy} \equiv \begin{cases} I_{n_x^D} & \text{if } D_x = \mathcal{TP}_y \\ \mathbf{0}_{n_x^D \times n_{\mathcal{TP}}^D(y)} & \text{if } D_x \neq \mathcal{TP}_y \end{cases} \tag{C15}$$

where $I_{n_x^D}$ is an identity matrix of size $n_x^D$; $\mathbf{0}_{n_x^D \times n_{\mathcal{TP}}^D(y)}$ is a matrix of zeros of size $n_x^D \times n_{\mathcal{TP}}^D(y)$; and $n_{\mathcal{TP}}^D(y) \equiv |\{d : d \in \mathcal{TP}_y\}|$ is the number of destinations in the $y$'th unique trade pattern $\mathcal{TP}_y$.

Let the projection matrix be $P_3 P_2$, where $P_3 \equiv I_{n^{DT}} - Z_3 (Z_3' Z_3)^{-1} Z_3'$. The first projection $P_2$

---

[14]In a vector form, $\iota_{n_t^D}' A_t$ indicates the set of destinations in year $t$.

is the same destination-demean progress, whereas the second projection $P_3$ applies demeaning at the destination-trade pattern level. As discussed in previous sections, the interactive construction of trade pattern fixed effects enables us to handle interactive error terms and reduce the time variation of the unobserved confounding variables.

To formally prove proposition 2, we need to show that

$$
\begin{aligned}
P_3 P_2 Z_1 &= \mathbf{0}, \\
P_3 P_2 Z_3 &= \mathbf{0}, \\
P_3 P_2 Z_2 &= \mathbf{0}.
\end{aligned}
\tag{C16}
$$

We begin by noting that the second relationship holds by definition (of $P_2$):

$$
P_3 P_2 Z_2 = [I_{n^{DT}} - Z_3 \left(Z_3' Z_3\right)^{-1} Z_3'][I_{n^{DT}} - Z_2 \left(Z_2' Z_2\right)^{-1} Z_2'] Z_2 = \mathbf{0}.
\tag{C17}
$$

We prove $P_3 P_2 Z_1 = \mathbf{0}$ and $P_3 P_2 Z_3 = \mathbf{0}$ by relying on two relationships that we state here and prove later in the text. First, the two projection matrices $T_3 \equiv Z_3 \left(Z_3' Z_3\right)^{-1} Z_3'$ and $T_2 \equiv Z_2 \left(Z_2' Z_2\right)^{-1} Z_2'$ commute:

$$
T_3 T_2 = T_2 T_3.
\tag{C18}
$$

Second, $T_3$ projects $Z_1$ to itself:

$$
T_3 Z_1 = Z_1.
\tag{C19}
$$

Given (C18) and (C19), it follows that

$$
\begin{aligned}
P_3 P_2 Z_1 &= [I_{n^{DT}} - T_3][I_{n^{DT}} - T_2] Z_1 \\
&= Z_1 - T_3 Z_1 + T_3 T_2 Z_1 - T_2 Z_1 \\
&= T_3 T_2 Z_1 - T_2 Z_1 \\
&= T_2 T_3 Z_1 - T_2 Z_1 \\
&= T_2 Z_1 - T_2 Z_1 \\
&= \mathbf{0}
\end{aligned}
\tag{C20}
$$

where the second equality is due to (C19); the third equality holds due to the commutativity (C18); the fourth equality applies (C19) one more time. Following the same procedure, it can be shown that $P_3 P_2 Z_3 = \mathbf{0}$.

We complete our proofs showing that (C18) and (C19) hold.

**Proof of** (C18):

*Proof.* We want to prove that the two projection matrices $Z_3 \left(Z_3'Z_3\right)^{-1} Z_3'$ and $Z_2 \left(Z_2'Z_2\right)^{-1} Z_2'$ commute. We do so by proving that the product of these two matrices $Z_3 \left(Z_3'Z_3\right)^{-1} Z_3'Z_2 \left(Z_2'Z_2\right)^{-1} Z_2'$ is symmetric.

$Z_3 \left(Z_3'Z_3\right)^{-1} Z_3'$ can be written as:

$$Z_3 \left(Z_3'Z_3\right)^{-1} Z_3' = \begin{bmatrix} B_1 \left(Z_3'Z_3\right)^{-1} B_1' & \cdots & B_1 \left(Z_3'Z_3\right)^{-1} B_{n^T}' \\ \vdots & \ddots & \vdots \\ B_1 \left(Z_3'Z_3\right)^{-1} B_{n^T}' & \cdots & B_{n^T} \left(Z_3'Z_3\right)^{-1} B_{n^T}' \end{bmatrix} \tag{C21}$$

The blocks of $Z_3 \left(Z_3'Z_3\right)^{-1} Z_3'$ can be further simplified using the following two observations. First, $\left(Z_3'Z_3\right)^{-1}$ is a $n^{\mathcal{DTP}} \times n^{\mathcal{DTP}}$ diagonal matrix with its elements indicating (the reverse of) the number of repetitions for each destination-trade pattern pair, i.e.,

$$
\begin{aligned}
\left(Z_3'Z_3\right)^{-1} &= \left(\sum_t B_t' B_t\right)^{-1} \\
&= \begin{bmatrix} \sum_t K_{t1}' K_{t1} & \cdots & \sum_t K_{t1}' K_{tn^{\mathcal{TP}}} \\ \vdots & \ddots & \vdots \\ \sum_t K_{tn^T}' K_{t1} & \cdots & \sum_t K_{tn^{\mathcal{TP}}}' K_{tn^{\mathcal{TP}}} \end{bmatrix}^{-1} \\
&= \begin{bmatrix} r_1^{\mathcal{TP}} I_{n_{\mathcal{TP}}^D(1)} & & \\ & \ddots & \\ & & r_{n^{\mathcal{TP}}}^{\mathcal{TP}} I_{n_{\mathcal{TP}}^D(n^{\mathcal{TP}})} \end{bmatrix}^{-1} \\
&= \operatorname{diag}\left(\frac{1}{r_1^{\mathcal{TP}}} I_{n_{\mathcal{TP}}^D(1)}, \ldots, \frac{1}{r_{n^{\mathcal{TP}}}^{\mathcal{TP}}} I_{n_{\mathcal{TP}}^D(n^{TP})}\right)
\end{aligned} \tag{C22}
$$

where $r_z^{\mathcal{TP}} \equiv |\{t : D_t = \mathcal{TP}_z\}|$ is the number of periods that the trade pattern $\mathcal{TP}_z$ is observed for $z \in \left\{1, \ldots, n^{\mathcal{TP}}\right\}$. The third equality holds as $K_{th}' K_{tj} = \mathbf{0} \ \forall h \neq j$ and $K_{th}' K_{tj} = I_{n_h^D} \ \forall h = j$ by definitions of (C14) and (C15).

Second, the $(h, j)$ block of $Z_3 \left(Z_3'Z_3\right)^{-1} Z_3'$, i.e., $B_h \left(Z_3'Z_3\right)^{-1} B_j'$, is equal to a matrix of zeros if the trade pattern of period $h$ is different from that of period $j$ and is equal to an identity matrix multiplied by a scalar if the trade pattern of the two periods is the same:

$$B_h \left(Z_3'Z_3\right)^{-1} B_j' = \sum_{z \in \{1,..,n^{\mathcal{TP}}\}} \frac{1}{r_z^{\mathcal{TP}}} K_{hz} I_{n_{\mathcal{TP}}^D(z)} K_{jz}' = \begin{cases} \frac{1}{r_h^D} I_{n_h^D} & \text{if } D_h = D_j \\ \mathbf{0}_{n_h^D \times n_j^D} & \text{if } D_h \neq D_j \end{cases} \tag{C23}$$

where $r_z^{\mathcal{D}} \equiv |\{t : D_t = D_z\}|$ is the number of periods that the trade pattern $D_z$ is observed.

Finally, from (C21) and (C9), $Z_3 \left(Z_3' Z_3\right)^{-1} Z_3' Z_2 \left(Z_2' Z_2\right)^{-1} Z_2'$ can be decomposed into $n^T \times n^T$ blocks:

$$
\begin{aligned}
T &\equiv Z_3 \left(Z_3' Z_3\right)^{-1} Z_3' Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' \\
&= \begin{bmatrix}
B_1 \left(Z_3' Z_3\right)^{-1} B_1' \frac{1}{n_1^D} \iota_{n_1^D} \iota_{n_1^D}' & \cdots & B_1 \left(Z_3' Z_3\right)^{-1} B_{n^T}' \frac{1}{n_{n^T}^D} \iota_{n_{n^T}^D} \iota_{n_{n^T}^D}' \\
\vdots & \ddots & \vdots \\
B_{n^T} \left(Z_3' Z_3\right)^{-1} B_1' \frac{1}{n_1^D} \iota_{n_1^D} \iota_{n_1^D}' & \cdots & B_{n^T} \left(Z_3' Z_3\right)^{-1} B_{n^T}' \frac{1}{n_{n^T}^D} \iota_{n_{n^T}^D} \iota_{n_{n^T}^D}'
\end{bmatrix}
\end{aligned}
\tag{C24}
$$

where block $(x, y)$ of $T$ is given by

$$
T(x, y) = B_x \left(Z_3' Z_3\right)^{-1} B_y' \frac{1}{n_y^D} \iota_{n_y^D} \iota_{n_y^D}'.
\tag{C25}
$$

From (C23), it is straightforward to see that $T(x, y) = T(y, x)'$. That is, if the trade pattern of period $x$ is the same as that of period $y$, then $T(x, y) = T(y, x)' = \frac{1}{r_x^{\mathcal{D}} n_x^D} \iota_{n_x^D} \iota_{n_x^D}' = \frac{1}{r_y^{\mathcal{D}} n_y^D} \iota_{n_y^D} \iota_{n_y^D}'$; if the trade pattern of period $x$ is different from that of period $y$, then $T(x, y) = T(y, x)' = \mathbf{0}_{n_x^D \times n_y^D}$.

Now, given that $Z_3 \left(Z_3' Z_3\right)^{-1} Z_3'$, $Z_2 \left(Z_2' Z_2\right)^{-1} Z_2'$, and $T$ are all symmetric, it follows that

$$
T = Z_3 \left(Z_3' Z_3\right)^{-1} Z_3' Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' = T' = Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' Z_3 \left(Z_3' Z_3\right)^{-1} Z_3'.
\tag{C26}
$$

$\square$

**Proof of** (C19):

*Proof.* From (C21) and the definition of $Z_1$ in (C3), we can write $T_3 Z_1$ as

$$
T_3 Z_1 = \begin{bmatrix}
\sum_t B_1 \left(Z_3' Z_3\right)^{-1} B_t' A_t \\
\vdots \\
\sum_t B_{n^T} \left(Z_3' Z_3\right)^{-1} B_t' A_t
\end{bmatrix}.
\tag{C27}
$$

Using (C23), we have

$$
B_x \left(Z_3' Z_3\right)^{-1} B_y' A_y = \begin{cases}
\frac{1}{r_x^{\mathcal{D}}} A_x = \frac{1}{r_y^{\mathcal{D}}} A_y & \text{if } D_x = D_y \\
\mathbf{0}_{n_x^D \times n^D} & \text{if } D_x \neq D_y
\end{cases}
\tag{C28}
$$

19

With (C28), it follows that

$$T_3 Z_1 = \begin{bmatrix} \sum_{t:D_t=D_1} \frac{1}{r_1^{\mathcal{D}}} A_1 \\ \vdots \\ \sum_{t:D_t=D_{n^T}} \frac{1}{r_{n^T}^{\mathcal{D}}} A_{n^T} \end{bmatrix} = \begin{bmatrix} A_1 \\ \vdots \\ A_{n^T} \end{bmatrix} = Z_1. \tag{C29}$$

$\square$

### C.1.3 A numerical example with the projection matrices to visualize differences across estimators

To clarify how the estimator works, we now spell out all the key matrices from above discussions and provide a numerical example. We use the data generating process discussed in subsection 3.3.2 of our paper, reproduced below for convenience:[15]

$$p_{dt} = \beta_0 + \beta_1 e_{dt} + \beta_2 m_{dt} \tag{C30}$$

$$p_{dt} = \begin{cases} \text{observed} & \text{if} \quad \gamma_0 + \gamma_1 e_{dt} + \gamma_2 m_{dt} < 0 \\ \text{missing} & \text{if} \quad \gamma_0 + \gamma_1 e_{dt} + \gamma_2 m_{dt} \geq 0 \end{cases} \tag{C31}$$

$$m_{dt} = \psi_d + \epsilon_t + \xi_d * \upsilon_t \tag{C32}$$

$$e_{dt} = \sigma_e (m_{dt} + u_{dt}) \tag{C33}$$

We reduce the number of destinations to 5 and the number of years to 4 to keep the size of the matrices tractable. To keep the example clean, we only allow for two distinct values of the factors affecting the time variation of the unobserved marginal cost (i.e., $\epsilon_t$ and $\upsilon_t$). We set $\gamma_0$ such that half of the observations (destination-year pairs) are dropped.

Table C1 shows one particular realization of such a data generating process. The firm exports in all four periods, and its decisions generate two unique trade patterns. In the first two years, the firm exports to destinations 2, 4 and 5. In the last two years, the firm exports only to destinations 4 and 5.

---

[15]We restate the simulation setting here: $\psi_d$, $\epsilon_t$, $\xi_t$, $\upsilon_t$ and $u_{dt}$ are simulated from a standard normal distribution. We set $\sigma_e$ to be 0.5 such that the bilateral exchange rate shocks are slightly less volatile than idiosyncratic marginal cost shocks. We set $\beta_1 = \beta_2 = 1$ such that an exchange rate appreciation of the home currency and a positive marginal cost shock increase the border price denominated in the home currency. This also implies a positive omitted variable bias. We set $\gamma_1 = -0.1$ and $\gamma_2 = 1$ such that the selection bias is also positive. The magnitude of $\gamma_1$ is set to be smaller than that of $\gamma_2$ to reflect the fact that the aggregate shocks (such as bilateral exchange rates) is less detrimental for the firm's entry decisions compared to idiosyncratic factors (such as the unobserved marginal cost).

| Year | Destination | Trade Pattern | $p_{d,t}$ | $e_{d,t}$ | $m_{d,t}$ | $\epsilon_t$ | $v_t$ |
|------|-------------|---------------|-----------|-----------|-----------|--------------|-------|
| 1 | 2 | 2_4_5 | -0.072 | 0.155 | -0.227 | 0.843 | 0.277 |
| 1 | 4 | 2_4_5 | 0.178 | -0.092 | 0.270 | 0.843 | 0.277 |
| 1 | 5 | 2_4_5 | -1.138 | -1.252 | 0.114 | 0.843 | 0.277 |
| 2 | 2 | 2_4_5 | 0.455 | 0.682 | -0.227 | 0.843 | 0.277 |
| 2 | 4 | 2_4_5 | 0.636 | 0.366 | 0.270 | 0.843 | 0.277 |
| 2 | 5 | 2_4_5 | 0.068 | -0.046 | 0.114 | 0.843 | 0.277 |
| 3 | 4 | 4_5 | -0.313 | 0.689 | -1.002 | -0.191 | 1.117 |
| 3 | 5 | 4_5 | -0.315 | 0.071 | -0.387 | -0.191 | 1.117 |
| 4 | 4 | 4_5 | -1.099 | -0.097 | -1.002 | -0.191 | 1.117 |
| 4 | 5 | 4_5 | -0.747 | -0.360 | -0.387 | -0.191 | 1.117 |

$Z_1$ is the matrix that contains the destination dummies. To economize on the matrix size, we only create dummies for destinations that are observed, i.e., we do not create dummies for destinations 1 and 3. For example, the first column of $Z_1$ reports the observations in which the firm sells to destination 2. From the matrix, we can see that the firm sells to destination 2 two times. $Z_2$ is the matrix that contains the year dummies. $Z_3$ gives our proposed destination-specific trade pattern dummies. As defined in (C14) and (C15), it is constructed by interacting the destination dummies with the trade pattern dummies. For example, the first three columns represent the dummy structure for the destinations related to the 2_4_5 trade pattern, i.e., $2 - 2\_4\_5$, $4 - 2\_4\_5$ and $5 - 2\_4\_5$. Similarly, the last two columns represent the dummy structure for the destinations related to the 4_5 trade pattern, i.e., $4 - 4\_5$ and $5 - 4\_5$.

$$
Z_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad
Z_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad
Z_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{C34}
$$

From these, we can see clearly that $P_2$ is a destination demean process.

$$P_2 = \begin{bmatrix} 0.67 & -0.33 & -0.33 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.33 & 0.67 & -0.33 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.33 & -0.33 & 0.67 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.67 & -0.33 & -0.33 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.33 & 0.67 & -0.33 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.33 & -0.33 & 0.67 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.50 & -0.50 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.50 & 0.50 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.50 & -0.50 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.50 & 0.50 \end{bmatrix} \quad \text{(C35)}$$

By way of example, for the first observation, $2/3p_{11} - 1/3p_{21} - 1/3p_{31} = p_{11} - \frac{1}{3}(p_{11} + p_{21} + p_{31})$.

As discussed in subsection C.1.1, $Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' Z_1$ follows a trade pattern structure and $\bar{Z}$ suggests an additive relationship between the destination dummies $Z_1$ and the trade pattern dummies $Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' Z_1$.

$$Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' Z_1 = \begin{bmatrix} 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 \\ 0 & 0.50 & 0.50 \\ 0 & 0.50 & 0.50 \\ 0 & 0.50 & 0.50 \\ 0 & 0.50 & 0.50 \end{bmatrix} \quad \bar{Z} = Z_1 - Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' Z_1 = \begin{bmatrix} 0.67 & -0.33 & -0.33 \\ -0.33 & 0.67 & -0.33 \\ -0.33 & -0.33 & 0.67 \\ 0.67 & -0.33 & -0.33 \\ -0.33 & 0.67 & -0.33 \\ -0.33 & -0.33 & 0.67 \\ 0 & 0.50 & -0.50 \\ 0 & -0.50 & 0.50 \\ 0 & 0.50 & -0.50 \\ 0 & -0.50 & 0.50 \end{bmatrix}$$
$$\text{(C36)}$$

As we can see from (C37), the projection $P$ does not follow a particular structure. Therefore, our two-step decomposition $P = P_1 P_2$ discussed in subsection C.1.1 helps to unveil the key economic

mechanisms behind the statistical projection.

$$P = \begin{bmatrix} 0.46 & -0.29 & -0.17 & -0.21 & 0.04 & 0.17 & -0.13 & 0.13 & -0.13 & 0.13 \\ -0.29 & 0.46 & -0.17 & 0.04 & -0.21 & 0.17 & 0.13 & -0.13 & 0.13 & -0.13 \\ -0.17 & -0.17 & 0.33 & 0.17 & 0.17 & -0.33 & 0 & 0 & 0 & 0 \\ -0.21 & 0.04 & 0.17 & 0.46 & -0.29 & -0.17 & -0.13 & 0.13 & -0.13 & 0.13 \\ 0.04 & -0.21 & 0.17 & -0.29 & 0.46 & -0.17 & 0.13 & -0.13 & 0.13 & -0.13 \\ 0.17 & 0.17 & -0.33 & -0.17 & -0.17 & 0.33 & 0 & 0 & 0 & 0 \\ -0.13 & 0.13 & 0 & -0.13 & 0.13 & 0 & 0.38 & -0.38 & -0.13 & 0.13 \\ 0.13 & -0.13 & 0 & 0.13 & -0.13 & 0 & -0.38 & 0.38 & 0.13 & -0.13 \\ -0.13 & 0.13 & 0 & -0.13 & 0.13 & 0 & -0.13 & 0.13 & 0.38 & -0.38 \\ 0.13 & -0.13 & 0 & 0.13 & -0.13 & 0 & 0.13 & -0.13 & -0.38 & 0.38 \end{bmatrix} \tag{C37}$$

Let $Y = [-0.072, 0.178, -1.138, 0.455, 0.636, 0.068, -0.313, -0.315, -1.099, -0.747]'$ and $X = [0.155, -0.092, -1.252, 0.682, 0.366, -0.046, 0.689, 0.071, -0.097, -0.360]'$. The OLS estimator is given by $(X'X)^{-1}X'Y$, which gives an estimate of $\widehat{\beta}_1 = 0.745$. The estimator applying $d$ and $t$ fixed effects is given by $(X'P'PX)^{-1}X'P'Y$, which gives $\widehat{\beta}_1 = 1.508$. The estimator applying $dD$ and $t$ fixed effects is given by $(X'P_2'P_3'P_3P_2X)^{-1}X'P_2'P_3'P_3P_2Y$, which gives the calibrated value of $\widehat{\beta}_1 = 1.000$.

## C.2   Identifying Markup Elasticities in Unbalanced Panels: Adding the Firm Dimension

Our discussions in the previous subsection generalise to more panel dimensions. We now allow for a third, firm (or firm-product) dimension.[16] The data structure can be viewed as a collection of two dimensional settings as previously presented in (C3). For a given firm $f$, let $n_f^D$ denote the total number of export destinations by the firm and $n_{ft}^D \left( n_{ft}^D \leq n_f^D \right)$ be the number of observed destinations in year $t$. Let $n_f^T$ denote the maximum number of exporting years and the $n_f^{DT} \equiv \sum_t n_{ft}^D$ be the number of observed transactions by firm $f$. Let $A_{ft}$ be the $\left( n_{ft}^D \times n_f^D \right)$ matrix obtained from the $(n_f^D \times n_f^D)$ identity matrix from which, for each firm $f$, rows corresponding to destinations not observed in year $t$ have been omitted. Consider

$$Z_f \equiv \begin{pmatrix} Z_{f,1}, & Z_{f,2} \\ n_f^{DT} \times n_f^D & n_f^{DT} \times n_f^T \end{pmatrix} \equiv \begin{bmatrix} A_{f1} & A_{f1}\iota_{n_f^D} & \\ \vdots & & \ddots \\ A_{fn_f^T} & & A_{fn_f^T}\iota_{n_f^D} \end{bmatrix} \tag{C38}$$

---

[16]We focus on the firm rather than the firm and product panel dimensions to simplify our notation, but all of the discussions here trivially generalize to the firm and product dimensions.

where $\iota_j$ is a vector of ones with length $j$. The matrix $Z_f$ gives the dummy-variable structure for firm $f$. Let

$$Z_1 \equiv \begin{bmatrix} Z_{1,1} & & \\ & \ddots & \\ & & Z_{n^F,1} \end{bmatrix}, \quad Z_2 \equiv \begin{bmatrix} Z_{1,2} & & \\ & \ddots & \\ & & Z_{n^F,2} \end{bmatrix} \tag{C39}$$

$$Z \equiv \begin{pmatrix} Z_1, & Z_2 \\ n^{FDT} \times n^{FD} & n^{FDT} \times n^{FT} \end{pmatrix} \tag{C40}$$

where $n^{FD} \equiv \sum_f n_f^D$ and $n^{FT} \equiv \sum_f n_f^T$. Define

$$P_2 \equiv I_{n^{FDT}} - Z_2 \left( Z_2' Z_2 \right)^- Z_2' \tag{C41}$$

$$\bar{Z} \equiv P_2 Z_1. \tag{C42}$$

Using the properties of block diagonal matrices, it is straightforward to show that $P$ is the projection matrix onto the null-space of $Z$

$$P \equiv P_2 - \bar{Z}(\bar{Z}'\bar{Z})^- \bar{Z}'. \tag{C43}$$

In a panel with firm (or firm-product) panel dimensions, the following two corollaries corresponding to propositions 1 and 2 hold:

**Corollary 1.** *In a three dimensional panel, factors varying at the $fd + ft$ panel dimensions can be eliminated using a two-step procedure in which, in the first step, all variables are demeaned across observed destinations within each firm-time pair and, in the second step, firm-destination ($fd$) and firm-trade pattern ($fD$) fixed effects are applied additively, i.e., $fd + fD$.*

**Corollary 2.** *In a three dimensional panel, factors varying at the $fdD + ft$ panel dimensions can be eliminated using a two-step procedure in which all variables are demeaned across observed destinations within each firm-time pair in the first stage and firm-destination specific trade patterns ($fdD$) are applied in the second stage. This procedure also eliminates all confounding factors that the $fd + ft$ fixed effects can address.*

### C.2.1 Proof of Corollary 1

As in the two dimensional panel case (see (C7)), the projection matrix $P$ can be decomposed into two projection matrices, which can be applied sequentially as a two-step estimator.

$$P \equiv P_2 - \bar{Z}(\bar{Z}'\bar{Z})^- \bar{Z}' = (I_{n^{FDT}} - \bar{Z}(\bar{Z}'\bar{Z})^- \bar{Z}')P_2 \equiv P_1 P_2 \tag{C44}$$

where $P_1 \equiv I_{n^{FDT}} - \bar{Z}(\bar{Z}'\bar{Z})^- \bar{Z}'$ and the second projection matrix $P_1$ implicitly applies trade pattern fixed effects as

$$P_2 Z_1 = Z_1 - Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' Z_1. \tag{C45}$$

To see that $Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' Z_1$ follows a trade pattern structure, note that the following relationships hold:

$$Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' Z_1 = \begin{bmatrix} Z_{1,2} \left(Z_{1,2}' Z_{1,2}\right)^{-1} Z_{1,2}' Z_{1,1} & & \\ & \ddots & \\ & & Z_{n^F,2} \left(Z_{n^F,2}' Z_{n^F,2}\right)^{-1} Z_{n^F,2}' Z_{n^F,1} \end{bmatrix} \tag{C46}$$

$$Z_{f,2} \left(Z_{f,2}' Z_{f,2}\right)^{-1} Z_{f,2}' Z_{f,1} = \begin{bmatrix} \frac{1}{n_{f1}^D} \iota_{n_{f1}^D} \iota_{n_{f1}^D}' A_{f1} \\ \vdots \\ \frac{1}{n_{fn_f^T}^D} \iota_{n_{fn_f^T}^D} \iota_{n_{fn_f^T}^D}' A_{fn_f^D} \end{bmatrix} \tag{C47}$$

As in the two panel dimension case, $Z_{f,2} \left(Z_{f,2}' Z_{f,2}\right)^{-1} Z_{f,2}' Z_{f,1}$ follows a trade pattern structure.

### C.2.2 Proof of Corollary 2

The proposed method eliminates potential confounding fixed effects in the conventional fixed effect model. Let $Z_{f,3}$ be the destination-specific trade pattern dummies of firm $f$ defined as in (C14) and (C15), and $Z_3$ be the matrix that represents the firm-destination specific trade pattern dummies which can be constructed as a block diagonal matrix of $Z_{f,3}$'s:

$$Z_3 \equiv \begin{bmatrix} Z_{1,3} & & \\ & \ddots & \\ & & Z_{n^F,3} \end{bmatrix} \tag{C48}$$

Using the properties of block diagonal matrices and firm level relationships (C18) and (C19), it is straightforward to show the following relationship holds:

$$Z_3 \left(Z_3' Z_3\right)^- Z_3' Z_1 = \begin{bmatrix} Z_{1,1} & & \\ & \ddots & \\ & & Z_{n^F,1} \end{bmatrix} \tag{C49}$$

$$Z_3 \left(Z_3' Z_3\right)^- Z_3' Z_2 \left(Z_2' Z_2\right)^{-1} Z_2' Z_1 = \begin{bmatrix} Z_{1,2} \left(Z_{1,2}' Z_{1,2}\right)^{-1} Z_{1,2}' Z_{1,1} & & \\ & \ddots & \\ & & Z_{n^F,2} \left(Z_{n^F,2}' Z_{n^F,2}\right)^{-1} Z_{n^F,2}' Z_{n^F,1} \end{bmatrix}.$$

(C50)

It follows that

$$[I_{n^{FDT}} - Z_3 \left(Z_3' Z_3\right)^- Z_3'][I_{n^{FDT}} - Z_2 \left(Z_2' Z_2\right)^{-1} Z_2'] Z_1 = \mathbf{0}. \tag{C51}$$

Similarly, it is straightforward to show that $[I_{n^{FDT}} - Z_3 \left(Z_3' Z_3\right)^- Z_3'][I_{n^{FDT}} - Z_2 \left(Z_2' Z_2\right)^{-1} Z_2'] Z_2 = \mathbf{0}$ and $[I_{n^{FDT}} - Z_3 \left(Z_3' Z_3\right)^- Z_3'][I_{n^{FDT}} - Z_2 \left(Z_2' Z_2\right)^{-1} Z_2'] Z_3 = \mathbf{0}$.

# D  Estimating Markup Elasticities with Heterogeneous Responses

In this subsection, we extend the results in sections B and C and discuss the estimated object captured by the OLS and fixed effect approaches when the markup elasticity is different across firms, products, destinations and time. We highlight two interrelated issues. The first issue arises because linear OLS or fixed effect estimators treat the heterogeneous coefficients as if they were homogeneous. The second issue which arises due to inaccurate first-order log approximations of nonlinear theoretical relationships.

## D.1  The implicit weight of observations

To introduce this issue, consider the following simple specification:

$$p_{fidt} = \beta_{fidt} e_{dt} + v_{fidt} \tag{D1}$$

where $p_{fidt}$ is the log price and $e_{dt}$ is the log bilateral exchange rate and $\beta_{fidt}$ as the markup elasticity, which is a function of the bilateral exchange rate as shown in (7); $v_{fidt}$ is an iid error.

The OLS estimate of $\beta$ is given by

$$
\begin{aligned}
\beta^{OLS} &= \frac{\sum_f \sum_i \sum_d \sum_t (p_{fidt} - \overline{p})(e_{dt} - \overline{e})}{\sum_f \sum_i \sum_d \sum_t (e_{dt} - \overline{e})^2} \\
&= \frac{\sum_f \sum_i \sum_d \sum_t (\beta_{fidt} e_{dt} + v_{fidt})(e_{dt} - \overline{e})}{\sum_f \sum_i \sum_d \sum_t (e_{d,t} - \overline{e})^2} \\
&= \frac{1}{n^F n^I} \sum_f \sum_i \left[ \frac{\sum_d \sum_t (\beta_{fidt} e_{dt} + v_{fidt})(e_{dt} - \overline{e})}{\sum_d \sum_t (e_{dt} - \overline{e})^2} \right] \\
&= \frac{1}{n^F n^I} \sum_f \sum_i \left[ \frac{\sum_d \sum_t \beta_{fidt} e_{dt}(e_{dt} - \overline{e})}{\sum_d \sum_t (e_{dt} - \overline{e})^2} + \frac{\sum_d \sum_t v_{fidt}(e_{dt} - \overline{e})}{\sum_d \sum_t (e_{dt} - \overline{e})^2} \right] \\
&= \frac{1}{n^F n^I} \sum_f \sum_i \left[ \sum_d \sum_t \beta_{fidt} w_{dt} + \frac{\sum_d \sum_t v_{fidt}(e_{dt} - \overline{e})}{\sum_d \sum_t (e_{dt} - \overline{e})^2} \right] \quad (D2)
\end{aligned}
$$

where $n^F, n^I, n^D, n^T$ represent the number of firms, products, destinations, and time periods respectively; $w_{dt} \equiv \frac{e_{dt}(e_{dt} - \overline{e})}{\sum_d \sum_t (e_{dt} - \overline{e})^2}$; $\overline{p} \equiv \frac{1}{n^F n^I n^D n^T} \sum_f \sum_i \sum_d \sum_t p_{fidt}$; $\overline{e} \equiv \frac{1}{n^D n^T} \sum_d \sum_t e_{dt}$.[17] Now, the second term in the bracket of (D2) is close to 0 under the assumption of no selection and omitted variable bias. Let's abstract from these biases, and focus on the first term in the bracket.

From this term in (D2), it is apparent that, when the markup elasticity is heterogeneous, $\beta^{OLS}$ is the exchange rate-deviation weighted sum of the $\beta_{fidt}$'s.

$$
\beta^{OLS} \approx \frac{1}{n^F n^I} \sum_f \sum_i \sum_d \sum_t \beta_{fidt} w_{dt} \neq \frac{1}{n^F n^I n^D n^T} \sum_f \sum_i \sum_d \sum_t \beta_{fidt} \quad (D3)
$$

As can be seen from the definition of $w_{dt}$, the OLS estimator gives a larger weight to high exchange rate values, that is, foreign currency appreciations. The result is different, for instance, from an observation weighted average of the $\beta_{fidt}$'s.

In general, with multiple regressors, the weights of an OLS estimator also depend on the variation of other independent variables and the coefficients in front of these variables. The OLS estimates capture

$$
\boldsymbol{\beta}^{OLS} \approx (X'X)^{-1} X'(X \circ B) \quad (D4)
$$

where $X$ is a $n^{FIDT} \times k$ matrix that stores the values of the $k$ independent variables; $B$ is a $n^{FIDT} \times k$ matrix that stores the heterogeneous coefficients for each of the independent variables; $\circ$ is the Hadamard (element-by-element) product. Similarly, the estimates of a FE estimator and

---

[17]We have assumed a balanced panel in the discussions of (D2) and (D3) for clarity. We discuss the general case in (D4), (D5) and (D6).

our TPSFE estimator captures

$$\boldsymbol{\beta}^{FE} \approx (X'P'PX)^{-1}X'P'(PX \circ B) \tag{D5}$$

$$\boldsymbol{\beta}^{TPSFE} \approx (X'P_2'P_3'P_3P_2X)^{-1}X'P_2'P_3'(P_3P_2X \circ B) \tag{D6}$$

where $P$ is the projection matrix that is required to perform the conventional FE estimator; $P_2$ represents a projection matrix that performs a destination demeaning operation and $P_3$ represents the second demeaning step of the TPSFE estimator at the firm-product-destination-trade pattern level.

For our purposes, there are at least two relevant takeaways from (D4), (D5) and (D6). First, even without the omitted variable and selection biases, the estimated coefficients from the three estimators can differ slightly due to the different weighting matrices applied to the coefficient matrix $B$. Second, in general, the estimated coefficients for all of the three estimators do not necessarily equal the unweighed average of the coefficients, i.e., $\frac{1}{n^{FIDT}}\iota'_{n^{FIDT}}B$, or any other average that an econometrician may take as the reference benchmark to assess estimation biases. When elasticities are heterogeneous, the assessment of the performance of an estimator may vary with the choice of the benchmark.

## D.2   Approximation bias

The second issue arises when non-linear relationships are approximated using log-linear equations. In the open marco literature, first order log approximations are widely used to derive theoretical relationships between variables. For example, in equation (6), we have shown a log linearised equation of the markup response as

$$\widehat{\Gamma}_{fiodt} = \frac{\omega_{fiodt}}{\rho_i + \omega_{fiodt}} \left( \widehat{\mathcal{E}}_{odt} - \widehat{\mathcal{MC}}_{fiodt} + \widehat{P}^N_{dt} \right) + \frac{\rho_i}{\rho_i + \omega_{fiodt}} \widehat{\tau}_{odt}.$$

Log-linearisation is obviously convenient here, as the coefficient in front of the exchange rate changes, $\frac{\omega_{fiodt}}{\rho_i + \omega_{fiodt}}$, directly gives the key parameter of interest, the markup elasticity to exchange rates. However, as is well known, estimating the relationship using logged variables can lead to a non-trivial bias even if all variables are directly observable and there is no selection bias. Concretely, when we regress the logged markup on the logged exchange rate, the logged marginal cost, the logged non-tradable price and the logged trade cost, we do not get the average of the markup elasticity to exchange rates *even after accounting for the weighting issue* analyzed in the previous subsection. A specific bias arises due to the fact that the high order terms of the approximation $O_{fiodt}$ are

correlated with the variables in the estimation equation (i.e., $\ln(\mathcal{E}_{odt}), \ln(\mathcal{MC}_{fiodt}), \ln(P_{dt}^N), \ln(\tau_{odt}))$.

$$\ln(\Gamma_{fiodt}) = \Gamma_{fiodt}^{Approx} + O_{fiodt} \tag{D7}$$

$$\Gamma_{fiodt}^{Approx} \equiv \frac{\omega_{fiodt}}{\rho_i + \omega_{fiodt}}\left[\ln(\mathcal{E}_{odt}) - \ln(\mathcal{MC}_{fiodt}) + \ln(P_{dt}^N)\right] + \frac{\rho_i}{\rho_i + \omega_{fiodt}}\ln(\tau_{odt}) \tag{D8}$$

To be clear, if we could estimate equation (D8) directly by regressing $\Gamma_{fiodt}^{Approx}$ on $\ln(\mathcal{E}_{odt}), \ln(\mathcal{MC}_{fiodt})$, $\ln(P_{dt}^N)$ and $\ln(\tau_{odt})$, then the estimates would only reflect the weighing problem discussed in the previous subsection—the estimated coefficients would be consistent with the formulas described by (D4), (D5) and (D6). However, the literature usually estimates markup elasticity by regressing $\ln(\Gamma_{fiodt})$ on the logged independent variables $(\ln(\mathcal{E}_{odt}), \ln(\mathcal{MC}_{fiodt}), \ln(P_{dt}^N)$ and $\ln(\tau_{odt}))$. The estimates are bound to suffer from an approximation bias, as the higher order terms are in general correlated with the first order terms.

Notably, in all simulations of the CD model, the weighting issue and the approximation biases always go in opposite directions and partially offset each other. If we take the unweighed mean of the theoretical markup elasticities as a reference benchmark, the difference between this and our estimated markup elasticities to exchange rates remains reasonably small after splitting the sample into high and low differentiation goods.

# E    The TPSFE Estimator Relative to De Loecker, Goldberg, Khandelwal and Pavcnik (2016)

In this subsection, we extend the framework of De Loecker, Goldberg, Khandelwal and Pavcnik (2016) to add a destination dimension, and discuss the structural assumptions that would be required for our main identification condition (B20) to be satisfied in this new framework.

## E.1    Structural Interpretation of Assumptions Required by Our Estimator

We start by writing the production function as follows:

$$Q_{fidt} = F_{fi}(\boldsymbol{V}_{fidt}, \boldsymbol{K}_{fidt})\Omega_{fit}\Psi_{fid} \tag{E1}$$

where $Q_{fidt}$ represents the quantity of exports for product $i$ from firm $f$ to destination $d$ at time $t$; $\boldsymbol{V}_{fidt}$ denotes a vector of variable inputs, $\{V_{fidt}^1, V_{fidt}^2, ..., V_{fidt}^v\}$; $\boldsymbol{K}_{fidt}$ denotes a vector of dynamic inputs; a firm-product pair make decisions on allocating its dynamic inputs across destinations in each time period, $\{K_{fidt}^1, K_{fidt}^2, ..., K_{fidt}^k\}$. We stress that the above function allows

for destination-specific inputs $\{\boldsymbol{V}_{fidt}, \boldsymbol{K}_{fidt}\}$ as well as destination-specific productivity differences, $\Psi_{fid}$, at the firm and product level. In addition, we allow for the production function and Hicks-neutral productivity to be firm-product specific.

Specifically, we posit the following:

1. The production technology is firm-product-specific.

2. $F_{fi}(.)$ is continuous and twice differentiable w.r.t. at least one element of $V_{fidt}$, and this element of $V_{fidt}$ is a static (i.e., freely adjustable or variable) input in the production of product $i$.

3. $F_{fi}(.)$ is constant return to scale.

4. Hicks-neutral productivity $\Omega_{fit}$ is log-additive.

5. The destination specific technology advantage $\Psi_{fid}$ takes a log-additive form and is not time varying.

6. Input prices $\boldsymbol{W}_{fit}$ are firm-product-time specific.

7. The state variables of the firm are

$$\boldsymbol{s}_{fit} = \{D_{fit}, \boldsymbol{K}_{fit}, \Omega_{fit}, \Psi_{fid}, \boldsymbol{G}_{fi}, \boldsymbol{r}_{fidt}\} \tag{E2}$$

where $\boldsymbol{G}_{fi}$ includes variables indicating firm and product properties, e.g., firm registration types, product differentiation indicators. $\boldsymbol{r}_{fidt}$ collects other observables including variables that track the destination market conditions, such as the bilateral exchange rate and destination CPI.

8. Firms minimize short-run costs taking output quantity, $Q_{fidt}$, and input prices, $\boldsymbol{W}_{fit}$, at time $t$ as given.

The assumptions 1, 2, 4, 8 are standard in the literature. De Loecker, Goldberg, Khandelwal and Pavcnik (2016) also posit them, but in our version we allow the production function to be firm specific and the Hicks-neutral productivity to be product-specific. Compared to the conditions assumed in the literature, assumption 5 is a relaxation: it allows for the possibility that (log-additive) productivity be destination-specific.

Assumptions 6 and 7 allow prices of inputs to be firm and product specific. These two conditions indicate that firms source inputs at the product level, and then allocate these inputs into production for different destinations. Note that the firm can arrange different quantities of inputs and have different marginal costs across destinations for the same product.

The assumption that is crucial to our identification is that the production technology is constant returns to scale (condition 3). This condition implies that the marginal cost at the firm-product-destination level does not depend on the quantity produced. If changes in relative demand and exports across destinations were systematically associated to changes in relative marginal costs, condition (B20) would be violated. As discussed in the next subsection, looking at the solution to the firms' cost minimization problem, condition 3 ensures that the difference in the marginal costs across destinations only reflects technology differences varying at the destination dimension.

## E.2   The cost minimization problem by firm-product pair

Write the cost function

$$
\mathcal{L}(\boldsymbol{V}_{fidt}, \boldsymbol{K}_{fidt}, \lambda_{fidt}) = \sum_{v=1}^{V} W_{fit}^v \sum_{d \in D_{fit}} V_{fidt}^v + \sum_{k=1}^{K} R_{fit}^k \left( \sum_{d \in D_{fit}} K_{fidt}^k - K_{fit}^k \right)
$$
$$
+ \sum_{d \in D_{fit}} \lambda_{fidt} [Q_{fidt} - F_{fi}(\boldsymbol{V}_{fidt}, \boldsymbol{K}_{fidt}) \Omega_{fit} \Psi_{fid}]
$$

where $K_{fit}^k$ is the accumulated capital input $k$ in the previous period; $K_{fidt}^k$ stands for the corresponding allocation for destination $d$; $R_{fit}^k$ is the implied cost of capital.[18]

The F.O.C.s of the cost minimization problem are

$$
\frac{\partial \mathcal{L}_{fit}}{\partial V_{fidt}^v} = W_{fit}^v - \lambda_{fidt} \Omega_{fit} \Psi_{fid} \frac{\partial F_{fi}(.)}{\partial V_{fidt}^v} = 0, \tag{E3}
$$

$$
\frac{\partial \mathcal{L}_{fit}}{\partial K_{fidt}^k} = R_{fit}^k - \lambda_{fidt} \Omega_{fit} \Psi_{fid} \frac{\partial F_{fi}(.)}{\partial K_{fidt}^k} = 0. \tag{E4}
$$

Conditions (E3) and (E4) need to hold across inputs and across destinations, which implies the following:

$$
\frac{W_{fit}^1}{W_{fit}^v} = \frac{\frac{\partial F_{fi}(.)}{\partial V_{fi1t}^1}}{\frac{\partial F_{fi}(.)}{\partial V_{fi1t}^v}} = \frac{\frac{\partial F_{fi}(.)}{\partial V_{fi2t}^1}}{\frac{\partial F_{fi}(.)}{\partial V_{fi2t}^v}} = ... = \frac{\frac{\partial F_{fi}(.)}{\partial V_{fidt}^1}}{\frac{\partial F_{fi}(.)}{\partial V_{fidt}^v}} \qquad \forall v = 1, ..., V; \quad d \in D_{fit}, \tag{E5}
$$

$$
\frac{W_{fit}^v}{R_{fit}^k} = \frac{\frac{\partial F_{fi}(.)}{\partial V_{f,i,1,t}^v}}{\frac{\partial F_{fi}(.)}{\partial K_{fi1t}^k}} = \frac{\frac{\partial F_{fi}(.)}{\partial V_{fi2t}^v}}{\frac{\partial F_{fi}(.)}{\partial K_{fi2t}^k}} = ... = \frac{\frac{\partial F_{fi}(.)}{\partial V_{fidt}^v}}{\frac{\partial F_{fi}(.)}{\partial K_{fidt}^k}} \qquad \forall v, k; \quad d \in D_{fit}. \tag{E6}
$$

Note that the production function is assumed to be firm-product specific and constant return to scale. Together with equations (E5) and (E6), these assumptions imply that the allocation of

---

[18]The assumption that the production function $F_{fi}(.)$ is firm-product-specific ensures the implied cost of capital $R_{fit}^k$ is not destination-specific.

variable inputs is inversely proportional to the ratio of the productivity deflated outputs across destinations, i.e.,

$$\frac{Q_{fidt}}{\Omega_{fit}\Psi_{fid}} = c \cdot \frac{Q_{fid't}}{\Omega_{fit}\Psi_{fid'}} \quad \rightarrow \quad c\boldsymbol{V}^*_{fidt} = \boldsymbol{V}^*_{fid't} \quad \text{and} \quad c\boldsymbol{K}^*_{fidt} = \boldsymbol{K}^*_{fid't}. \tag{E7}$$

Utilizing the relationship of (E7) and the assumption that $F_{fi}(.)$ is constant return to scale, it is straightforward to see

$$\frac{\partial F_{fi}(\boldsymbol{V}^*_{fidt}, \boldsymbol{K}^*_{fidt})}{\partial V^v_{fidt}} = \frac{\partial F_{fi}(c\boldsymbol{V}^*_{fidt}, c\boldsymbol{K}^*_{fidt})}{\partial (cV^v_{fidt})} = \frac{\partial F_{fi}(\boldsymbol{V}^*_{fid't}, \boldsymbol{K}^*_{fid't})}{\partial V^v_{fid't}}. \tag{E8}$$

Rearranging (E3) and (E8) yields:

$$\begin{aligned} \lambda_{fidt} &= \left( \frac{\Omega_{fit}\Psi_{fid}}{W^v_{fit}} \frac{\partial F_{fi}(\boldsymbol{V}^*_{fidt}, \boldsymbol{K}^*_{fidt})}{\partial V^v_{fidt}} \right)^{-1} \\ &= \left( \frac{\Omega_{fit}\Psi_{fid}}{W^v_{fit}} \frac{\partial F_{fi}(\boldsymbol{V}^*_{fid't}, \boldsymbol{K}^*_{fid't})}{\partial V^v_{fid't}} \right)^{-1}. \end{aligned} \tag{E9}$$

Therefore, the relative marginal cost across destinations is static, depending on the relative productivity difference across destinations, i.e.,

$$\frac{\lambda_{fidt}}{\lambda_{fid't}} = \frac{\Psi_{fid'}}{\Psi_{fid}} \tag{E10}$$

Although the marginal cost is firm-product-destination specific and time-varying, the relative marginal cost is not. Therefore, condition (B20) is satisfied.

## E.3   An alternative approach

An alternative approach to reconcile our work with De Loecker, Goldberg, Khandelwal and Pavcnik (2016) consists of directly redefining what a product variety is in their model. Namely, if one redefines a product-destination pair as a variety, i.e., $j = \{i, d\}$, then the original setting and assumptions will go through without any change.

We argue that this approach is not very useful, for two reasons. The first one is practical. De Loecker, Goldberg, Khandelwal and Pavcnik (2016) define a product variety as a two-digit industry. The need to define a product at the industry level is mainly due to data limitations. If one adopts a more refined product definition, for instance, the estimator by De Loecker, Goldberg, Khandelwal and Pavcnik (2016) would suffer from a small sample problem—there would not be enough power to estimate. The small sample problem will be much more severe if one defines a

product-destination pair as a variety. This is due not only to the smaller number of observations in each cell, but also to the frequent changes in the set of destinations a firm exports a product to.

The second reason is related to conceptual assumptions regarding production functions. De Loecker, Goldberg, Khandelwal and Pavcnik (2016) rely on the assumption that the production function is the same for single- and multi-product firms. When redefining a product-destination pair as a variety, the identification condition would require the production function to be product-destination specific and invariant along the firm dimension. In the context of our problem, controlling for firm-product level marginal cost is the primary concern. We think that keeping the flexibility of the production function at the product level is extremely valuable.

# F  General Model-Free Relationships

## F.1  The separation of the marginal cost and the markup

We start by deriving a general expression of a firm's profit-maximizing price. Please note that variables in the following derivation are in levels rather than logarithms. Write:

$$\max_p q(p, \xi)p - c[q(p, \xi), \zeta]. \tag{F1}$$

The firm takes its demand function, $q(p, \xi)$, and cost function, $c[q(p, \xi), \zeta]$, as given and maximises its profit by choosing its optimal price $p$. $\xi$ and $\zeta$ are exogenous demand and supply function shifters respectively.

The first order condition of the firm is given by

$$\frac{\partial q(p, \xi)}{\partial p}p + q(p, \xi) = \frac{\partial c[q(p, \xi), \zeta]}{\partial q(p, \xi)}\frac{\partial q(p, \xi)}{\partial p} \tag{F2}$$

From this equation, we can derive the optimal price as

$$p^* = \frac{\vartheta(p^*, \xi)}{\vartheta(p^*, \xi) - 1}mc[q(p^*, \xi), \zeta]. \tag{F3}$$

where $\vartheta(p, \xi) \equiv -\frac{\partial q(p, \xi)}{\partial p}\frac{p}{q(p, \xi)}$, $mc[q(p, \xi), \zeta] \equiv \frac{\partial c[q(p, \xi), \zeta]}{\partial q(p, \xi)}$.

## F.2 The equilibrium relationship between quantity and price under pure supply versus demand shocks

**Proposition 3.** *If changes in the equilibrium price and quantity are solely driven by shocks to the supply side, the following expression holds*

$$\frac{d\log(q^*)}{d\log(p^*)} = -\vartheta(p^*, \xi) \tag{F4}$$

*Proof.*

$$d\log(q(p^*(\xi,\zeta),\xi)) = \frac{1}{q(p^*(\xi,\zeta),\xi)} dq(p^*(\xi,\zeta),\xi)$$

$$= \frac{1}{q(p^*(\xi,\zeta),\xi)} \left( \frac{\partial q(p^*(\xi,\zeta),\xi)}{\partial p^*(\xi,\zeta)} dp^*(\xi,\zeta) + \frac{\partial q(p^*(\xi,\zeta),\xi)}{\partial \xi} d\xi \right) \tag{F5}$$

$$d\log(p^*(\xi,\zeta)) = \frac{1}{p^*(\xi,\zeta)} dp^*(\xi,\zeta) \tag{F6}$$

Substituting equation (F6) into (F5) and applying the condition $d\xi = 0$ completes the proof. $\square$

**Proposition 4.** *If changes in the equilibrium price and quantity are solely driven by shocks to the demand side, the following expression holds*

$$\frac{d\log(q^*)}{d\log(p^*)} = \frac{\varphi_q(p^*, \xi)}{\varphi_p(\xi, \zeta)} - \vartheta(p^*, \xi) \tag{F7}$$

where $\varphi_q(p^*, \xi) \equiv \frac{\partial q(p^*, \xi)}{\partial \xi} \frac{\xi}{q(p^*, \xi)}$ and $\varphi_p(\xi, \zeta) \equiv \frac{\partial p^*(\xi, \zeta)}{\partial \xi} \frac{\xi}{p^*(\xi, \zeta)}$

*Proof.*

$$d\log(q(p^*(\xi,\zeta),\xi)) = \frac{1}{q(p^*(\xi,\zeta),\xi)} \left( \frac{\partial q(p^*(\xi,\zeta),\xi)}{\partial \xi} d\xi + \frac{\partial q(p^*(\xi,\zeta),\xi)}{\partial p^*(\xi,\zeta)} dp^*(\xi,\zeta) \right)$$

$$= (\varphi_q(p^*, \xi) - \vartheta(p^*, \xi)\varphi_p(\xi, \zeta)) \frac{d\xi}{\xi} \tag{F8}$$

$$d\log(p^*(\xi,\zeta)) = \frac{1}{p^*(\xi,\zeta)} dp^*(\xi,\zeta)$$

$$= \frac{1}{p^*(\xi,\zeta)} \left( \frac{\partial p^*(\xi,\zeta)}{\partial \xi} d\xi \right)$$

$$= \varphi_p(\xi,\zeta) \frac{d\xi}{\xi} \tag{F9}$$

$\square$

# G Data

## G.1 Chinese Customs Data

China's export growth exploded over 2000-2014 (see table G1). Statistics from customs data on firms, HS08 products, and firm-products highlight the growth at the extensive margin, including both net entry of firms, and net entry of firm-products. The total number of active exporters almost quintupled over our sample period, from 62,746 in 2000 to 295,309 in 2014. The number of annual transactions at the firm-HS08 product level increased at roughly the same pace as the number of exporters, from about 904 thousand in 2000 to 4.56 million in 2014. The value of total exports measured in dollars increased ten-fold from 2000 to 2014.

Table G1: Chinese exports: firms, products and values, 2000-2014

|  | HS08 Products | Firms | Firm-HS08 Product Pairs | Observations | Value (billions US$) |
|---|---|---|---|---|---|
| 2000 | 6,712 | 62,746 | 904,111 | 1,953,638 | 249 |
| 2001 | 6,722 | 68,487 | 991,015 | 2,197,705 | 291 |
| 2002 | 6,892 | 78,607 | 1,195,324 | 2,672,837 | 325 |
| 2003 | 7,013 | 95,683 | 1,475,588 | 3,328,320 | 438 |
| 2004 | 7,017 | 120,567 | 1,826,966 | 4,125,819 | 593 |
| 2005 | 7,125 | 142,413 | 2,277,801 | 5,252,820 | 753 |
| 2006 | 7,171 | 171,169 | 2,907,975 | 6,312,897 | 967 |
| 2007 | 7,172 | 193,567 | 3,296,238 | 7,519,615 | 1,220 |
| 2008 | 7,213 | 206,529 | 3,244,484 | 7,995,266 | 1,431 |
| 2009 | 7,322 | 216,219 | 3,363,610 | 8,263,509 | 1,202 |
| 2010 | 7,363 | 234,366 | 3,847,708 | 9,913,754 | 1,577 |
| 2011 | 7,404 | 254,617 | 4,153,534 | 10,645,699 | 1,898 |
| 2012 | 7,564 | 266,842 | 4,171,770 | 11,057,899 | 2,016 |
| 2013 | 7,579 | 279,428 | 4,140,897 | 11,643,683 | 2,176 |
| 2014 | 7,641 | 295,309 | 4,555,912 | 12,297,195 | 2,310 |
| 2000-2014 | 10,002 | 581,141 | 22,820,644 | 108,465,375 | 17,453 |

## G.2 The evolution of exports by private, state-owned and foreign-invested firms in China

In figure G1, we lay out some basic facts about the evolution of different types of firms among Chinese exporters. In the Chinese Customs Database, firms report their registration type in one of the following eight categories: state-owned enterprise, Sino-foreign contractual joint venture,
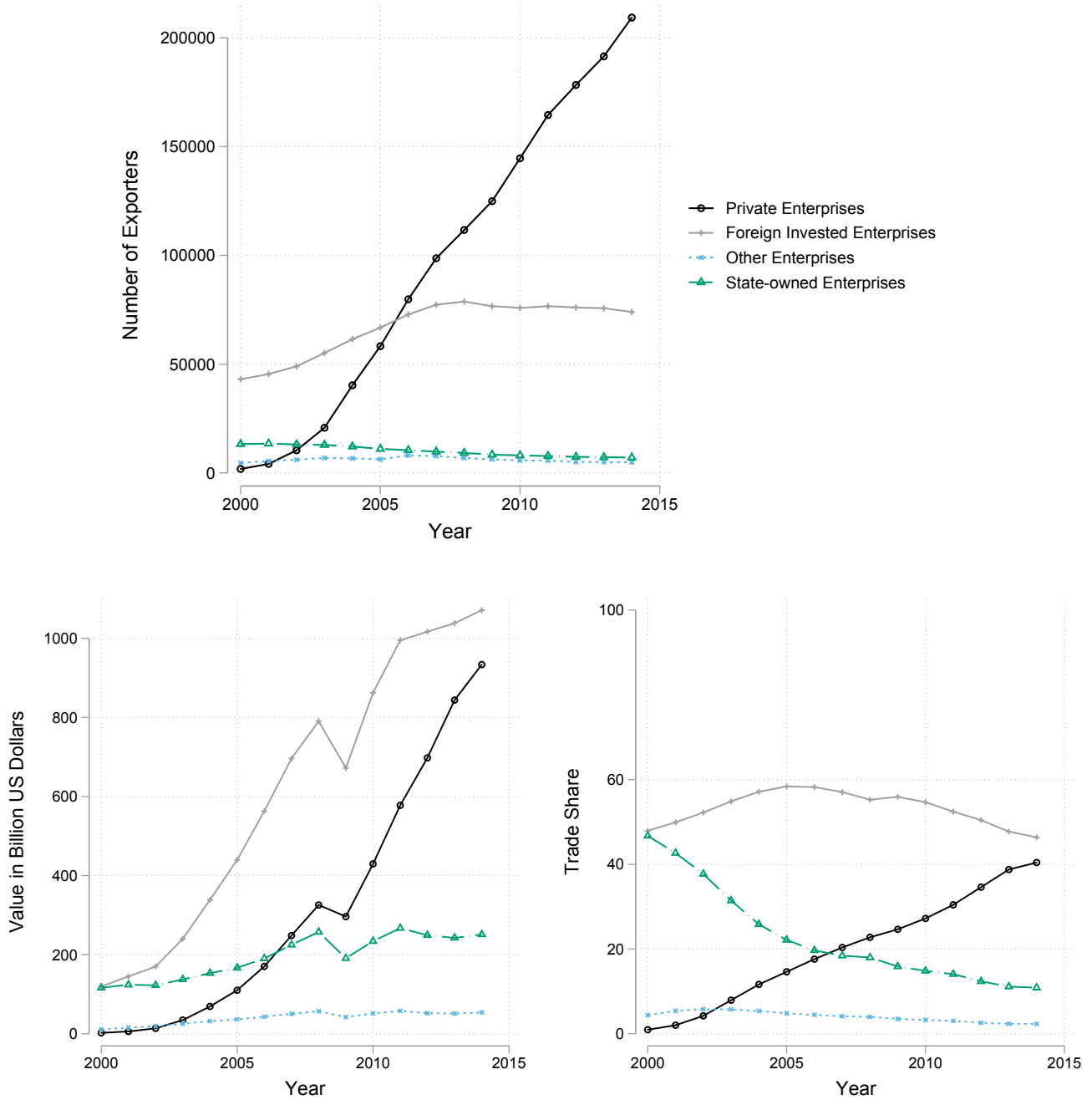
Figure G1: The changing face of Chinese exporters, 2000-2014

Note: Calculations based on the universe of all exporters from the customs database of China. Three types of foreign invested enterprises are reported in our dataset, namely wholly foreign owned enterprises (coded as "4"), sino-foreign joint ventures by jointed equity (coded as "3") and by contractual arrangements that specify the division of tasks and profits (coded as "2"). The last type is quantitatively small in firm number and trade values.

Sino-foreign equity joint venture, wholly foreign owned enterprise, collective enterprise, private enterprise, individual business, and "other" enterprise. We combine Sino-foreign contractual joint ventures, Sino-foreign equity joint ventures, and wholly foreign owned enterprises into a single category - foreign invested enterprises (FIEs). Firms with other ownership structures, including collectives, individual businesses, and "other" enterprises, are lumped together under the descriptor "Other" enterprises.

A well-known fact is the extraordinary rate of entry into export activity by private enterprises. This is apparent in the top panel of the figure. From being a small and neglectable group in 2000, the number of private enterprises directly exporting goods from China to the rest of the world rose to over 200,000 by 2014.[19] Perhaps less known and understood, however, is the economic weight of a different category of exporters *from* China, the foreign-invested enterprises (FIEs). After a slow and steady rise between 2000 and 2006, their number stabilized at about 75,000 firms—dwarfing the presence of state-owned enterprises (SOEs). Indeed, in spite of the attention paid to them by the media, there were only 10,000 registered SOEs at the start of our sample period. This number gradually fell over time, as successive policy initiatives favored their privatization, or led some of them to exit from foreign markets (top panel, figure G1).

The key message from the top panel of figure G1 is reinforced by the evidence on export values and shares by different types of firms, shown in the bottom panel. By export value and share of total exports, the most important single group of exporters from China is that of foreign-invested enterprises. In 2014, the value of their exports was over US $1 trillion (bottom left panel of figure G1). Over the period, exports from China that originated from firms that are wholly or partially owned by foreigners fluctuated between 45 and 58% of China's total exports.[20]

Conversely, the weight of SOEs, which were essentially at par with FIEs in 2000, declined dramatically from 2000 to 2007 and then settled into a slow and steady negative trend (bottom left panel, figure G1). This is clear evidence that the role of SOEs in foreign trade has been far less dynamic than that of other types of firms. However, the diminishing weight of SOEs in foreign trade has been more than made up by private firms—reflecting both entry of new firms into export markets and privatization of SOEs. By the end of the sample, private firms account for a striking 40% of Chinese exports. We stress nonetheless that this large shift in export shares between SOEs

---

[19]At the start of our sample period, export activity was highly regulated in China with most rights to export held by SOEs—only a very limited number of private enterprises were able to export directly. The result of this was that in the earlier years post-2000 private enterprises desiring to export their merchandise exported through SOEs.

[20]The importance of foreign involvement in Chinese exports has previously been documented by Koopman, Wang and Wei (2014). Based on an accounting framework methodology and product-level trade flows, they show that 29.3 percent of Chinese export value comes from foreign, rather than domestic Chinese, value-added. This is not inconsistent with our estimates; our complementary contribution is to document foreign engagement based on *ownership* of exporting firms, rather than through the origin of the value-added content of exported goods.

and private firms has not (so far at least) dented the share of exports by FIEs, which has remained quite stable over our sample.

The question is whether, against this evolution in the number of exporters and export shares by ownership, there are significant differences in strategic pricing.

## G.3    Macroeconomic Data

Macroeconomic variables on nominal bilateral exchange rates, CPI of all destination countries (normalized so that CPI=100 in 2010 for all series), real GDP in constant 2005 US dollars, and the import to GDP ratio come from the World Bank. We construct the nominal bilateral exchange rate in renminbi per unit of destination currency from China's official exchange rate (rmb per US$) and each destination country's official exchange rate in local currency units per US$ (all series are the yearly average rate). These variables are available for 152 destination countries in our sample. For the 17 eurozone countries which we aggregate into a single economic entity, we use the CPI index, bilateral exchange rate and import-to-GDP ratio for the euro area from the World Bank. We construct a measure of real GDP in local currency for the eurozone using the reported GDP in constant US dollars (2010) variable and the 2010 euro-dollar rate.

In our empirical analysis, we focus on nominal rather than real bilateral exchange rates. Estimation using real exchange rates implicitly imposes a one-to-one linear relationship between each nominal bilateral exchange rate and the ratio of CPI indices (i.e., destination CPI/origin CPI). Real exchange rate series which embed this restriction are highly correlated with nominal exchange rates. Since nominal exchange rate series are significantly more volatile over time than the ratio of CPI indices, movements in the real exchange rate are primarily driven by fluctuations in nominal exchange rates. It is not clear if restricting these two variables with significantly different volatilities into a one-to-one linear relationship is justified in exchange rate pass through studies. Throughout our analysis, we enter nominal bilateral exchange rates and destination CPI index as two separate variables.

In all regressions, we enter variables in logged levels. A problem arising from using logged levels rather than time differences is that nominal series, such as exchange rates and CPI indices, cannot be compared directly across countries. To address this compatibility problem, note that the nominal series can be re-written as a comparable measure plus an unobserved destination specific drift, i.e.,

$$e_{dt}^{nominal} = e_{dt}^{comparable} + \mu_d.$$

Under trade pattern fixed effects, the time-invariant destination-specific drift is absorbed into the fixed effects, which enables us to correctly disentangle the effect of nominal exchange rate

fluctuations from destination CPI movements.

## G.4 Trade Pattern Statistics by Product Differentiation

We calculate the trade pattern statistics reported in table 2 separately for high- and low-differentiation goods defined by our CCHS classification. Inspecting Tables G2 and G3, we do not find significant differences in the statistics of market changes for high- and low-differentiation goods in our sample.

Table G2: Number of Unique Trade Patterns - High Differentiation Goods

| Number of Unique Trade Patterns ($y$) | Total Number of Exporting Years ($x$) | | | | | | | | | | | | | | Share |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | |
| 1 | 35.6 | 26.6 | 22.1 | 19.1 | 16.4 | 13.9 | 11.6 | 10.7 | 9.3 | 8.1 | 6.5 | 5.7 | 5.8 | 5.1 | 22.8 |
| 2 | 64.4 | 23.7 | 16.4 | 12.9 | 10.7 | 8.9 | 7.6 | 7.0 | 6.2 | 5.5 | 4.7 | 4.8 | 4.5 | 4.4 | 27.7 |
| 3 | | 49.7 | 20.3 | 14.1 | 10.9 | 8.8 | 6.9 | 6.2 | 5.3 | 4.8 | 3.8 | 3.8 | 3.4 | 3.4 | 14.6 |
| 4 | | | 41.2 | 17.7 | 12.2 | 9.2 | 7.0 | 6.0 | 5.1 | 4.4 | 3.7 | 3.2 | 2.7 | 3.1 | 9.1 |
| 5 | | | | 36.2 | 15.8 | 11.2 | 8.3 | 6.4 | 5.0 | 4.4 | 3.7 | 3.0 | 2.7 | 2.4 | 6.3 |
| 6 | | | | | 34.0 | 14.7 | 9.9 | 7.6 | 6.1 | 4.8 | 3.5 | 3.0 | 2.4 | 2.3 | 4.7 |
| 7 | | | | | | 33.3 | 13.6 | 9.2 | 7.1 | 5.4 | 4.6 | 3.6 | 3.0 | 2.3 | 3.7 |
| 8 | | | | | | | 35.1 | 13.7 | 9.1 | 7.0 | 5.3 | 4.5 | 3.3 | 2.3 | 3.1 |
| 9 | | | | | | | | 33.1 | 13.3 | 9.2 | 6.5 | 5.0 | 3.7 | 2.8 | 2.2 |
| 10 | | | | | | | | | 33.5 | 13.1 | 9.0 | 6.8 | 4.8 | 3.1 | 1.7 |
| 11 | | | | | | | | | | 33.2 | 12.9 | 9.1 | 6.0 | 3.5 | 1.3 |
| 12 | | | | | | | | | | | 35.6 | 13.6 | 7.8 | 5.3 | 1.0 |
| 13 | | | | | | | | | | | | 33.9 | 13.2 | 6.6 | 0.7 |
| 14 | | | | | | | | | | | | | 36.5 | 11.8 | 0.5 |
| 15 | | | | | | | | | | | | | | 41.5 | 0.5 |
| Total | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |

Note: We start from the whole sample of all firms selling *high differentiation* goods and drop firm-product pairs that only exported once in their lifetime. For each firm-product pair, we calculate its total number of exporting years and the number of unique trade patterns in its lifetime and then put it into the relevant cells of the table. The last column gives the share of firm-product pairs with $y$ number of unique trade patterns.

Table G3: Number of Unique Trade Patterns - Low Differentiation Goods

| Number of Unique Trade Patterns ($y$) | Total Number of Exporting Years ($x$) | | | | | | | | | | | | | | | Share |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | |
| 1 | 36.1 | 26.6 | 22.6 | 19.5 | 16.9 | 14.1 | 12.0 | 10.0 | 8.3 | 7.3 | 5.9 | 5.3 | 4.4 | 4.4 | 23.9 |
| 2 | 63.9 | 23.0 | 16.5 | 13.1 | 10.9 | 9.2 | 7.7 | 6.5 | 5.8 | 5.2 | 4.4 | 3.8 | 3.1 | 3.3 | 29.1 |
| 3 | | 50.4 | 20.3 | 14.1 | 11.1 | 8.9 | 7.2 | 6.3 | 5.4 | 4.6 | 3.9 | 3.2 | 2.7 | 2.8 | 15.4 |
| 4 | | | 40.6 | 17.6 | 12.2 | 9.4 | 7.4 | 6.3 | 5.1 | 4.3 | 3.4 | 2.6 | 2.6 | 2.4 | 8.8 |
| 5 | | | | 35.7 | 15.9 | 11.1 | 8.4 | 6.7 | 5.4 | 4.6 | 3.8 | 2.8 | 2.6 | 2.3 | 6.0 |
| 6 | | | | | 33.1 | 15.0 | 10.2 | 7.7 | 6.2 | 5.2 | 3.9 | 3.0 | 2.4 | 2.1 | 4.4 |
| 7 | | | | | | 32.3 | 14.0 | 9.9 | 7.3 | 5.6 | 4.5 | 3.8 | 2.8 | 2.1 | 3.3 |
| 8 | | | | | | | 33.0 | 13.7 | 9.6 | 7.0 | 5.2 | 3.9 | 3.2 | 2.3 | 2.6 |
| 9 | | | | | | | | 32.9 | 13.6 | 9.0 | 6.8 | 5.1 | 3.7 | 2.5 | 1.9 |
| 10 | | | | | | | | | 33.1 | 13.2 | 8.7 | 6.8 | 5.3 | 3.3 | 1.4 |
| 11 | | | | | | | | | | 33.9 | 13.2 | 8.9 | 6.9 | 3.5 | 1.1 |
| 12 | | | | | | | | | | | 36.2 | 13.7 | 8.9 | 5.0 | 0.8 |
| 13 | | | | | | | | | | | | 37.1 | 14.0 | 7.5 | 0.6 |
| 14 | | | | | | | | | | | | | 37.3 | 12.4 | 0.4 |
| 15 | | | | | | | | | | | | | | 44.2 | 0.4 |
| | | | | | | | | | | | | | | | |
| Total | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |

Note: We start from the whole sample of all firms selling *low differentiation* goods and drop firm-product pairs that only exported once in their lifetime. For each firm-product pair, we calculate its total number of exporting years and the number of unique trade patterns in its lifetime and then put it into the relevant cells of the table. The last column gives the share of firm-product pairs with $y$ number of unique trade patterns.

## G.5 Additional Information on the CCHS Classification

### G.5.1 The use of measure words in Chinese grammar

To illustrate how measure words encode meaning in Chinese, consider the problem of counting three small objects. Chinese grammar requires the use of a measure word between the number and the noun being counted. Thus, to say "three ballpoint pens," or "three kitchen knives," one would say the English equivalent of "three long-thin-cylindrical-objects [zhī, 支] ballpoint pens" and "three objects-with-a-handle [bǎ, 把] kitchen knives."[21] Both of these objects, ballpoint pens and kitchen knives, are measured with count/discrete classifiers (zhī and bǎ, respectively) and are, in our classification, high differentiation goods. In contrast, products reported with mass/continuous classifiers including kilograms (cereal grains, industrial chemicals), meters (cotton fabric, photographic film), and cubic meters (chemical gases, lumber) are low differentiation goods. Because measure words encode physical features of the object being counted, they allow us to identify when statistical reporting is for a high versus low differentiation good. According to Cheng and Sybesma (1999), "...the distinction between the two types of classifiers is made with explicit reference to two different types of nouns: nouns that come with a built-in semantic partitioning and nouns that do not – that is, count nouns and mass nouns."

### G.5.2 Comparison to quantity-reporting in other customs systems

While the proposed CCHS classification of goods could lead to some amount of mis-classification because there are some count nouns which exhibit low levels of differentiation and some mass nouns which are quite differentiated, a Chinese-linguistics-based approach to goods classification is still valuable for several reasons. First, nouns with built-in semantic partitioning such as televisions, microscopes and automobiles are high differentiation goods regardless of whether their trade is reported in metric tonnes or units. This is a key advantage of relying on Chinese measure words to classify tradeable goods: measure words clearly identify objects that inherently are semantically partitioned (i.e. are distinct objects), relative to goods that exist as partitionable masses. Second, the use of reported quantity data in other countries' customs systems to identify discrete objects could be less accurate or consistent for a number of reasons discussed below. Finally, the choice of the measure word is predetermined in the minds of Chinese speakers by grammatical rules that have existed for centuries. This choice is clearly exogenous to and predates modern statistical reporting systems.

Like Chinese, Japanese requires the use of measure words between a number word and a

---

[21]English uses measure words; "two dozen eggs" and "a herd of cattle" are two examples. The difference lies in the extent to which unique measure words exist for Chinese nouns and the fact that proper Chinese grammar always requires the use of the appropriate measure word when counting.

noun when counting. Documentation for Japanese trade declarations instructs that the WCO measurement unit "NO" (the English abbreviation for number of items) subsumes 11 indigenous Japanese measure words used with discrete nouns (個、本、枚、頭、羽、匹、台、両、機、隻、着). We interpret these instructions from Japanese customs declarations as a validation of our approach of using count classifiers in the Chinese Customs Database to identify discrete products in the Harmonized System. However, because the official measure of discrete items used in Japanese customs data is an English word, we cannot build a linguistics-based classification of discrete and continuous goods directly from measure words in Japanese data. This is one reason why we prefer to build the classification from Chinese rather than Japanese trade data.[22]

Although goods are inherently discrete (e.g., televisions, automobiles) or continuous (e.g., grain, liquid industrial chemicals), in some customs datasets, discrete products might only be reported by net weight rather than by net weight AND countable units, or quantity reporting could be inconsistent. While the WCO has recommended since 2011 that *net weight* be reported for *all transactions* and supplementary units, such as units/pieces, be reported for specific Harmonized System products, these recommendations are *non-binding*. At one end of the spectrum, EU member states follow their own variation of the WCO guidelines and report net weight as well as a supplemental quantity unit for specific CN products. At the other end, administrative customs data for Egyptian exports over 2005-2016 lists 32 distinct measures of quantity with Egyptian statistics reporting only one measure of quantity per transaction, rather than the two, net mass and supplementary unit, recommended by the WCO. Overall, 87% of Egyptian export observations report net mass (net pounds) as the unit of quantity, only 0.006% report "pieces" as the unit of quantity, and the remainder are scattered across official WCO and alternative measures. Authors' calculations from EID-Exports-2005-2016 obtained from http://erfdataportal.com.

### G.5.3 The dispersion of prices for high and low differentiation goods: A telling example

To provide intuitive evidence about the relevance of our classification in studies of pricing to market, we offer a case study of price adjustments by firms producing two different products – one low differentiation good and one high differentiation good. We select, respectively, canned tomato paste (measured in kilograms) and wheeled tractors (measured with liàng, 辆).

In figure G2, we plot the dispersion of price residuals across destinations for the top three exporters of tomato paste (upper panel) and wheeled tractors (lower panel) in 2007 and 2008. For each annual observation of a sale to a destination, we calculate the deviation of the sales price from its mean across all destinations within the firm-product-year triplet (where sales price is the

---

[22]We thank Taiji Furusawa, Keiko Ito, and Tomohiko Inui for answering our questions about the use of measure words in Japanese trade data.

Figure G2: Price dispersion across destinations for top three firms in 2007 and 2008
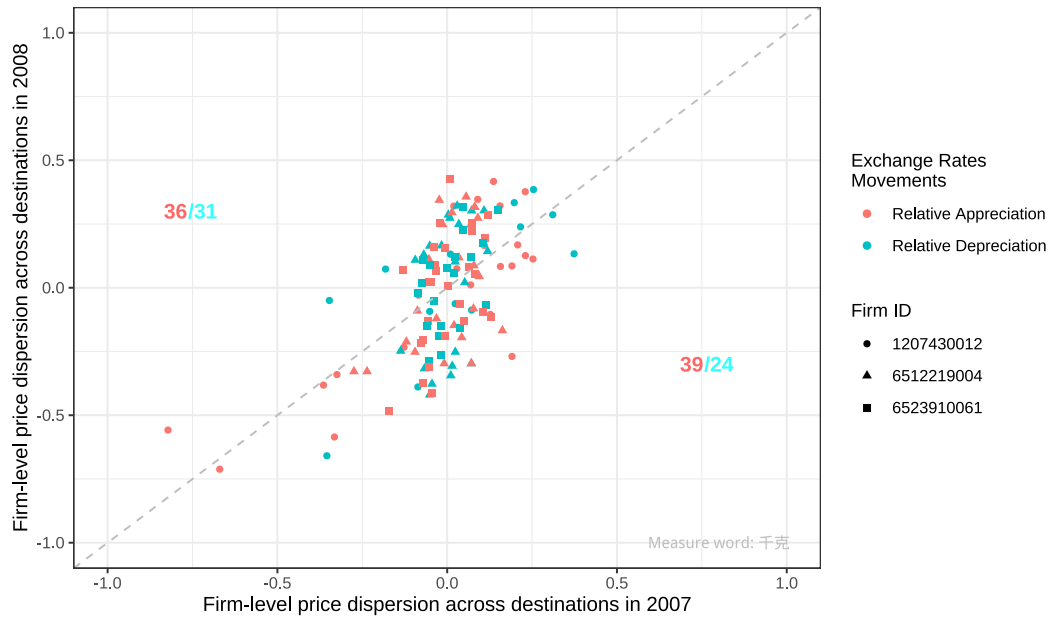


Example 1: Canned Tomato Paste (a low differentiation product)
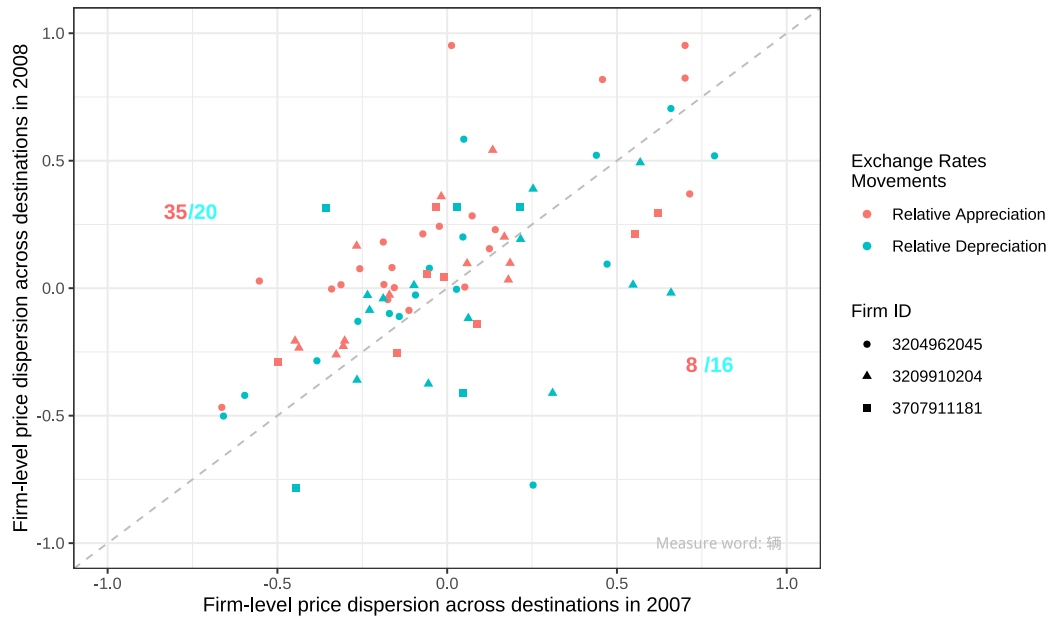


Example 2: Wheeled Tractors (a high differentiation product)

Note: Firm-level price dispersion for tomato paste (HS20029010) and wheeled tractors (HS87019011) is calculated as the deviation from the mean log unit value, denominated in RMB, across destinations at the firm-product-year level, i.e., $uv_{ifdt} - \overline{uv}_{ift}$. For this figure, we begin with a balanced panel of firm-product-destination observations for two consecutive years, 2007 and 2008, and plot the observations of residual price dispersion for the top three firms based on the number of observations in the constructed balanced panel. Red observations are for destinations whose currency appreciated relative to the renminbi between 2007 and 2008 while blue observations are for destinations whose currencies depreciated.

44

log unit value in renminbi), i.e. $uv_{fidt} - \overline{uv}_{fit}$, and plot these deviations using different shapes (i.e., triangle, square, and circle) for each firm. The x-axis measures positive and negative deviations of the sales price from the mean value in 2007; the y-axis measures the deviations from the mean in 2008.[23] Any observation on the 45 degree line is a product whose relative premium or discount in its destination $d$ did not change between 2007 and 2008. Thus, a point lying on the 45 degree line at 0.2 represents a product that was sold in some destination $d$ at a 20% premium over the firm's mean price in both 2007 and 2008. An observation plotted *above* the 45 degree line depicts a product-destination whose price residual increased between 2007 and 2008 *relative* to the firm's sales of the good in other destinations. Conversely, an observation plotted below the 45 degree line represents a product-destination that saw its relative price fall.

We color-code each point representing a firm-product-destination triplet according to whether the destination's currency appreciated or depreciated over 2007-2008 relative to the other destinations the firm was selling to. Red indicates relative appreciation, blue relative depreciation. Above and below the 45 degree line, we report the number of observations marked by red dots, corresponding to bilateral appreciations, in ratio to the number of observations marked by blue dots corresponding to depreciations.

As apparent from these graphs, first, the relative price residuals for many firm-product-destination triplets, measured in the producer's currency, change from year to year. Second, the low differentiation good, tomato paste, exhibits less dispersion in price residuals across destinations than the high differentiation good, wheeled tractors. Third and most importantly, for high differentiation goods, appreciation of the destination market currency relative to the renminbi is associated with an increase in relative price residuals (red dots are denser above the 45 degree line), while depreciation of the destination market currency is associated with a decrease in relative price residuals. No such clear pattern emerges between relative price changes and relative currency changes for the low differentiation good, tomato paste.

### G.5.4 An example of the fine detail in Chinese measure words

To illustrate the variety of count classifiers used for similar objects, note that "Women's or girls' suits of synthetic fibres, knitted or crocheted" (HS61042300) and "Women's or girls' jackets & blazers, of synthetic fibres, knitted or crocheted" (HS61043300) are measured with two distinct Chinese count classifiers, "tào, 套" and "jiàn, 件," respectively. Further, table G4 documents the intrinsic information content of the measurement units for HS04 product groups 8211 and 8212. The Chinese language descriptions of all of these HS08 products conveys the similarity

---

[23]The magnitude of price dispersion within a year across destinations for wheeled tractors is of the same order of magnitude as that found in European automobile prices in an important study of international market segmentation by Goldberg and Verboven (2001).

across products; each Chinese description contains the Chinese character 'dao' (刀), which means 'knife' and is a part of longer compound words including table knife and razor. Interestingly, three different Chinese count classifiers, "tào, 套," "bǎ, 把," and "piàn, 片," are used to count sets of knives (HS82111000), knives and razors (HS82119100 - HS82121000), and razor blades (HS82122000), respectively.

Two further points can be drawn from this table. First, this table illustrates that while Chinese customs statistics are reported for eight digits, in many cases, the final two digits of Chinese customs codes are 00, indicating that the eight digit code is identical to the corresponding six-digit code in the universal Harmonized System. This exemplifies a wider observation that only a single Chinese measure word is used to report quantity for all products in most six-digit HS code. By extension, Chinese measure words can be used to develop a universal classification for the Harmonized System at the six-digit product level. Second, the discrete noun "knife" or 'dao' (刀) appears in the description of every product reported below. This suggest that it would be theoretically possible to develop a binary classification system of Harmonized System products as discrete versus continuous through the use of natural-language processing software that is trained to recognize discrete nouns in any language. In this light, the use of Chinese measure words to identify discrete nouns can be seen as a shortcut in which the linguistic classification of Chinese measure words replaces the data training step.

Table G4: Examples of count classifiers in the Chinese Customs Database

| Quantity Measure | HS08 Code | English Description | Chinese Description |
|---|---|---|---|
| tào, 套 | 82111000 | Sets of assorted knives | 成套的刀 |
| bǎ, 把 | 82119100 | Table knives having fixed blades | 刃面固定的餐刀 |
| bǎ, 把 | 82119200 | Other knives having fixed blades | 其他刃面固定的刀 |
| bǎ, 把 | 82119300 | Pocket & pen knives & other knives with folding blades | 可换刃面的刀 |
| bǎ, 把 | 82121000 | Razors | 剃刀 |
| piàn, 片 | 82122000 | Safety razor blades, incl razor blade blanks in strips | 安全刀片, 包括未分开的刀片条 |

The most frequently used mass classifier is kilograms. Examples of other mass classifiers include meters for "Knitted or crocheted fabric of cotton, width ≤ 30cm" (HS60032000), square meters for "Carpets & floor coverings of man-made textile fibres" (HS57019010), and liters for "Beer made from malt" (HS22030000).

### G.5.5 Integrating the CCHS classification with UN Broad Economic Categories

In table G5, we provide a breakdown of our CCHS classification within the UN's Broad Economic Categories (BEC) of intermediate, consumption and other goods. The majority of intermediate goods are low differentiation and the majority of consumption goods are high differentiation, but all BEC groups include both high differentiation and low differentiation goods.

Table G5: Classification of differentiated goods: CCHS vs. BEC

(a) Share of goods by classification: observation weighted

| | Corsetti-Crowley-Han-Song (CCHS) | | |
| | Low Differentiation / (Mass nouns) | High Differentiation / (Count nouns) | |
|---|---|---|---|
| **BEC** | | | |
| Intermediate | 29.8 | 2.7 | 32.5 |
| Consumption | 14.3 | 20.1 | 34.4 |
| Other† | 15.0 | 18.1 | 33.1 |
| | 59.1 | 40.9 | 100.0 |

(b) Share of goods by classification: value weighted

| | Corsetti-Crowley-Han-Song (CCHS) | | |
| | Low Differentiation / (Mass nouns) | High Differentiation / (Count nouns) | |
|---|---|---|---|
| **BEC** | | | |
| Intermediate | 26.0 | 3.9 | 29.9 |
| Consumption | 8.6 | 14.0 | 22.6 |
| Other† | 12.6 | 34.9 | 47.5 |
| | 47.2 | 52.8 | 100.0 |

Notes: Share measures are calculated based on Chinese exports to all countries including Hong Kong and the United States during periods 2000-2014. †: The "Other" category refers to capital goods and unclassified products by BEC classification, such as nuclear weapons.

### G.5.6 Variation in the CCHS classification across industrial sectors

For twenty industrial sectors, Table G6 reports the share of products in each sector that are classified as high differentiation according to the Corsetti, Crowley, Han, and Song (CCHS) classification. For the 36 measure words in our estimation dataset, we categorize goods measured with the 24 count classifiers as high differentiation, while goods measured with 12 mass classifiers are

Table G6: CCHS product classification across sectors

| Sector (HS chapters) | Sector's share of total exports | Value share of CCHS high differentiation products within sector |
|---|---|---|
| 1-5 Live animals; animal products | 0.8 | 4.0 |
| 6-14 Vegetable products | 1.0 | 0.6 |
| 15 Animal/vegetable fats | 0.0 | 0.0 |
| 16-24 Prepared foodstuffs | 1.4 | 0.0 |
| 25-27 Mineral products | 2.1 | 0.0 |
| 28-38 Products of chemical and allied industries | 4.6 | 0.2 |
| 39-40 Plastics/rubber articles | 3.4 | 15.0 |
| 41-43 Rawhides/leather articles, furs | 1.6 | 58.6 |
| 44-46 Wood and articles of wood | 0.8 | 0.5 |
| 47-49 Pulp of wood/other fibrous cellulosic material | 0.8 | 0.0 |
| 50-63 Textile and textile articles | 13.2 | 68.4 |
| 64-67 Footwear, headgear, etc. | 2.9 | 43.5 |
| 68-70 Misc. manufactured articles | 1.8 | 3.2 |
| 71 Precious or semiprec. stones | 1.4 | 0.0 |
| 72-83 Base metals and articles of base metals | 7.7 | 1.9 |
| 84-85 Machinery and mechanical appliances, etc. | 42.2 | 73.1 |
| 86-89 Vehicles, aircraft, etc. | 4.7 | 66.1 |
| 90-92 Optical, photographic equipment etc. | 3.5 | 79.7 |
| 93 Arms and ammunition | 0.0 | 82.5 |
| 94-96 Articles of stone, plaster, etc. | 6.0 | 65.0 |
| 97 Works of art, antiques | 0.1 | 60.8 |

Source: Compiled by the authors from exports of Chinese Customs Database, 2000-2014, using the Corsetti, Crowley, Han and Song (CCHS) classification.

treated as low differentiation.[24] Column one lists the HS chapters that define the sector. The second column provides the sector's share in China's total exports over 2000-2014. Quantitatively, important export sectors with large shares of high differentiation goods include optical and photographic equipment (79.7 percent), machinery and mechanical appliances (73.1 percent), textiles and apparel (68.4 percent), vehicles and aircraft (66.1 percent), stone and plaster articles (65.0 percent), leather goods (58.6 percent), and plastics and rubber articles (15.0 percent). The share of high differentiation products across sectors varies widely, but lines up with our prior Machinery and mechanical appliances and vehicles and aircraft are dominated by CCHS high differentiation goods while virtually all chemicals and base metal products are low differentiation.

---

[24]We thank Prof. Lisa Lai-Shen Cheng for her feedback on our classification of measure words from the Chinese Customs Database into count and mass classifiers.

### G.5.7 Applying Rauch's classification to Chinese exports

In order to provide a Rauch classification for HS08 products in the Chinese Customs Database, it was first necessary to concord the SITC Rev. 2 product codes from Rauch's classification to universal HS06 product codes. At the HS06 level, 80% of products map into a unique category – differentiated, reference priced or organized exchange – but 20% of products have no unique mapping and are left unclassified. As noted in table 6, when applied to the universe of Chinese exports at the HS08 level, the 1-to-many and many-to-many concordance issue means approximately 12% of firm-product observations cannot be classified into Rauch categories.

Table G7: Mapping HS06 (2007) products to Rauch categories (Rauch's liberal classification)

|  | Number of HS06 codes | Percent of HS06 codes |
|---|---|---|
| HS06 codes with a unique Rauch classification | 4,386 | 79.98 |
| HS06 codes with multiple Rauch classifications | 1,098 | 20.02 |
| Total | 5,484 | 10.00 |

### G.5.8 Integrating the CCHS and Rauch classification systems

According to the Rauch classification system, products traded on organized exchanges are generally regarded as commodities whose prices are expected to fluctuate with global supply and demand. Reference price products are list-price goods: firms producing them compete somewhat directly by supplying at the price published in an industry trade publication. These goods are thought to offer a very limited scope for market power in pricing. Conversely, differentiated goods are defined as goods for which prices are not publicly negotiated—which indicates limited direct competition among firms and greater scope for charging markups. As argued above, our linguistics based classification allows us to refine the Rauch classification by distinguishing differentiated goods using two finer categories, and by classifying goods unclassified under Rauch.

To highlight the contribution of our product-feature-based classification system relative to Rauch (1999)'s market-structure based classification, we now integrate the two in our empirical analysis. Results are shown in table G8.

Table G8: Markup Elasticity by Rauch Classification

| Category | All | HD Goods | LD Goods | n. of obs |
|---|---|---|---|---|
| 2000 − 2005 | | | | |
| | | | | |
| Differentiated Products | 0.07*** | 0.10*** | 0.04 | 3,339,574 |
| | (0.02) | (0.03) | (0.03) | [812,719] |
| Organized Exchange | 0.05 | - | 0.06 | 36,656 |
| | (0.09) | | (0.09) | [11,945] |
| Reference Priced | -0.01 | 0.28 | -0.03 | 332,678 |
| | (0.08) | (0.22) | (0.08) | [88,809] |
| 2006 − 2014 | | | | |
| | | | | |
| Differentiated Products | 0.12*** | 0.20*** | 0.07*** | 15,722,023 |
| | (0.01) | (0.02) | (0.02) | [3,927,425] |
| Organized Exchange | 0.00 | - | 0.00 | 99,373 |
| | (0.09) | | (0.09) | [28,086] |
| Reference Priced | 0.14*** | 0.10 | 0.13*** | 1,537,937 |
| | (0.04) | (0.18) | (0.04) | [364,723] |

Note: Estimates based on specification (17) and the sample of multi-destination trade flows at the firm-product-time level to 152 destinations excluding Hong Kong and the United States. The bilateral exchange rate is defined as RMBs per unit of destination currency; an increase means an appreciation of the destination currency. Robust standard errors are reported in parentheses. The actual number of observations used for identification is reported in the brackets of the last column. Statistical significance at the 1, 5 and 10 percent level is indicated by ***, **, and *.

The most important takeaway from table G8 is that the estimated markup elasticity of "differentiated" goods according to the Rauch classification, 12% in the later period, is an average of very different elasticities for high and low differentiation goods, 20% and 7% respectively. Unsurprisingly, our estimates of markup elasticities are zero for goods traded in organized exchanges, which in our classification are treated as low differentiation goods. Note that for organized exchange-traded goods we can expect prices in renminbi to change with their international market prices, whose movements may be correlated with bilateral exchange rates. For reference-priced goods, consistent with our hypothesis, we find no markup adjustment for the subset of high differentiation goods in this set. Results are less straightforward however for the low-differentiation goods—we find some degree of markup adjustment, although only in the later period.

## G.6 Quantity elasticities: alternative measures

In Table G9, we estimate the quantity elasticity to bilateral exchange rates by directly regressing quantities on exchange rates, i.e., by using estimation equation A4 specified in Appendix A. This specification of our TPSFE estimator is very close to the specification (2) in Berman, Martin and Mayer (2012). Based on our data, we cannot estimate firm-level productivities (as these authors do) to include in the analysis as regressors. We can nonetheless rank firms according to their total export revenue at the product level, and based on this ranking divide firms in three bins—small, medium, large. Further, we can refine this classification distinguishing firms by product differentiation and end-use. As shown in the Table G9, we find that the larger exporters tend to adjust quantity less to exchange rate movements, in line with the results in Berman, Martin and Mayer (2012). The difference across estimates for small, medium and large firms is admittedly less clear cut, compared to Berman, Martin and Mayer (2012)—possibly due to the difference in the information used to classify firms (our is export-revenue-based, rather than value-added based), and the fact that we do not have a proxy for productivity to include in the regressions. Yet, we consistently find more muted quantity responses to exchange rates for consumption goods and high differentiation goods, compared to intermediates and low-differentiation goods. Contrasting results for high versus low differentiation goods, we find stark differences in the quantity elasticities by firm size for the first group of goods only. Conversely, for low differentiation goods, whose markets can be expected to be more competitive, quantity elasticities tend to remain relatively high and significant for all firms, independent of their size. The last six rows of the table confirm that, in most cases, quantity elasticities are higher for low differentiation goods relative to high differentiation goods.

Table G9: Quantity Elasticities by Firm and Product Types (2006 – 2014)

| Product Category | All | Firm Size | | | n. of obs |
| | | Small | Medium | Large | |
|---|---|---|---|---|---|
| All | 0.16*** | 0.17*** | 0.17*** | 0.11* | 19,272,657 |
| | (0.02) | (0.03) | (0.04) | (0.06) | [4,624,185] |
| High Differentiation (HD) | 0.14*** | 0.20*** | 0.12 | 0.02 | 8,263,242 |
| | (0.03) | (0.05) | (0.07) | (0.10) | [1,849,755] |
| Low Differentiation (LD) | 0.17*** | 0.16*** | 0.19*** | 0.16** | 11,009,415 |
| | (0.03) | (0.04) | (0.05) | (0.07) | [2,774,430] |
| Intermediate | 0.17*** | 0.13** | 0.20*** | 0.19* | 6,288,252 |
| | (0.03) | (0.06) | (0.07) | (0.10) | [1,519,759] |
| — HD | 0.13 | 0.08 | 0.13 | 0.27 | 577,029 |
| | (0.13) | (0.19) | (0.26) | (0.36) | [129,561] |
| — LD | 0.17*** | 0.13** | 0.21*** | 0.19* | 5,711,223 |
| | (0.04) | (0.06) | (0.08) | (0.11) | [1,390,198] |
| Consumption | 0.10*** | 0.12* | 0.11 | 0.02 | 6,133,394 |
| | (0.04) | (0.06) | (0.08) | (0.10) | [1,675,480] |
| — HD | 0.05 | 0.15* | 0.04 | -0.14 | 3,580,469 |
| | (0.05) | (0.09) | (0.11) | (0.14) | [982,799] |
| — LD | 0.15*** | 0.10 | 0.20* | 0.17 | 2,552,925 |
| | (0.05) | (0.09) | (0.11) | (0.14) | [692,681] |

Note: Estimates based on specification (A4) and the sample of multi-destination trade flows at the firm-product-time level to 152 destinations excluding Hong Kong and the United States. The bilateral exchange rate is defined as RMBs per unit of destination currency; an increase means an appreciation of the destination currency. Robust standard errors are reported in parentheses. The actual number of observations used for identification is reported in the brackets of the last column. Statistical significance at the 1, 5 and 10 percent level is indicated by ***, **, and *.

## G.7 Comparison of estimators after dropping the singleton trade patterns

In Table G10, we repeat our comparison of estimators in Table 12 using the stage 8 sample, excluding the non-repetitive trade patterns. We report the statistics associated with the Stage 8 sample in Table G12 and conduct a wide range of checks on the distribution of relevant variables across the Stage 7 and Stage 8 samples in our Online Appendix. In G11, we report the number of unique trade patterns associated with the estimation sample of Stage 8.

Table G10: Comparison across Estimators after Dropping the Singleton Trade Patterns

| | (1) TPSFE | (2) HDFE $(fid+t)$ | (3) HDFE $(fit+d)$ | (4) HDFE $(fit+fid)$ | (5) HDFE $(fit+fid)$ in Two Steps | (6) (5) without Trade Pattern $D$ in the second step | n. of obs |
|---|---|---|---|---|---|---|---|
| Destination Demean Fixed Effects | Yes $fidD$ | No $fid+t$ | No $fit+d$ | No $fit+fid$ | Yes $fid+fiD$ | Yes $fid$ | |
| **2006-2014, HD** | | | | | | | |
| SOE | 0.41*** | 0.35*** | 0.35*** | 0.41*** | 0.41*** | 0.02*** | 284,677 |
| | (0.05) | (0.04) | (0.04) | (0.05) | (0.04) | (0.01) | |
| FIE | 0.33*** | 0.38*** | 0.26*** | 0.33*** | 0.33*** | 0.02*** | 445,035 |
| | (0.04) | (0.03) | (0.03) | (0.04) | (0.03) | (0.01) | |
| PE | 0.10*** | 0.13*** | 0.07*** | 0.11*** | 0.11*** | 0.02*** | 1,149,655 |
| | (0.02) | (0.02) | (0.01) | (0.02) | (0.02) | (0.00) | |
| Small | 0.05 | 0.07** | 0.08*** | 0.04 | 0.04 | 0.01 | 682,355 |
| | (0.04) | (0.03) | (0.02) | (0.04) | (0.03) | (0.01) | |
| Medium | 0.18*** | 0.23*** | 0.14*** | 0.18*** | 0.18*** | 0.02*** | 654,521 |
| | (0.04) | (0.03) | (0.02) | (0.04) | (0.03) | (0.01) | |
| Large | 0.34*** | 0.44*** | 0.21*** | 0.34*** | 0.34*** | 0.03*** | 613,937 |
| | (0.03) | (0.03) | (0.02) | (0.03) | (0.03) | (0.01) | |
| Intermediate | 0.12* | 0.09 | 0.07 | 0.11 | 0.11** | 0.02* | 137,328 |
| | (0.07) | (0.07) | (0.05) | (0.07) | (0.06) | (0.01) | |
| Consumption | 0.33*** | 0.38*** | 0.24*** | 0.35*** | 0.35*** | 0.04*** | 1,040,714 |
| | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.00) | |
| **2006-2014, LD** | | | | | | | |
| SOE | 0.09** | 0.09*** | 0.10*** | 0.09** | 0.09*** | 0.03*** | 361,675 |
| | (0.04) | (0.03) | (0.03) | (0.04) | (0.03) | (0.01) | |
| FIE | 0.20*** | 0.29*** | 0.17*** | 0.22*** | 0.22*** | 0.01 | 597,446 |
| | (0.04) | (0.03) | (0.02) | (0.04) | (0.03) | (0.01) | |
| PE | 0.01 | -0.01 | 0.02 | 0.01 | 0.01 | 0.02*** | 1,846,478 |
| | (0.02) | (0.01) | (0.01) | (0.02) | (0.01) | (0.00) | |
| Small | 0.05 | 0.04* | 0.02 | 0.04 | 0.04 | 0.01* | 1,025,635 |
| | (0.03) | (0.02) | (0.02) | (0.03) | (0.02) | (0.01) | |
| Medium | 0.04 | 0.04 | 0.05*** | 0.05* | 0.05** | 0.02*** | 970,470 |
| | (0.03) | (0.02) | (0.01) | (0.03) | (0.02) | (0.01) | |
| Large | 0.13*** | 0.19*** | 0.08*** | 0.13*** | 0.13*** | 0.02*** | 899,710 |
| | (0.03) | (0.02) | (0.01) | (0.03) | (0.02) | (0.01) | |
| Intermediate | 0.04** | 0.08*** | 0.04*** | 0.04** | 0.04*** | 0.02*** | 1,445,816 |
| | (0.02) | (0.01) | (0.01) | (0.02) | (0.01) | (0.00) | |
| Consumption | 0.09*** | 0.13*** | 0.05*** | 0.09*** | 0.09*** | 0.02*** | 720,020 |
| | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.01) | |

Note: Each cell reports the estimated markup elasticity from the estimation method specified on top of each column. Each row indicates a different subsample. Within in a row, all methods are applied based on the same sample. The number of observations in the last column corresponds to Stage 8 of the data cleaning procedure specified in appendix G.9. Robust standard errors are reported in parentheses. Statistical significance at the 1, 5 and 10 percent level is indicated by ***, **, and *.

Table G11: Number of Unique Trade Patterns Associated with the Firm-product Pairs in the Stage 8 Sample (2006-2014)

| No. of Unique Trade Patterns | Observations | % | Cumulative % |
|---|---|---|---|
| 1 | 4,420,003 | 91.33 | 91.33 |
| 2 | 376,796 | 7.79 | 99.12 |
| 3 | 40,301 | 0.83 | 99.95 |
| 4 | 2,192 | 0.05 | 100.00 |
| 5 | 41 | 0.00 | 100.00 |
| Total | 4,839,333 | 100.00 | |

Note: After removing observations associated with singleton trade patterns, we are left with a sample of approximately 4.8 million observations. We calculate the number of unique trade patterns for each firm-product pairs in this sample. The table reports the number of observations of firm-product pairs associated with 1, 2, etc. unique trade patterns.

## G.8   In which currency do exporters from China invoice?

The Chinese Customs Authority reports the value of export shipments in US dollars, but does not provide any information about whether the trade was invoiced in US dollars, renminbi, another vehicle currency or the currency of the destination. We turn to the customs records of Her Majesty's Revenue and Customs (HMRC) in the United Kingdom, one of China's major destination markets, to shed light on this issue.

We interpret the widespread prevalence of dollar invoicing for a country that issues its own vehicle currency as suggestive that Chinese exports to other countries, including those that do not issue vehicle currencies, are likely predominately invoiced in US dollars.

Since 2010, HMRC has recorded the invoicing currency for the vast majority of import and export transactions between the UK and non-EU trading partners.[25] Figure G3 presents the shares of import transactions and import value into the UK from China by invoicing currency.[26] Results are reported for three currencies, the euro (EUR), pound sterling (GBP), and the US dollar (USD). All transactions that use other currencies of invoice, for example, the Swiss franc, Japanese yen or

---

[25]The reporting requirements for invoice currency are described in *UK Non-EU Trade by declared currency of Invoice (2016)*, published 25 April 2017. See page 7: "Only data received through the administrative Customs data collection has a currency of invoice declared... For Non-EU import trade, businesses must submit the invoice currency when providing customs declarations. However, 5.0 per cent of Non-EU import trade value [in 2016] did not have a currency... This was accounted for by trade reported through separate systems, such as parcel post and some mineral fuels. For Non-EU export trade, businesses are required to declare invoice currency for declarations with a value greater than £100,000. As a result of this threshold and trade collected separately (reasons outlined above) 10.1 per cent of Non-EU export trade [in 2016] was declared without a currency."

[26] To construct this figure, we begin with the universe of UK import transactions for goods originating from China over 2010-2016. Then, we aggregate all transactions within a year that are reported for a firm-CN08product-quantity measure-currency quadruplet to an annual observation for that quadruplet. The variable "quantity measure" records whether a transaction for a CN08 product is reported in kilograms or a supplementary quantity unit like "items" or "pairs." This leaves us with 2.004 million annual transactions which we use to construct figure G3.

Figure G3: Invoicing currencies for UK imports from China
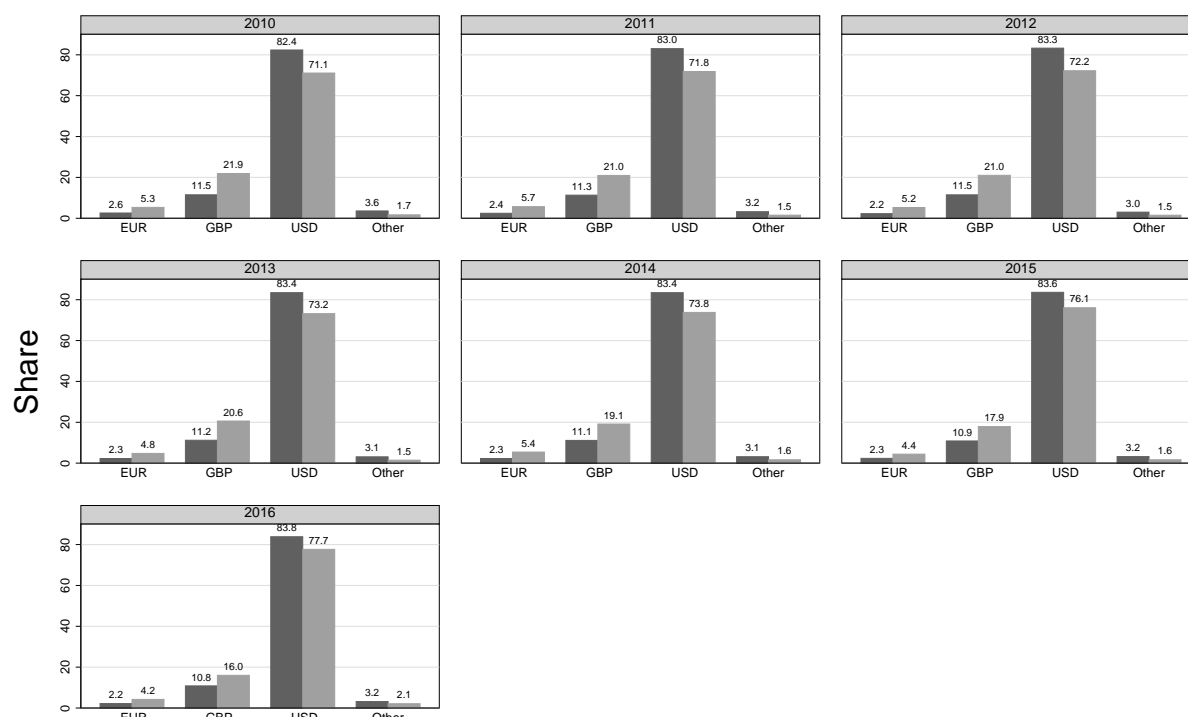


Black: Share of Transactions; Grey: Share of Trade Value
Source: Calculations based on HMRC administrative datasets.

Chinese renminbi, are aggregated into the category "Other."[27] In each graph, the dark bar refers to the share of transactions and the light grey bar refers to the share of import value reported in the relevant currency.

The first point to note is that virtually all of the UK's imports from China are invoiced in one of three major currencies: the pound sterling (GBP), the US dollar (USD), or the euro (EUR). Very little trade is invoiced in any other currency, including the Chinese renminbi.

The second striking point is that the most important currency for Chinese exports to the UK is the US dollar. The dollar's prominence as the invoicing currency of choice for Chinese exports to the UK rose over 2010-2016 with the share of import value growing from 71.1% to 77.7%. The share of transactions invoiced in US dollars was stable at around 83% throughout the 2010-2016 period.[28] Over this same period, the pound's importance as an invoicing currency for imports from China fell. While the share of transactions invoiced in sterling held steady at 10-12% over the period, the share of import value fell from a high of 21.9% in 2010 to a low of 16.0% by

---

[27] We do not report the number of transactions for which the currency is not reported; the number of transactions with no currency reported falls below HMRC Datalab's threshold rule of firms in at least one year and is, for confidentiality reasons, omitted from the figure.

[28] See also Goldberg and Tille (2008) and Goldberg and Tille (2016) who document relatively large shares of exports invoiced in dollars for many countries.

2016. The importance of the euro as an invoicing currency for Chinese exports to Britain was low throughout the 2010-2016 period.

This evidence is relevant to our empirical analysis insofar as a firm that invoices in a vehicle currency, say dollars, also prices its good in that currency. Suppose that the firm sets one single price for its product in dollars: this practice (arguably maximizing the markup relative to global demand) would rule out destination specific adjustment in markups. In this case, our TPSFE estimation should yield insignificant results. The same would be true if firms set different dollar prices across markets (in line with evidence of deviations from the law of one price), but do not adjust them in response to fluctuations in the exchange rate.

This suggests that our TPSFE estimator of markup elasticities can provide evidence on a relevant implication of what Gopinath has dubbed the 'International Price System.' Specifically, our empirical findings can inform us about the possibility of dollar invoicing translating into a 'reference price system' in which firms do not exploit market-specific demand elasticities, but price in relation to global demand. If a reference price system dominates, we would expect to observe firms setting one prevailing price in the global market for manufactured goods as they do for commodities.

## G.9   Price Changes and Trade Pattern Dummies

In this subsection, we show how we build our (unbalanced) panel. We will rely on an example to explain how we identify price changes at the firm-product destination level and trade patterns across destinations at the firm-product level in the data.

Consider a firm exporting a product to five countries, A through E, over 6 time periods. In the following matrix, $t = 1, 2, 3, ...$ indicates the time period and A, B, C, D, E indicates the country. Empty elements in the matrix indicate that there was no trade.

$$
\begin{array}{cccccc}
t = 1 & A & B & & & \\
t = 2 & A & B & C & & E \\
t = 3 & A & B & C & D & \\
t = 4 & A & & C & D & E \\
t = 5 & A & B & C & & \\
t = 6 & A & B & C & D &
\end{array}
$$

The following matrix records export prices by destination country and time:

$$
\begin{bmatrix}
p_{A,1} & p_{B,1} & \cdot & \cdot & \cdot \\
p_{A,2} & p_{B,2} & p_{C,2} & \cdot & p_{E,2} \\
p_{A,3} & p_{B,3} & p_{C,3} & p_{D,3} & \cdot \\
p_{A,4} & \cdot & p_{C,4} & p_{D,4} & p_{E,4} \\
p_{A,5} & p_{B,5} & p_{C,5} & \cdot & \cdot \\
p_{A,6} & p_{B,6} & p_{C,6} & p_{D,6} & \cdot
\end{bmatrix}
$$

Suppose the pricing currency is the dollar and we want to identify price changes in dollars. First, we compare export prices denominated in dollars over time and at the firm-product-destination level as illustrated in the following figure. Price changes less than 5% are marked with "x".
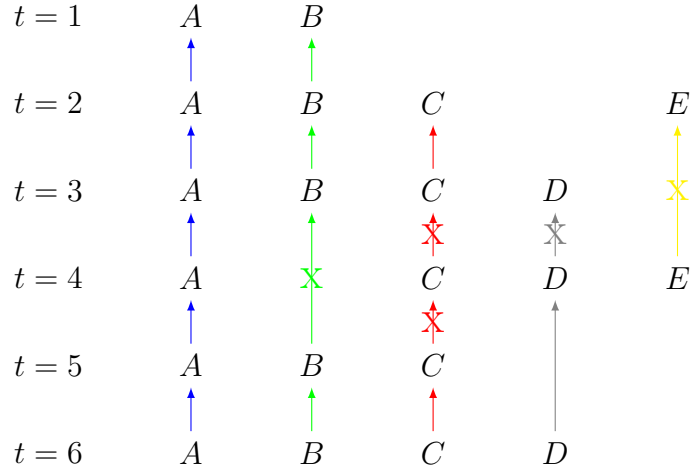


We then set the batch of individual prices associated with a price changes below $\pm 5\%$ ($p_{B,5}, p_{C,4}, p_{D,4}, p_{E,4}$) to missing. This gives

$$
\begin{bmatrix}
p_{A,1} & p_{B,1} & \cdot & \cdot & \cdot \\
p_{A,2} & p_{B,2} & p_{C,2} & \cdot & \cdot \\
p_{A,3} & p_{B,3} & p_{C,3} & p_{D,3} & p_{E,3} \\
p_{A,4} & \cdot & \cdot & \cdot & \cdot \\
p_{A,5} & \cdot & p_{C,5} & \cdot & \cdot \\
p_{A,6} & p_{B,6} & p_{C,6} & p_{D,6} & \cdot
\end{bmatrix}
$$

Note that we did not treat $p_{C,5}$ as missing at this stage. This is because $|p_{C,5} - p_{C,3}|$ could be $> 5\%$ even if both $|p_{C,4} - p_{C,3}| < 5\%$ and $|p_{C,5} - p_{C,4}| < 5\%$.[29] Rather, we repeat the above step using the remaining observations as illustrated below.

---

[29]Variables are in logs.

$t = 1$      $A$      $B$

$t = 2$      $A$      $B$      $C$          $E$

$t = 3$      $A$      $B$      $C$      $D$

$t = 4$      $A$

$t = 5$      $A$          $C$

$t = 6$      $A$      $B$      $C$      $D$

In this example, we indeed find $|p_{C,5} - p_{C,3}| > 5\%$ and the remaining pattern is given as follows. As no prices are sticky, we can stop the iteration.[30] Note that as no price changes can be formulated for the single trade record $p_{E,2}$, this observation is dropped from our sample.

$$
\begin{bmatrix}
p_{A,1} & p_{B,1} & \cdot & \cdot & \cdot \\
p_{A,2} & p_{B,2} & p_{C,2} & \cdot & \cdot \\
p_{A,3} & p_{B,3} & p_{C,3} & p_{D,3} & \cdot \\
p_{A,4} & \cdot & \cdot & \cdot & \cdot \\
p_{A,5} & \cdot & p_{C,5} & \cdot & \cdot \\
p_{A,6} & p_{B,6} & p_{C,6} & p_{D,6} & \cdot
\end{bmatrix}
$$

Now we have identified the universe observations with price changes. The next step is to formulate the trade pattern dummy.

$t = 1$      $A$      $B$

$t = 2$      $A$      $B$      $C$

$t = 3$      $A$      $B$      $C$      $D$

$t = 4$      $A$

$t = 5$      $A$          $C$

$t = 6$      $A$      $B$      $C$      $D$

In this example, we find 5 trade patterns, i.e., $A - B$, $A - B - C$, $A - B - C - D$, $A$, $A - C$, but only one pattern, $A - B - C - D$, which appears at least two times. To compare the change in relative prices across destinations, we require the same trade pattern be observed at least two times in the price-change-filtered dataset. Essentially, by formulating trade pattern fixed effects,

---

[30]In the real dataset, the algorithm often needs to iterate several times before reaching this stage.

we are restricting the comparison within a comparable environment. Firms switch trade patterns for a reason. Restricting the analysis to the same trade pattern also controls for other unobserved demand factors affecting the relative prices.

## G.10 Data cleaning process and the number of observations

Table G12: Key Statistics for Our Data Cleaning Process

| Stage | Observations | Value (Billions US$) | Destinations | Products (HS06) | Products (HS08) | Products (Refined†) | Firms |
|---|---|---|---|---|---|---|---|
| | | | | *Number of Unique Values* | | | |
| 0 | 108,465,375 | 17,453 | 246 | 5,899 | 10,002 | - | 581,141 |
| 1 | 92,308,538 | 11,553 | 244 | 5,880 | 9,959 | - | 545,175 |
| 2 | 92,177,750 | 11,546 | 243 | 5,875 | 9,954 | 20,472 | 545,133 |
| 3 | 83,439,493 | 11,546 | 227 | 5,875 | 9,954 | 20,472 | 545,133 |
| 4 | 76,662,842 | 10,878 | 155 | 5,867 | 9,929 | 20,334 | 531,505 |
| 5 | 72,025,441 | 9,004 | 155 | 5,867 | 9,929 | 20,334 | 531,505 |
| 6 | 49,722,707 | 7,228 | 155 | 5,445 | 9,040 | 17,232 | 355,843 |
| 7 | 23,552,465 | 5,980 | 152 | 5,041 | 8,076 | 14,560 | 237,933 |
| 8 | 5,912,633 | 1,213 | 152 | 5,000 | 7,955 | 14,111 | 209,003 |

† A refined product is defined as 8-digit HS code + a form of commerce dummy. More precisely, this could be described as a variety but we used the term product throughout the paper.

**Stage 0:** Raw data
**Stage 1:** Drop exports to the U.S. and Hong Kong
**Stage 2:** Drop if the destination identifier, product identifier or value of exports is missing; drop duplicated company names
**Stage 3:** Collapse at the firm-product-destination-year level; integrating 17 eurozone countries into a single economic entity
**Stage 4:** Drop observations if bilateral exchange rates or destination CPI is missing
**Stage 5:** Filtering price changes (in logs, denominated in dollar) < 0.05 at the firm-product-destination level following the method described by G.9
**Stage 6:** Drop single-destination firm-product-year triplets
**Stage 7:** Drop single-year firm-product-destination triplets
**Stage 8:** Formulating trade pattern; Drop single-year firm-product-trade-pattern triplets
(Finally, we drop "single-year firm-product-trade-pattern triplets." Including these observations will not change the estimates obtained from the TPSFE estimator because they do not provide the within firm, product and destination *intertemporal variation* upon which the estimator relies.)

# References

**Balazsi, Laszlo, Laszlo Matyas, and Tom Wansbeek.** 2018. "The Estimation of Multidimensional Fixed Effects Panel Data Models." *Econometric Reviews*, 37(3): 212–227.

**Berman, Nicolas, Philippe Martin, and Thierry Mayer.** 2012. "How Do Different Exporters React to Exchange Rate Changes?" *The Quarterly Journal of Economics*, 127(1): 437–492.

**Charbonneau, Karyne B.** 2017. "Multiple Fixed Effects in Binary Response Panel Data Models." *The Econometrics Journal*, 20(3): S1–S13.

**Cheng, L. Lai-Shen, and Rint Sybesma.** 1999. "Bare and Not-so-bare Nouns and the Structure of NP." *Linguistic Inquiry*, 30(4): 509–542.

**Correia, Sergio.** 2017. "Linear Models with High-dimensional Fixed Effects: An Efficient and Feasible Estimator." *Working Paper*.

**De Loecker, Jan, Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik.** 2016. "Prices, Markups, and Trade Reform." *Econometrica*, 84(2): 445–510.

**Fernández-Val, Iván, and Martin Weidner.** 2016. "Individual and Time Effects in Nonlinear Panel Models with Large N, T." *Journal of Econometrics*, 192(1): 291–312.

**Goldberg, Linda, and Cedric Tille.** 2016. "Micro, Macro, and Strategic Forces in International Trade Invoicing: Synthesis and Novel Patterns." *Journal of International Economics*, 102(C): 173–187.

**Goldberg, Linda S., and Cédric Tille.** 2008. "Vehicle Currency Use in International Trade." *Journal of International Economics*, 76(2): 177–192.

**Goldberg, Pinelopi K., and Frank Verboven.** 2001. "The Evolution of Price Dispersion in the European Car Market." *The Review of Economic Studies*, 68(4): 811–848.

**Guimaraes, Paulo, and Pedro Portugal.** 2011. "A Simple Feasible Procedure to Fit Models with High-dimensional Fixed Effects." *The Stata Journal*, 10(4): 628–649.

**Heckman, James J.** 1979. "Sample Selection Bias As a Specification Error." *Econometrica: Journal of the Econometric Society*, 153–161.

**Honoré, Bo, Francis Vella, and Marno Verbeek.** 2008. "Attrition, Selection Bias and Censored Regressions." In *The Econometrics of Panel Data*. 385–418. Springer.

**Koopman, Robert, Zhi Wang, and Shang-Jin Wei.** 2014. "Tracing Value-added and Double Counting in Gross Exports." *The American Economic Review*, 104(2): 459–94.

**Kyriazidou, Ekaterini.** 1997. "Estimation of a Panel Data Sample Selection Model." *Econometrica: Journal of the Econometric Society*, 1335–1364.

**Levinsohn, James, and Amil Petrin.** 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *The Review of Economic Studies*, 70(2): 317–341.

**Matyas, Laszlo.** 2017. *The Econometrics of Multi-dimensional Panels.* Springer.

**Olley, Steven, and Ariel Pakes.** 1996. "The Dynamics of Productivity in the Telecomunications Equipment Industry." *Econometrica*, 64: 1263–97.

**Rauch, James E.** 1999. "Networks Versus Markets in International Trade." *Journal of International Economics*, 48(1): 7–35.

**Rios-Avila, Fernando.** 2015. "Feasible Fitting of Linear Models with N Fixed Effects." *The Stata Journal*, 15(3): 881–898.

**Verbeek, Marno, and Theo Nijman.** 1996. "Incomplete Panels and Selection Bias." In *The Econometrics of Panel Data.* 449–490. Springer.

**Wansbeek, Tom, and Arie Kapteyn.** 1989. "Estimation of the Error-components Model with Incomplete Panels." *Journal of Econometrics*, 41(3): 341–361.

**Wooldridge, Jeffrey M.** 2009. "On Estimating Firm-level Production Functions Using Proxy Variables to Control for Unobservables." *Economics Letters*, 104(3): 112–114.