Trade and Innovation: The Role of Scale and Competition Effects*

Kevin Lim† Daniel Trefler‡ Miaojie Yu §

June 13, 2018

Abstract

This paper studies the effects of scale and competition on firm-level innovation in China. Using both econometrics and a calibrated structural model, we disentangle the mechanisms via which trade affects innovation, focusing on scale effects (impact on market size) and competition effects (impact on markups). The structural model also examines heterogeneity of these affects across firms, which leads to a new mechanism for competition effects: firms can escape the competition by innovating into a market segment where competition is less intense. The econometric estimates and simulations of the calibrated structural model indicate that both scale and competition effects are important for understanding how trade affects innovation in China. In particular, scale effects of trade on innovation are positive in the aggregate, whereas competition effects are negative. However, when firms can innovate to escape the competition, greater competition induced by lower trade barriers can lead firms to increase innovation rather than reduce it. Finally, the calibrated model allows us to examine the impact of reductions in trade costs between China and the OECD on quality, productivity, markups and innovation around the world.

*We are indebted to Scott Orr, David Rivers and Frederic Warzynski for discussions about productivity and markup estimation. Scott Orr developed the code for estimation. We are grateful to discussants at the 2017 AEA meeting (Murat Celik) and Stanford (Nick Bloom) as well as seminar participants at Peking University (especially Loren Brandt and Justin Lin), CIFAR, Science Po, the St. Louis Fed, and the Bank of Canada lecture at the 2018 CEA meeting. Trefler thanks the Canadian Institute for Advance Studies (CIFAR) Program in Institutions, Organizations and Growth, the Social Sciences and Humanities Research Council of Canada (SSHRC) and the Bank of Canada for generous intellectual and financial support.

†Department of Economics, University of Toronto

‡Rotman School of Management, University of Toronto, CIFAR, NBER (dtrefler@rotman.utoronto.ca).

§China Center for Economic Research, National School of Development, Peking University

1 Introduction

This paper studies how trade affects firm-level innovation in China through two channels: scale and competition. On the one hand, an increase in the size of the market available to a firm can raise the returns to successful innovation and hence induce greater investment in innovative activities. At the same time, firms in a larger market face tougher competition, which may either incentivize or disincentivize innovation. We conjecture that these market size and competition effects are precisely what drive innovation in China.

To investigate, we study Chinese firm-level data matched with R&D data, patent data, and international transactions data. We use the data to examine whether rising rates of innovation by Chinese firms can be explained by improved access to foreign markets, and whether China's rising productivity and quality can be explained by rising rates of innovation. Econometric evidence strongly suggests that increases in foreign market size have positive effects on firm innovation, while greater competition within China as well as fro other Chinese firms in export markets reduces innovation by Chinese firms in the aggregate. These econometrics effects are essentially difference-in-difference results that compare across firms and time the effects of scale and competition. They are thus *relative* effects. A calibrated model is needed to estimate *level* effects.

Motivated both by the empirical evidence and the need for results about level effects, we develop a dynamic structural trade model that features both endogenous competition and innovation. In the model, firms choose R&D investments to move up a product grade ladder, where grades differ endogenously in terms of competitiveness and profitability. The incentives for innovation depend on the size of the market and the levels of competition within each grade, which in turn depend on the trade environment. We calibrate the key parameters of the model using the matched Chinese firm-level data, and simulate counterfactuals to study both the aggregate effects of trade on innovation as well the decomposition of these effects into scale and competition effects. Simulations of the calibrated structural model indicate that both scale and competition effects are important for understanding how trade affects innovation in China. In particular, when firms can innovate to escape the competition, greater competition induced by lower trade barriers can lead firms to increase rather than reduce innovation.

The contributions of this paper to the literature on trade and innovation are thus threefold. First, it extends the body of work that studies the interaction between market size and firm-level innovation to the context of international trade by Chinese firms. In a domestic setting, Acemoglu and Linn (2004) find large effects of potential market size (driven by US demographic changes) on innovation by pharmaceutical firms, while Beerli et al. (2018) find positive effects of domestic market size on innovation by Chinese firms across durable goods sectors. In an international trade setting, Lileeva and Trefler (2010) find positive effects of lower US tariffs on innovation by Canadian plants, while Bustos (2011) finds positive effects of reductions in Brazilian tariffs through the MERCOSUR trade agreement on innovation by Argentinian firms. Similarly, Aw et al. (2011) find that larger export markets for Taiwanese electronics firms leads to greater investments in innovation, while Coelli et al. (2018) find large effects of tariff reductions on firm-level innovation worldwide as measured by patent data. Our results show that these positive scale effects of trade on innovation characterize innovative behavior by Chinese firms as well.

Second, we expand on the area of the literature focusing on the interaction between competition and firm-level innovation. In particular, we study a model with both endogenous competition (variable markups) as well as a motive for firms to innovate in order to move into market segments with less competition. In this sense, we embed the "escape-the-competition" motive for innovation emphasized by Aghion et al. (2001, 2005) into a general equilibrium trade model, and show that this mechanism is important for understanding innovation by Chinese firms. Our study of both scale and competition effects is similar in spirit to work by Aghion et al. (2017) and Impullitti and Licandro (2018), although the key economic mechanisms differ in a meaningful way. In Aghion et al. (2017), there is no "escape-the-competition" motive for innovation, and greater competition unambiguously disincentivizes innovation. In Impullitti and Licandro (2018), competition can induce greater innovation amongst oligopolistic firms due to improvements in static efficiency, although the extensive margin of competition (number of rival producers) is not considered. Fieler and Harrison (2018), who also study Chinese firm-level data, find that tariff reductions in a firm's downstream sector lead to increases in firm productivity, and argue that this can be rationalized by firms innovating to escape the indirect competition that propagates upstream via input-output linkages. In contrast, we abstract from input-output linkages to focus on 'escape the competition' responses of innovation to direct competition within a firm's own market.

By focusing on heterogeneous effects of competition across firms, we also aim to provide some resolution to the question of whether trade-related competition induces or reduces innovation. As yet, the empirical evidence is mixed: for instance, Autor et al. (2017) find that greater competition from Chinese imports led US firms to reduce innovation (as measured by patents), whereas Bloom et al. (2016) find that rising competition from Chinese imports led to an increase in innovative activities within firms most affected by Chinese import competition. Within the Chinese market, Bombardini et al. (2018) find that increased foreign import competition induced by China's accession to the WTO encouraged innovation for only the most productive Chinese firms. These findings are consistent with our model once the combined effects of scale and competition across firms in different market segments are considered. These results also have important policy implications, as Akcigit et al. (2017) show how R&D subsidies in response to foreign competition can be welfare-improving in the long-run, while import tariffs create large dynamic losses.

Finally, we contribute to the literature by studying both scale and competition effects in a general equilibrium setting. In this vein, Atkeson and Burstein (2010) argue that although lower trade barriers can encourage innovation, the resulting welfare gains are small because of offsetting general equilibrium effects that operate via firm entry. Building on this work, Atkeson and Burstein (2018) argue that there is limited scope for innovation subsidies to generate increases in aggregate productivity and output. However, these theoretical analyses are conducted in an environment with constant markups, and hence consider only the scale effects of trade on innovation. Our work aims to extend these general equilibrium results by considering economies with endogenous competition as well.

The outline of the paper is as follows. Section 2 begins by describing the data. Section 3 then discusses econometric evidence for scale and competition effects of trade on innovation. Section 4 then develops the closed economy structural model, while section 5 extends the model to an open economy. Next, section 6 describes calibration of the model's parameters, and section 7 discusses the counterfactual exercises that we employ the model to study. Finally, section 8 concludes.

2 Data

We use the 2000–2006 Chinese Manufacturing Enterprises (CME) database, which includes all state-owned enterprises (SOEs) and large non-SOEs whose annual sales are more than RMB 5 million (approximately $600,000US). We clean the data as in Brandt et al. (2014), Brandt et al. (2012) and Brandt et al. (2017). Of note, we delete firms that report less than 8 employees in all years and, when calculating productivity and markups, drop firms that switch 2-digit industries. In addition, we clean up the data following Feenstra et al. (2014), meaning, we delete observations with (a) incomplete or internally inconsistent financial variables, (b) fewer than 8 employees, and (c) invalid entry for year.[1]

We merge theses data with export and import data at the HS8 level from the Chinese General Administration of Customs. We match the CME and customs data following Yu (2015), matching firm name or zip code or telephone number. We are able to match 76,946 firms, which is more than 40% of the firms and 53% of the export value in the customs data.[2] There are two sources of export data because the CME itself reports the total value of a firm's exports (not disaggregated by HS8 or destination). If a firm is not matched to the customs database but reports zero exports (as opposed to missing exports) then we treat it as a non-exporter.[3] See the online appendix for details of the CME, the customs data and the matching algorithm.

If the CME firm is matched to the customs data then we use the customs data. This is 16% of our sample. If it is not matched then we use CME exports. These exports are sometimes missing and we set them to zero in cases where the firm always reports either zero or missing exports and never positive exports.

Our key variables are exports, quality, markups, and three measures of innovation. We discuss each of these in turn.

2.1 Innovation data

We use three measures of innovation. The first is patents, which we merge in with the CME-customs matched data. Unmatched firms are assumed to have no patents. Because a small number of firms have thousands of patents, we top code patents at 50; however, our results are not sensitive to this. The second innovation measure is R&D. We work with 100(R&D/Sales) and because a few firms report inexplicably high values we top code the data at 20%. The third measure is the share of total sales that are generated by new products. These data are from a new-products question in the CME survey.

We note that the patent and R&D data are skewed, with very few firms reporting positive amounts of one or the other. We therefore also use the principal component of the three measures. Specifically, we calculate the principal component separately by 2-digit CIC industry.

[1] In this draft of the paper, we also delete observations with missing firm identification. This will change in future drafts.

[2] The 53% is comparable to the match in the Canadian database. The 'lost' export value is due to the fact that many firms export via trade intermediaries.

[3] There are almost no instances in which a matched firm has (non-zero) customs exports and zero CME exports.

2.2 Markups (and RTFP)

We estimate markups using De Loecker and Warzynski (2012) and so must first estimate TFP. TFP estimation is described in detail Orr et al. (forthcoming). We have checked that the data used produces almost identical aggregates to those used in We start by dropping firms from the data based on four criteria that are relevant for productivity analysis. First, the firms must have complete data on sales, employment, material costs, and capital. Second, they cannot have 'holes' over time, i.e., if they appear in years t_0 and t_1 then they must appear in all years between t_0 and t_1. Third, they cannot switch industries or cities over time (city switching is very rare). Industries are defined at the 2-digit Census Industry Classification level (CIC-2). Fourth, we drop Tobacco (CIC-2 = 16) because it has too few firms. This leaves us with 772,788 firm-years and 298,259 firms in 28 industries.

To prepare the data for estimation, we deflate each firm's sales using an industry-level price deflator. We deflate materials input expenditures at the industry level using input price deflators that have been filtered through the Chinese input-output tables.[4] We estimate the production function by CIC-2 industry for Cobb-Douglas and translog, and for value-added and gross-output production functions. As discussed in Orr et al. (forthcoming), the translog gross-output production function estimates are most sensible as judged by input elasticities, returns to scale, and stability across specifications.[5] In particular, we consider five different variants of the proxy-variable approach:

1. Case 1 (Vanilla): This specification is exactly as in Ackerberg et al. (2015).

2. Case 2 (Exporting): Same as vanilla except we allow the law of motion for firm level productivity to depend on lagged export status. This controls for learning-by-exporting effects as in De Loecker and Warzynski (2012) and De Loecker (2013).

3. Case 3 (Attrition): Same as vanilla, except we include the Olley and Pakes (1996) selection correction terms to correct for attrition bias.

4. Case 4 (Over-identification): Same as vanilla, except we include lagged capital and lagged capital square as extra instruments.

5. Case 5 (Full Model): The case 2-4 modifications of case 1 are all introduced simultaneously.

[4]We measure labour input using employment and thus do not need a labour deflator. However, capital is simply measured in RMB.

[5]The Cobb-Douglas elasticities (coefficients on capital, labour and materials) are very similar to those reported in Brandt et al. (2014) and Brandt et al. (2017). We are extremely grateful to these authors for providing time-consistent definitions of firms as well as concordances, deflators, and capital stock adjustments. See Orr et al. (forthcoming).

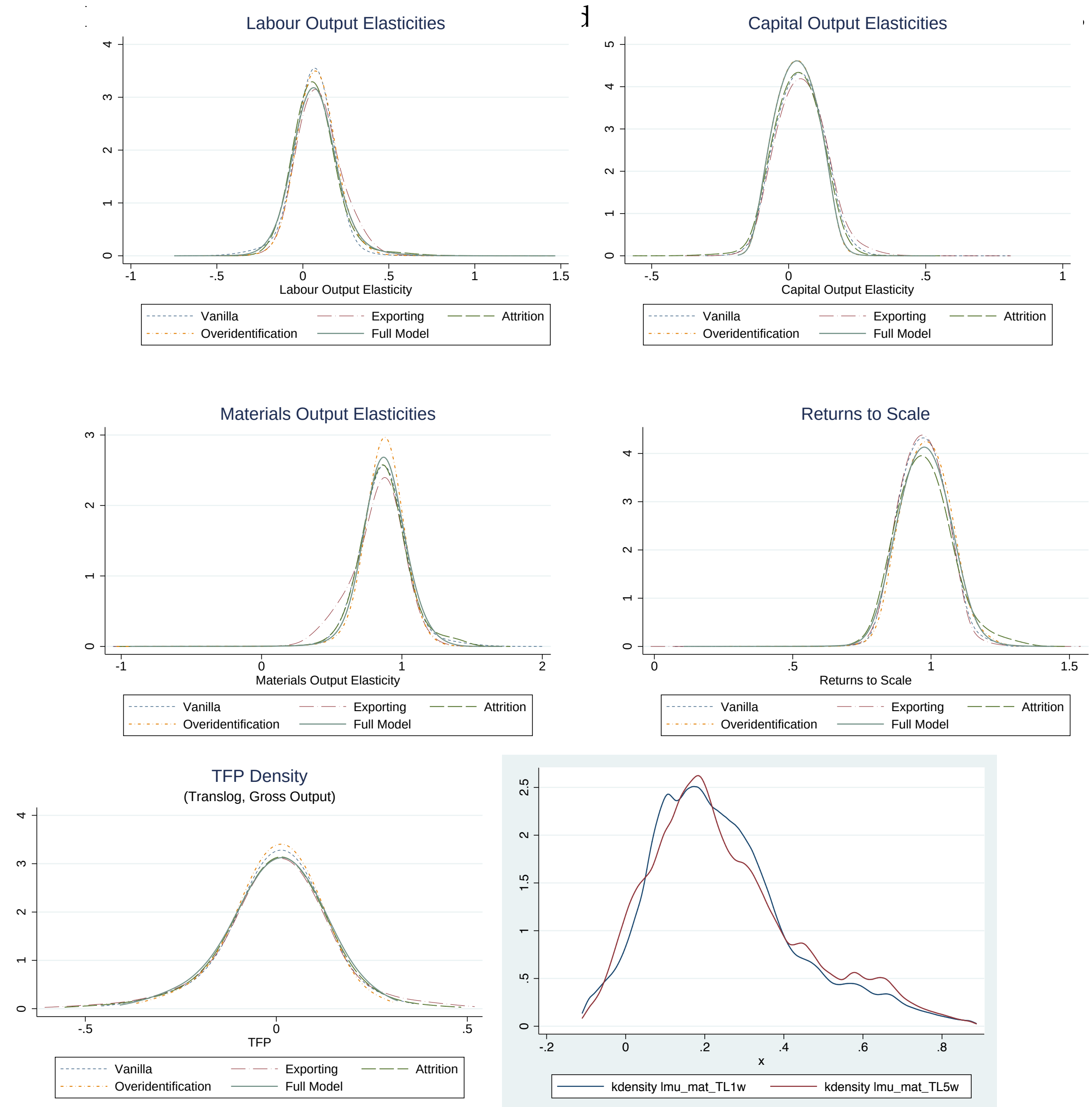
Figure 1 reports histograms for the elasticity of output with respect to labour, capital and materials. Each panel reports five histograms, one for each specification listed above and, as is apparent from the fact that the five histograms sit on top of each other, the choice of specification makes little difference. As is standard, the labour and capital output elasticities tend to be close to zero (and infrequently negative for some firm-year observations). The returns to scale tend to be strongly concentrated around 1, which is reassuring. Finally, there is little variation across specifications in the distribution of revenue TFP.

With revenue TFP estimates in hand we estimate markups using the approach in De Loecker and Warzynski (2012). Since labour shares are notoriously low in the CME (see e.g., Brandt et al., 2014), we follow Brandt et al. (2017) in basing markups on material inputs. These appear in the bottom right panel of table 1. We only report case 1 and case 5, but the other three cases are very similar. Note that the log markups are relatively close to 0, with most markups being less than 50%. This is much more sensible than the large markups reported in other research.

2.3 Quality

We will need quality to motivate our model. We note that our method only allows us to calculate quality for firms that are matched to the customs data because these are the only firms for which we have quantity and price (unit value) data rather than just revenue data. As a result, we use the quality data to motivate our results, but most of our empirical work will be based on the larger CME sample.

Are starting point is not the demand system of our theory, but the Berry (1994) method. Given the richness of our data, we are also able to improve on the implementation proposed by Khandelwal (2010) and to construct a novel instrument that avoids some of the criticisms of existing instruments (Ackerberg et al., 2007). Consider a Chinese firm f that exports an HS8 good h to destination d in year t. A market is a triplet (h, d, t) and let Ω_{hdt} be the set of firms selling into the market. In what follows, we will repeatedly see this market triplet. How much consumers in d buy will depend not just on prices in the market, but on outside options. For outside options, it will be enough here to model the upper-tier nests. Let $\mathbb{H}$ be an upper-tier nest, which in practice is an HS2 or HS4 category. Firm f's market share is $p_{fhdt}q_{fhdt}/\sum_{f'\in\Omega_{hdt}}(p_{f'hdt}q_{f'hdt})$. This is a core object in demand estimation. Interestingly, our rich data allow us to model the denominator $\sum_{f'\in\Omega_{hdt}}(p_{f'hdt}q_{f'hdt})$ as a fixed effect α_{hdt}. Notice also that Berry's random component of utility will also be subsumed by this fixed effect so that we do not have to estimate this



Notes: Each panel displays 5 histograms and each histogram corresponds to one of the 5 cases listed in the text. Letting β_{ij} be the coefficient in the translog production function on the interaction between the log of inputs i and j, the panels report labour output elasticities ($\beta_L + 2\beta_{LL}l_{it} + \beta_{LK}k_{it}$), capital output elasticities ($\beta_K + 2\beta_{KK}k_{it} + \beta_{LK}l_{it}$), materials output elasticity ($\beta_M + 2\beta_{MM}m_{it} + \beta_{LM}m_{it} + \beta_{KM}k_{it}$), returns to scale (the sum of the labour, capital and materials elasticities), revenue TFP, and log markups. Since output elasticities, returns to scale and revenue TFP vary across firm-year observations, the histograms are estimated from the 772,788 firm-year observations.

term. Thus, we are left with

$$\ln q_{fhdt} = \alpha_{hdt} + \beta \ln p_{fhdt} + \lambda^*_{fhdt} \tag{1}$$

where λ^*_{fhdt} is a measure of the quality of what firm f sells into market (h, d, t).[6]

Aggregating quality from the firm-market level (λ^*_{fhdt}) to the firm level is problematic because quality is never comparable across goods h. To partially address this, we demean quality using the average level of quality in market (h, d, t), i.e., we use $\lambda^*_{fhdt} - \overline{\lambda}^*_{hdt}$. We define a firm's quality in year t as

$$\lambda_{ft} \equiv \sum_{(h,d)} \omega_{fhdt} \left(\lambda^*_{fhdt} - \overline{\lambda}^*_{hdt} \right) \tag{2}$$

where ω_{fhdt} is Chinese firm f's exports in year t to market (h, d, t) as a share of its total exports in year t:

$$\omega_{fhdt} \equiv \frac{p_{fhdt} q_{fhdt}}{\sum_{(h',d')} p_{fh'd't} q_{fh'd't}}$$

We now turn to instruments. We need a pure supply shock and must avoid demand shocks. One common assumption is that supply shocks are spatially correlated while demand shocks are not. This leads to the Hausman-Nevo instrument which uses the prices of firms in nearby regions as an instrument for the firm's price. We do not think this is a good assumption in our context. Our rich data allow us to take a different approach that has not appeared in the literature. While a firm may be a large employer in its industry within a city, the firm is typically a small employer in its city overall. Consider a firm in a 2-digit CME industry in a city and calculate the average wage paid by firms in that city who are *not* in that industry. This is our instrument.

Finally, we follow Khandelwal (2010) in winsorizing price. We do so by first demeaning price within a market (h, d, t), then winsorizing prices above the 95th percentile and below the 5th percentile. Finally, we add the market mean back in.

The results appear in table 1. In our main results below we will define nests at the HS2 level, which has almost 100 products. In order to present the results more, clearly, here we first present results at a more aggregate level of HS sections. The table presents estimates of β in equation (1). Consider the first row, which pools all firms exporting chemicals (HS2 codes 28–38). The 'Second Stage' presents the IV estimate of β. 'OLS'

[6]We also exploit information about the mode of transportation m, e.g., air or waterborne. This amounts to treating the market not as an (h, d, t) tuplet but as an (m, h, d, t) tuplet. It makes no difference whether we aggregate over mode, but we think that an HS8 product shipped by air may be quite different than one shipped by sea. At any rate, this is a minor point empirically.

presents the OLS estimates of β. 'First Stage' and 'Reduced Form' are the coefficients on the instrument when the dependent variable are price and quantity, respectively. Notice that the IV estimates is always negative and more negative than the OLS estimate, as expected. The IV estimate is also almost always less than -1 which means that demand is elastic as required. Notice that there are large numbers of observations, large numbers of firms, and large numbers of markets. We can reject endogeneity ('K-P') and the joint hypothesis of endogeneity and the exclusion restriction ('A-R').

We now turn to the specifications that we use to generate our quality measures. We consider four specifications. We define the nest either at HS2 or HS4 and we either have just one instrument or consider a second instrument. The second instrument is the average number of export destinations per HS2 in a firm's city-year, excluding export destinations exported to by firms in target firm's own 2-digit CME industry. As in Melitz (2003) and Melitz and Ottaviano (2008), the more destinations exported to, the more productive is the region or the lower are the exporting fixed costs of the region, both of which are 'supply shocks.' For the case where we have one instrument and HS2 nests, 70 of 85 HS2 products have negative IV price elasticities. At the 5% significance level, 49 are negative and only 1 is positive.

Figure 2 shows the distribution of quality across all four specifications. As is apparent, they are very similar in distributions. Further, in the empirics to come, we get identical results for all four.

3 Econometric Results

Let f index firms, i index 4-digit CME manufacturing industries (of which there are about 450), c index cities (of which there are about 370) and t index years (2000-2006). We are interested in the impact of variables x_{ft} that measure scale and competition on outcomes y_{ft} that include quality, markups, RTFP, and innovation. We consider regressions of the form

$$y_{ft} = \alpha_f + \alpha_{it} + \alpha_{ct} + \beta x_{ft} + \delta W_{ft} + \varepsilon_{ft}\ . \tag{3}$$

α_f is a firm fixed effect. W_{ft} collects three time-varying firm characteristics: (1) a binary variable for whether the firm is a state-owned enterprise (SOE), (2) a binary variable for whether the firm has foreign investors, and (3) the average wage of firms in the same city as f, but not in the same 2-digit industry as f. The latter controls for cost shocks.[7] There

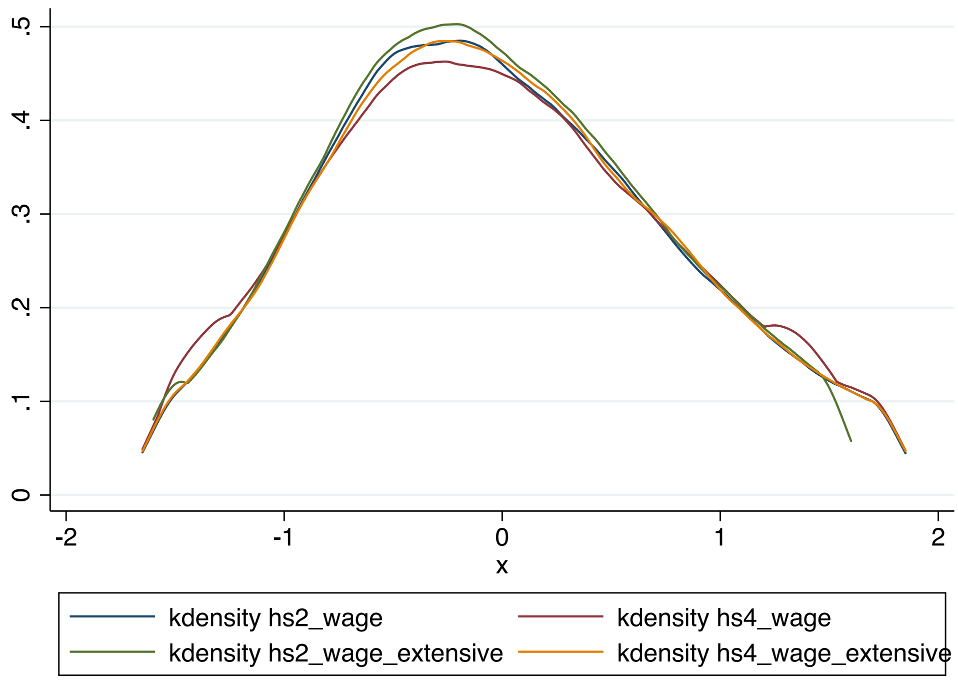
[7]We choose 2- rather than 4-digit because, for example, wage pressures likely apply to a broader 2-digit industry such as automotive rather than a narrower 4-digit industry such as a specific auto part.

Table 1: Demand Estimation and Quality

		Section (HS2)	Second Stage, DV: quantity, Log price coeff.		OLS, DV: quantity, Log price coeff.		First Stage, DV: price, Instrument coeff.		Reduced Form, DV: quantity, Instrument coeff.		Observations	Firms	No. of markets (h,d,t)	K-P	A-R
6	28-38	Chemicals	-1.08***	(0.18)	-0.86***	(0.024)	0.28***	(0.041)	-0.31***	(0.063)	455,689	5,594	43,818	47.13	0.000
7	39-40	Plastics	-1.75***	(0.22)	-1.02***	(0.028)	0.44***	(0.050)	-0.77***	(0.11)	614,486	12,626	26,832	79.85	0.000
8	41-43	Plastics	-2.11***	(0.53)	-0.70***	(0.024)	0.28***	(0.090)	-0.60***	(0.103)	226,476	4,576	6,160	9.95	0.000
9	44-46	Wood products	-2.89***	(0.63)	-0.57***	(0.044)	0.20***	(0.040)	-0.57***	(0.087)	142,445	2,846	6,696	23.79	0.000
10	47-49	Pulp and paper	-1.09***	(0.41)	-0.60***	(0.043)	0.36***	(0.061)	-0.39**	(0.154)	144,603	6,717	8,238	35.41	0.011
11	50-63	Textiles	-1.58***	(0.09)	-0.78***	(0.018)	0.43***	(0.028)	-0.68***	(0.053)	2,038,845	14,369	80,867	242.20	0.000
12	64-67	Footwear	-1.27***	(0.21)	-0.67***	(0.031)	0.40***	(0.072)	-0.51***	(0.105)	352,585	4,383	10,167	30.63	0.000
13	68-70	Non-ferrous prod.	-1.27***	(0.22)	-0.778***	(0.027)	0.32***	(0.050)	-0.41***	(0.080)	223,144	4,084	14,665	41.09	0.000
15	72-83	Ferrous metal prod.	-2.31***	(0.21)	-1.10***	(0.026)	0.35***	(0.035)	-0.80***	(0.089)	599,325	11,528	41,014	96.74	0.000
16	84-85	Machinery	-1.06***	(0.10)	-0.79***	(0.011)	0.77***	(0.050)	-0.82***	(0.097)	1,792,339	15,096	104,956	240.40	0.000
17	86-89	Transport equip.	-0.80**	(0.31)	-0.74***	(0.045)	0.39***	(0.065)	-0.31***	(0.119)	218,281	3,109	13,142	36.01	0.008
18	90-92	Optical	-0.93***	(0.11)	-0.70***	(0.017)	0.76***	(0.084)	-0.70***	(0.102)	306,112	4,517	23,028	81.00	0.000
20	94-96	Misc manuf.	-1.08***	(0.13)	-0.71***	(0.015)	0.41***	(0.042)	-0.44***	(0.063)	834,724	9,233	31,407	93.02	0.000

Notes: This table presents estimates of β in equation (1). Consider the first row, which pools all firms exporting chemicals (HS2 codes 28–38). The 'Second Stage' presents the IV estimate of β. 'OLS' presents the OLS estimates of β. The instrument Z is the average wage of firms in the same city but not in the same 2-digit CME industry. 'First Stage' is the coefficient on the instrument of a regression of price on Z. 'Reduced Form' is a regression of quantity on Z. Standard errors appear in parentheses and these are clustered at the firm level. 'K-P' is the Kleibergen-Paap F statistic for weak instruments. 'A-R' is the Anderson-Rubin p value for the joint null of weak instruments and the exclusion restriction.

Figure 2: Distribution of Quality across Four Specifications



Notes: This figure is a kernel density for four different quality measures. Quality is at the firm level (see equation 2) and there are 105,093 firm-year observations for each density. Two densities are based on HS2 and two on HS4. Two densities are based on the wage instrument and two are based on both the wage and extensive-margin instruments.

are likely pre-trends that we cannot control for because the data do not go back to a time when China had very little exposure to the international economy (say, 1990). Further, in an economy as dynamic as China's there are likely to be unobservable shocks that vary over time and across both industries and cities. It is thus essential to control for industry trends and local (city) trends as much as possible. Possible city trends include differential migration patterns and differential improvements in local amenities such as universities. We therefore include industry-year fixed effects α_{it} at the most detailed level available (4-digit industry) as well as city-year fixed effects α_{ct}. Such extensive fixed effects are uncommon in research using Chinese firm-level data.[8]

Throughout, we cluster standard errors by firm so as to allow for serial correlation. We have also briefly experimented with two-way clustering and initial results suggest that this makes no difference.

3.1 Scale: Exporting

In this section we examine an increase in the size of the market a firm faces. There are two ways in which a firm might experience an increase in the demand for its products. The first comes from an increase in the domestic or Chinese demand for its products, the second from foreign demand. Our data record each firm's total sales and exports at the 4-digit industry level so that we can compute a firm's domestic sales. Unfortunately, these domestic sales are subsumed in our industry-year fixed effects α_{it}.[9]

Firm-level exports X_{ft} provide sample variation that goes beyond 4-digit industries both because some of the export data are at the HS8 level and because even within a 4-digit industry not all firms export. Exports are not perfect for our needs because they are in value terms and thus capture both the scale of the foreign market and the markups charged in the foreign market, i.e., X_{ft} is not a pure measure of scale or competition as defined above, but is a combination of the two. That said, the scale component likely dominates the competition component by orders of magnitude. A simple calculation suggests that 83% of the vast growth in Chinese exports is due to growth in quantities

[8]We do not include industry-year fixed effects α_{ic} because these are largely subsumed by the firm fixed effects i.e., there are relatively few firms that switch industry or city. Preliminary research indicates that none of our results change if in addition we add industry-city-year fixed effects where the industry is at the 2-digit level. This lack of change mitigates some if not most of our concerns about pre-trends and unobservables.

[9]An alternative approach is to find a time- and firm-varying variable. One such variable is the a count of the number of kilometres of arterial highway within 20 kilometres of the firm's address. As these highways develop, the firm experiences a firm-specific increase in its market potential. See Liu et al. (2018). Initial explorations of the highway data point to the presence of scale effects from domestic market size.

rather than markups. This should come as no surprise.[10]

A firm's exports are endogenous and we create a unique instrument comprised of two elements, a standard Bartik instrument and a predicted probability of exporting. We first describe the Bartik component. Consider a firm that first exported in year t_0. Let ω_{fht_0} be the share of the firm's export sales in year t_0 that are accounted for by good h. If the firm is matched to the customs data then h is an HS8 product. If not, then h is a 4-digit industry code and $\omega_{fht_0} = 1$ if the firm is in industry h and equals 0 otherwise. Let m_{ht} be world imports of Chinese good h in year t. This is a measure of the demand shock for Chinese good h in year t: The larger is m_{ht}, the larger is a Chinese firm's potential foreign market. Aggregating up from the product level to the firm level using the firm's initial product weights yields:

$$\sum_h \omega_{fht_0} m_{ht} \ . \tag{4}$$

This is the Bartik component.[11]

Whether the firm takes advantage of foreign demand shocks depends on the fixed and marginal costs of exporting. We assume that the probability a firm exports depends on its current RTFP (as in Melitz 2003), on its past history of exporting (as in Roberts and Tybout 1997), and on its city (some cities are closer to ports than others). For a firm in year t and city c, we compute the firm's percentile in the distribution of RTFP among firms in that year in that city. Denote this percentile by $perc^{RTFP}_{ft}$.[12] In the first year we observe the firm (year t_0) we regress an exporter dummy on $perc^{RTFP}_{ft_0}$. This generates a prediction $p^X_{ft_0}$ that the firm exports in year t_0. In subsequent years ($t > t_0$), we regress the exporter dummy on both $perc^{RTFP}_{ft}$ and a dummy for whether the firm exported in the previous year. This generates a prediction p^X_{ft} that the firm exports in year $t > t_0$.

[10]Denote export quantities, prices, and costs by q, p, and c , respectively. Export values are $X = p \cdot q$ and markups are $\mu = p/c - 1$. Thus, $\Delta \ln X = \Delta \ln q + \Delta \ln p = \Delta \ln q + \Delta \ln(p/c) + \Delta \ln c = \Delta \ln q + \Delta \ln(\mu + 1) + \Delta \ln c$. In our data between 2000 and 2006, $\Delta \ln X = 1.44$ and $\Delta \ln(\mu + 1) = 0.05$. Further, the input with by far the largest cost share is materials and its deflator rose by 0.14 log points i.e., $\Delta \ln c = 0.14$. Hence $\Delta \ln q = 1.44 - 0.05 - 0.14 = 1.25$ and $\Delta \ln q / \Delta \ln X = 0.83$. That is, the dominant driver of exports has been quantities (scale) rather than markups (competition) or input costs.

[11]We use world imports from China. We could also consider world imports from countries that are 'similar' to China. The idea is that this would net out China-specific costs shocks. **We have experimented with this and obtained similar results.** However, there are limitations of this approach. First, the most similar countries to China are low- and middle-income countries such as Viet Nam and Hungary. But as Schott (2008) and Sutton and Trefler (2016) have documented, 2000–2006 world imports from these countries *combined* covered a much smaller range of goods than world imports from China. (In the formula above, m_{ht} was zero for a majority of h.) Restated, these countries are not similar to China. Second, our industry-year fixed effects and city-year fixed effects capture most if not all of the China-specific costs shocks that lead to exclude China when computing the demand shock. For example, suppose a Chines industry improves its technology rapidly or a Chinese city builds a technology park. These are subsumed in the industry-year and city-year fixed effects.

[12]We suppress the city subscript because firms do not switch cities so that c is implied by f.

Multiplying the probability that a firm exports by the size of foreign demand shock gives us our instrument for the log of firm exports $\ln(1 + X_{ft})$:

$$Z_{ft} \equiv p^X_{ft} \;\cdot\; \ln\left(1 + \sum\nolimits_h \omega_{fht_0} M_{ht}\right) \tag{5}$$

3.1.1 Scale: Impact of Exporting on Quality, Markups and RTFP

Armed with this instrument, we present the estimates. The first two columns of table 2 display OLS and IV estimates of equation (3) where the dependent variable is quality. There are 78,285 firm-year observations.[13] There are 25,406 firm fixed effects, 1,949 industry-year fixed effects and 1,329 city-year fixed effects. Standard errors clustered by firm are reported in parentheses. The key independent variable is the log of exports $\ln(1 + X_{ft})$. Its coefficient is statistically significant and positive in both the OLS and IV specifications. The IV result implies that increased exporting causally leads to higher quality. IV is smaller than OLS, as would be expected if higher quality caused more exporting.

The first stage appears in the bottom panel. This is a regression of log exports on the instrument Z_{ft} and all the exogenous second-stage variables. The instrument coefficient of 0.021 has a small standard error of 0.0013. The panel also reports the Kleibergen-Paap F statistic for weak instruments (which should be greater than 20) and the Anderson-Rubin p value for the joint null of weak instruments and the exclusion restriction (it should be less than 0.01).[14]

The 'RF' column reports the reduced form. It is an 'intention to treat' that captures the impact of the instrument directly on the outcome. The instrument is significant and positive, as expected.

The next three columns of table 2 report results for the log of markups. The dataset is much larger, with 791,229 firm-year observations spread across 222,428 firms, 3,395 industry-year pairs, and 2,380 city-year pairs. The IV and OLS results are very similar, the first stage is strong, and the reduced form is positive as expected. This leads us to conclude that an increase in exporting induced by a positive foreign demand shock leads to increased markups.[15]

[13] This is a restricted sample because quality is only available for firms that have been matched to the customs data. Also note that in reporting the number of observations, we drop firms that only appear in one year.

[14] We apologize that, due to time constraints, we are not reporting on coefficient magnitudes.

[15] Of lesser importance, the coefficient on city wage is negative. Recall that in each year t, this is the

Table 2: Dependent Variables: Quality, Markups and RTFP

	Quality			Markups			RTFP		
	OLS	IV	RF	OLS	IV	RF	OLS	IV	RF
ln (1+*Exports*)$_{ft}$	0.262***	0.090***		0.033***	0.036***		-0.072***	-0.209***	
	(0.003)	(0.032)		(0.004)	(0.012)		(0.004)	(0.015)	
Instrument for *ln* (1+*Exports*)$_{ft}$			0.0019***			0.007***			-0.038***
			(0.0007)			(0.002)			(0.003)
Controls (f,t)									
SOE Status	0.061	0.051	0.047	-0.587***	-0.588***	-0.585***	-0.101	-0.088	-0.105
	(0.045)	(0.044)	(0.046)	(0.064)	(0.064)	(0.064)	(0.083)	(0.084)	(0.083)
Foreign Invested	0.004	0.005	0.004	-0.110**	-0.111**	-0.094**	-0.181***	-0.107*	-0.210***
	(0.009)	(0.009)	(0.010)	(0.045)	(0.046)	(0.045)	(0.060)	(0.060)	(0.060)
City Wage	0.199	0.177	0.162	-1.902***	-1.925***	-1.914***	-2.300***	-2.250***	-2.317***
	(0.291)	(0.304)	(0.320)	(0.386)	(0.386)	(0.386)	(0.504)	(0.504)	(0.504)
R^2	0.830	0.813	0.791	0.902	0.902	0.902	0.881	0.881	0.881
# observations	78,285	78,285	78,285	791,229	791,229	791,229	791,229	791,229	791,229
# firm f FEs		25,406			222,428			222,428	
# year-ind4 (i,t) FEs		1,949			3,395			3,395	
# year-city (c,t) FEs		1,329			2,380			2,380	
First Stage									
Instrument for *ln* (1+*Exports*)$_{ft}$		0.021***			0.182***			0.182***	
		(0.0013)			(0.0022)			(0.0022)	
Kleibergen-Paap (F)		260			6,757			6,757	
Anderson-Rubin (p)		0.008			0.004			0.000	

Notes: This table presents estimates of equation (3) where the dependent variable is quality, log markups or RTFP. For each dependent variable, there are three columns: OLS, IV, and reduced form. Exports $\ln(1+X_{ft})$ are the endogenous variable. The first stage appears in the bottom panel where the coefficient on the equation (5) instrument for $\ln(1+X_{ft})$ is reported. The panel also reports the Kleibergen-Paap F statistic for weak instruments (we require $F > 20$) and the Anderson-Rubin p value for the joint null of weak instruments and the exclusion restriction (we require $p < 0.010$). All specifications have firm fixed effects, 4-digit industry-year fixed effects, and city-year fixed effects. 'SOE Status' is a binary variable equalling one if the firm is a state-owned enterprise. 'Foreign Invested' is a binary variable equalling one if the firm is at least partially foreign-owned. For each year t, (City Wage)$_{ft}$ is the average wage of firms in the same city as f, but *not* in the same two-digit industry as f. The coefficients and standard errors in the log markup and RTFP columns are multiplied by 100. Standard errors appear in parentheses and are clustered at the firm level.

The last three columns of table 2 report results for revenue TFP. If higher RTFP causally leads to more exporting then we expect the estimated coefficient on exporting in the RTFP equation to be smaller (in this case more negative) for IV than OLS. This is what we find. The surprising finding is that differential increases in exporting lead to differential decreases in RTFP, contrary to what Lileeva and Trefler (2010) and others have found. However, since exporting raises quality, a likely explanation is that higher-quality goods are more technically difficult to produce and so exporting reduces RTFP by shifting the firm's focus away from cost cutting and towards higher quality.[16]

3.1.2 A Diff-in-Diff Caveat: Levels versus Relatives

We conclude this discussion with a very important caution. Our econometric estimates are essentially difference in difference estimates i.e., they compare firms whose exports have increased with those whose exports have not increased. We thus cannot answer questions about the *level* effects of exporting on quality, markups, or RTFP. All we can speak to is the *differential* effect of exporting on quality, markups and RTFP. Thus, for example, we cannot conclude from the negative RTFP coefficient that exporting lowers the level of RTFP. Discussion of level effects must await the calibrated model.

3.1.3 Scale: The Impact of Exporting on Innovation

Since so many firms report zero for one or more of our three measures of innovation, we start by computing the principal component of the three measures: (1) number of patents, (2) R&D as a share of sales, and (3) new product sales as a share of total sales. The principal component is estimated separately for each 2-digit industry. Table 3 reports the results of using this principal component as a dependent variable. The first stage is again strong, the OLS and IV results are very similar, and the OLS, IV and reduced-form results are all positive and statistically significant. This means that exporting has a positive causal impact on innovation. Again, this is a differential effect rather than a level

---

wage paid by firms in the same city as firm f, but not in the same 2-digit industry as firm f. The fact that the coefficient is negative suggests that, even after controlling for city-year and industry-year fixed effects, there are some industry-city-year unobservables that matter for markups. These unobservables may have to do with wages of workers with certain skills or the price of certain types of industrial land. Note that none of our results change when we exclude the city wage variable, suggesting that the unobservables correlated with city wage are not correlated with our dependent variables. Nor do results change when 2-digit industry-city-year fixed effects are added.

[16]Alternatively, higher quality requires more expensive inputs and, absent firm-level output and input price deflators, quality may lead to downward bias in the measurement of RTFP. This happens if lack of firm-level input deflators dominates lack of firm-level output deflators.

effect.

To investigate further, table 4 reports the results separately for each of the three measures of innovation. Exporting has a positive causal effect on all three measures.

Table 3: Principal Component of Innovation

	Innovation - Principal Component		
	OLS	IV	RF
$ln(1+Exports)_{ft}$	0.010***	0.013***	
	(0.0006)	(0.0018)	
Instrument for			0.0026***
$ln(1+Exports)_{ft}$			(0.00037)
Controls (f,t)			
SOE Status	0.042***	0.042***	0.043***
	(0.009)	(0.009)	(0.009)
Foreign Invested	0.037***	0.036***	0.042***
	(0.007)	(0.007)	(0.007)
City Wage	0.078*	0.078*	0.079*
	(0.046)	(0.046)	(0.046)
R^2	0.662	0.662	0.662
# observations	762,702	762,702	762,702
# firm f FEs		247,298	
# year-ind4 (i,t) FEs		2,923	
# year-city (c,t) FEs		2,040	
First Stage			
Instrument for		0.205***	
$ln(1+Exports)_{ft}$		(0.0024)	
Kleibergen-Paap (F)		7,266	
Anderson-Rubin (p)		0.000	

Notes: This table presents estimates of equation (3) where the dependent variable is the principal component of (1) number of patents, (2) R&D as a share of sales, and (3) new product sales as a share of total sales. The principal component is calculated separately by 2-digit industry. The table has the same structure as table 2: See the notes to that table for details. Standard errors appear in parentheses and are clustered at the firm level.

Table 4: Patenting, R&D and New Product Sales

	Patents			R&D / Sales			(New Product Sales) / Sales		
	OLS	IV	RF	OLS	IV	RF	OLS	IV	RF
ln (1+*Exports*)$_{ft}$	0.942***	2.125***		0.372***	0.624***		0.122***	0.101***	
	(0.121)	(0.375)		(0.052)	(0.185)		(0.008)	(0.025)	
Instrument for *ln* (1+*Exports*)$_{ft}$			0.437***			0.128***			0.021***
			(0.077)			(0.038)			(0.005)
Controls (f ,t)									
SOE Status	-1.658	-1.785	-1.594	8.630***	8.603***	8.659***	-0.070	-0.067	-0.058
	(1.242)	(1.242)	(1.241)	(1.124)	(1.124)	(1.125)	(0.123)	(0.123)	(0.123)
Foreign Invested	-2.268*	-2.922**	-1.901	6.243***	6.104***	6.403***	-0.037	-0.025	0.022
	(1.298)	(1.309)	(1.298)	(0.733)	(0.742)	(0.733)	(0.101)	(0.102)	(0.101)
City Wage	0.056	-0.074	0.137	5.560	5.532	5.594	1.146	1.148	1.158
	(7.820)	(7.818)	(7.821)	(4.134)	(4.134)	(4.135)	(0.730)	(0.730)	(0.731)
R^2	0.629	0.629	0.629	0.570	0.570	0.570	0.687	0.687	0.687
# observations	762,702	762,702	762,702	762,702	762,702	762,702	762,702	762,702	762,702
# firm f FEs		247,298			247,298			247,298	
# year-ind4 (i ,t) FEs		2,923			2,923			2,923	
# year-city (c ,t) FEs		2,040			2,040			2,040	
First Stage									
Instrument for *ln* (1+*Exports*)$_{ft}$		0.205***			0.205***			0.205***	
		(0.0024)			(0.0024)			(0.0024)	
Kleibergen-Paap (F)		7,266			7,266			7,266	
Anderson-Rubin (p)		0.000			0.001			0.000	

Notes: This table presents estimates of equation (3) where the dependent variable is either (1) number of patents, (2) R&D as a share of sales, or (3) new product sales as a share of total sales. See table 2 for details. The coefficients and standard errors in the R&D and new sales columns are multiplied by 100. Standard errors appear in parentheses and are clustered at the firm level.

3.1.4 Scale: Heterogeneous Impacts of Exporting

Finally, we explore how the causal impacts of exporting vary by firm characteristics. Here we follow Aghion et al. (2017) in placing firms into quintiles based on sales and then re-estimating the above models separately by quintiles. Quintiles are computed by 4-digit industry using sales in the first year the firm is observed. Quintile 1 is small sales and quintile 5 is high sales. The results appear in table 5. Only the IV estimate of the coefficient on exports $\ln(1 + X_{ft})$ are reported.

Before examining the results a caveat about the quality results is needed. While for the most part a quintile has one fifth of the observations, this is not the case for quality where there are very few small firms. The quality results are thus not trustworthy. Indeed, they are never significant.

Looking across the other dependent variables, it is clear that there is substantial heterogeneity. This will be an important feature of the model below. In addition, larger firms experience a relatively more positive impact of exporting on innovation, log markups, and RTFP.[17] This has a clear LATE interpretation: exporting is done by the firms which expect a larger positive impact from exporting. This type of heterogeneity is consistent with the heterogeneity discussed in Lileeva and Trefler (2010).

[17]We repeat that the level of the coefficients less important than the differential effect across deciles.

Table 5: Heterogeneous Impacts of Exporting

	$\ln(1+\text{Exports}_{ft})$		Std. Err.	Obs	R^2	K-P (F)	A-R (p)
PC - Innovation							
Quintile 1	-0.002		0.008	130,006	0.615	384	0.82
Quintile 2	0.008		0.005	144,955	0.621	723	0.14
Quintile 3	0.005		0.004	153,796	0.632	1,202	0.26
Quintile 4	0.009	***	0.003	159,665	0.648	1,890	0.01
Quintile 5	0.020	***	0.003	166,414	0.728	2,859	0.00
Quality							
Quintile 1	0.060		0.278	3,649	0.864	4	0.83
Quintile 2	0.029		0.124	8,264	0.833	18	0.82
Quintile 3	0.060		0.086	12,431	0.807	45	0.51
Quintile 4	0.100		0.072	18,718	0.823	54	0.20
Quintile 5	0.038		0.062	31,480	0.825	69	0.55
Markups							
Quintile 1	-0.200	***	0.063	133,062	0.865	395	0.00
Quintile 2	-0.050		0.041	152,876	0.906	636	0.22
Quintile 3	0.030		0.030	160,819	0.913	1,107	0.31
Quintile 4	0.049	**	0.023	166,137	0.917	1,708	0.03
Quintile 5	0.083	***	0.020	169,900	0.921	2,636	0.00
RTFP							
Quintile 1	-0.492	***	0.072	133,062	0.875	395	0.00
Quintile 2	-0.289	***	0.050	152,876	0.879	636	0.00
Quintile 3	-0.219	***	0.036	160,819	0.872	1,107	0.00
Quintile 4	-0.186	***	0.029	166,137	0.870	1,708	0.00
Quintile 5	-0.170	***	0.025	169,900	0.890	2,636	0.00

Notes: This table presents estimates of equation (3) for various dependent variables: the principal component of innovation, quality, log markups, and RTFP. What is new is that the equation is estimated separately by quintiles of the distribution of firm sales. Quintile 1 is the smallest sales. Each row is an IV regression and only the IV coefficient on $\ln(1 + X_{ft})$ is reported. The coefficients and standard errors for log markups and RTFP are multiplied by 100. Standard errors appear in parentheses and are clustered at the firm level.

3.2 Competition 1: Competition in China's Domestic Markets

Typical measures of competition include concentration ratios, the Herfindahl index, the Lerner price-cost index, and measures of entry. All of these are industry-year specific and so are subsumed within our industry-year fixed effects. To examine the role of competition we therefore need measures that have firm-level element. To this end, we classify firms by their deciles within the distribution of some firm characteristics such as sales and interact this firm-specific variable with industry-year competition measures. This gives us a triple difference in which we examine effects of competition by small and large firms.

Two choices must be made. First, we use sales as the firm characteristics. However, preliminary results point to similar results using markups or RTFP in place of sales. Second, we use firm entry as our measure of competition. Again, preliminary results point to similar results using the 4-firm concentration ratio and the Herfindahl index. We prefer to use entry because this is what is most salient in light of the results by Brandt et al. (2017) on the importance of entry for understanding productivity dynamics in China.

Turning to the details of constructing our competition measure, we start by calculating sales deciles separately by 4-digit industry and year. Let $i(f)$ be the 4-digit industry of firm f in the year the firm is first observed, let $G_{i(f)}$ be the distribution of sales for those firms who are in $i(f)$, let d index deciles, let d_f be firm f's decile, and let D_f^d be a binary variable that equals one if firm f was in the dth decile of $G_{i(f)}$ in its first year i.e., if $d = d_f$. Next, we calculate the number of active firms by 4-digit industry and year (N_{it}) and interact it with the firm's decile to obtain a firm-specific measure of competition for firm f in year t. We can do this in two ways. One is to include 10 regressors of the form $N_{it} \cdot D_f^d$, $d = 1, \ldots, 10$. A more parsimonious approach which is valuable if the coefficients on $N_{it} \cdot D_f^d$ are monotonic in d is to use $N_{it} \cdot d_f$. Finally, to emphasize that the level effect is not identified because it is subsumed in the industry-year fixed effects, we use $N_{it} \cdot (d_f - 1)$ rather than $N_{it} \cdot d_f$ so that by construction the regressor is zero for the first decile.[18]

N_{it} is exogenous to the firm and D_f^k is predetermined so that we treat $N_{it} \cdot D_f^k$ is exogenous, i.e., no instrument is needed. Co-moving unobservables are a threat to identification so once again we include the most detailed possible industry-year and city-year fixed effects. Identification is threatened only by unobservables that operate at the industry-city-year level and that are not picked up by our city-industry wage variable.

Table 6 reports the IV results. Consider the 'Quality' column of the top panel. The IV coefficient on $N_{it} \cdot (d_f - 1)$ is positive for quality, indicating that larger firms respond to

[18] Two observations: (1) Results using $\ln N_{it}$ in place of N_{it} are identical. (2) With firm fixed effects, using N_{it} is similar to using $(N_{it} - N_{i,t-1}) \cdot D_f^d$ or $(\ln N_{it} - \ln N_{i,t-1}) \cdot D_f^d$. Indeed, we obtain similar results using these variables. The model below points to using N_{it} rather than $N_{it} - N_{i,t-1}$.

Table 6: Impacts of Competition in Chinese Markets on Quality, Markups, and RTFP

	Quality		Markups		RTFP	
N_{it} x (d_f - 1)	0.091	***	0.006	***	-0.013	***
	(0.0055)		(0.0002)		(0.0002)	
Controls (f,t)						
SOE Status	0.046		-0.006	***	-0.001	
	(0.0459)		(0.0006)		(0.0008)	
Foreign Invested	0.008		-0.001	***	-0.002	***
	(0.0102)		(0.0004)		(0.0006)	
City Wage	0.114		-0.019	***	-0.024	***
	(0.3203)		(0.0038)		(0.0050)	
R^2	0.793		0.904		0.883	
# observations	78,285		791,229		791,229	
# firm f FEs	25,406		222,428		222,428	
# year-ind4 (i,t) FEs	1,949		3,395		3,395	
# year-city (c,t) FEs	1,329		2,380		2,380	
Interacting N_{it} with 10 Decile Dummies (D_f^d)						
N_{it} x D_f^1	0.00		0.00		0.00	
N_{it} x D_f^2	0.06		-0.02	**	0.00	
N_{it} x D_f^3	0.17	***	-0.03	**	0.00	
N_{it} x D_f^4	0.28	***	-0.03	**	-0.01	**
N_{it} x D_f^5	0.34	***	-0.03	**	-0.02	**
N_{it} x D_f^6	0.43	***	-0.02	**	-0.03	***
N_{it} x D_f^7	0.51	***	-0.01	*	-0.05	***
N_{it} x D_f^8	0.57	***	0.01	*	-0.07	***
N_{it} x D_f^9	0.71	***	0.04	**	-0.09	***
N_{it} x D_f^{10}	0.84	***	0.08	**	-0.13	***

Notes: This table presents estimates of equation (3) for the dependent variables listed at the top of the table. Each column presents OLS estimates of equation (3) for two different specifications. In the first specification, the top panel, the key independent variable is the number of firms in the industry-year (N_{it}) interacted with the firm's decile in the distribution of sales, $d_f = 1, \ldots, 10$. More precisely, it is $N_{it} \cdot (d_f - 1)$ so that by construction the first decile has a response normalized to 0. This is because only the cross-decile differential effect is identified, not the level effect. Decile 1 contains the smallest firms. In the bottom panel, the key independent variables are N_{it} interacted with a set of 10 binary variables D_f^d where $D_f^d = 1$ if firm f is in decile d. The coefficients and standard errors for log markups amd RTFP are multiplied by 100. Standard errors appear in parentheses and are clustered at the firm level.

competition by raising quality more (or by lowering quality less) than small firms.[19] The bottom panel shows that this effect is monotonic across deciles.

The RTFP column also shows substantial and monotonic heterogeneity: In response to competition, larger firms reduce RTFP by as much as (or increase RTFP by less than) smaller firms. This is consistent with the idea that in response to competition, firms upgrade quality at the cost of lower RTFP. We will allow for this heterogeneity in our model.

The markup column shows that there is substantial heterogeneity in markup responses to competition, but this heterogeneity is not monotonic. We will thus need to think about a model which allows for non-monotonic markup responses.

Table 7 has the same structure as table 6, but with dependent variables that are related to innovation. We see heterogeneity in responses across all the innovation measures. In response to competition, larger firms either increase competition by more than smaller firms or reduce innovation by less than smaller firms. The heterogeneity is more pronounced for some measures of innovation than for others and, not unexpectedly, patent heterogeneity is particularly skewed in the 10th decile.

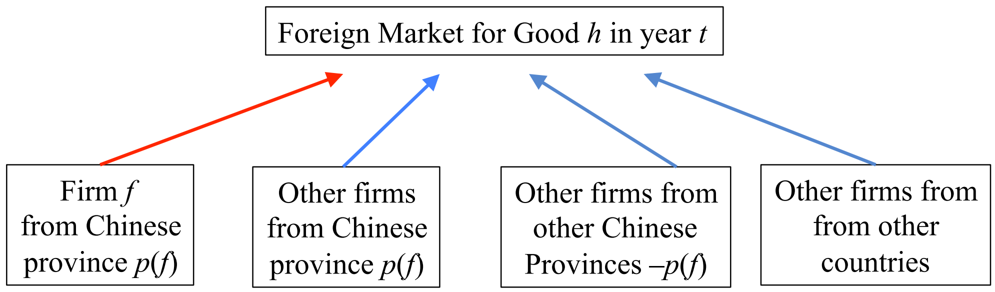
[19]The expression in parentheses is a reflection of the fact that the econometrics do not deliver level effects.

Table 7: Impacts of Competition in Chinese Markets on Innovation

	Innovation: Principal Component	Patents	R&D / Sales	New Product Sales / Sales
N_{it} x (d_f - 1)	0.023 ***	0.024 ***	0.005 ***	0.210 ***
	(0.0021)	(0.0023)	(0.0013)	(0.0323)
Controls (f,t)				
SOE Status	0.043 ***	-0.016	0.087 ***	-0.064
	(0.0086)	(0.0124)	(0.0112)	(0.1238)
Foreign Invested	0.042 ***	-0.019	0.064 ***	0.020
	(0.0068)	(0.0129)	(0.0073)	(0.1010)
City Wage	0.081 *	0.003	0.056	1.173
	(0.0462)	(0.0782)	(0.0413)	(0.7313)
R^2	0.662	0.629	0.570	0.687
# observations	762,702	762,702	762,702	762,702
# firm f FEs	247,298	247,298	247,298	247,298
# year-ind4 (i,t) FEs	2,923	2,923	2,923	2,923
# year-city (c,t) FEs	2,040	2,040	2,040	2,040
Interacting N_{it} with				
N_{it} x D_f^1	0.00	0.00	0.00	0.00
N_{it} x D_f^2	0.05 ***	0.01 **	0.03 ***	0.30
N_{it} x D_f^3	0.08 ***	0.02 ***	0.02 **	0.83 ***
N_{it} x D_f^4	0.07 ***	0.02 ***	0.02 ***	0.64 ***
N_{it} x D_f^5	0.07 ***	0.04 ***	0.03 ***	0.59 **
N_{it} x D_f^6	0.08 ***	0.04 ***	0.02 **	0.79 ***
N_{it} x D_f^7	0.10 ***	0.06	0.03 ***	0.94 ***
N_{it} x D_f^8	0.11 ***	0.10 ***	0.02 **	0.96 ***
N_{it} x D_f^9	0.17 ***	0.13	0.04 ***	1.55 ***
N_{it} x D_f^{10}	0.35 ***	0.35 ***	0.09 ***	3.02 ***

Notes: This table has the exact same structure as table 6, except that the dependent variables are different. These are listed in the header. The coefficients and standard errors for R&D and new sales are multiplied by 100. See the notes to table 6 for details. Standard errors appear in parentheses and are clustered at the firm level.

Figure 3: Competition 2 Measure



3.3 Competition 2: Competition in Foreign Markets

In this section we turn to an alternative measure of competition that has not been considered before. We examine the impact of Chinese firms competing with each other in foreign markets. To motivate our approach consider a world geography in which there are countries and, within China, there are provinces. We treat countries and Chinese provinces as the geographic units so that Chinese provinces are like countries in the model, meaning, each has its own unique cost structures. This is illustrated in figure 3. There is a foreign market for some good h in year t. We are interested in the competition that a target firm f in province $p(f)$ faces when exporting into this foreign market. This competition comes from firms in the same province $p(f)$, from firms in other provinces $-p(f)$, and from firms in other countries. As in our model, the degree of competition comes from underlying cost differences in the various provinces and countries. For firms in other countries, the data are not at the firm level and so are subsumed into the industry-year fixed effects. For firms from the same province, they compete with f but because the also share observed and unobserved attributes common to all firms in province $p(f)$, their exports into the foreign market will capture common unobserved shocks and not just competition. For firms in other provinces $-r(f)$ their exports into the foreign market will be closer to a pure competition effect.

Operationally, let $X_{-f,ht}$ be the exports of good h in year t summed over all Chinese firms who are *not* in firm f's province. Then:

$$(Competition\ into\ Foreign\ Markets)_{ft} \equiv \ln\left(1 + \sum_h \omega_{fht_0} X_{-f,ht}\right) \tag{6}$$

where ω_{fht_0} was introduced in the discussion leading up to equation (4). For brevity, we referred to this more simply as $Export\ Competition_{ft}$.

Table 8 presents the results of regressing our standard dependent variables on $Export$

Table 8: Impacts of Competition in Foreign Markets

	Quality	Markups	RTFP	Innovation: Principal Component	Patents	R&D / Sales	(New Product Sales) / (Sales)
Export Competition $_{ft}$	0.036***	-0.017**	-0.025***	-0.017***	-2.469***	-0.013***	-0.132***
	(0.005)	(0.008)	(0.008)	(0.002)	(0.450)	(0.002)	(0.030)
Controls (f,t)							
SOE Status	0.044	-0.627***	0.244**	0.044***	-1.707	0.084***	-0.024
	(0.044)	(0.102)	(0.106)	(0.008)	(1.226)	(0.011)	(0.123)
Foreign Invested	-0.000	-0.028	0.133**	0.037***	-1.622	0.057***	-0.013
	(0.009)	(0.068)	(0.062)	(0.006)	(0.012)	(0.007)	(0.101)
City Wage	0.176	-3.334***	-2.187***	0.069	-0.002	0.035	1.140
	(0.309)	(0.601)	(0.685)	(0.045)	(0.076)	(0.040)	(0.715)
R^2	0.787	0.963	0.897	0.659	0.624	0.568	0.683
# observations	84,771	633,811	633,811	781,137	781,137	781,137	781,137
# firm f FEs	27,174	195,721	195,721	251,005	251,005	251,005	251,005
# year-ind4 (i, t) FEs	1,989	2,409	2,409	2,964	2,964	2,964	2,964
# year-city (c,t) FEs	1,354	2,360	2,360	2,041	2,041	2,041	2,041

Notes: Each column is an OLS estimate of equation (3). The dependent variable is listed in the header. The key independent variable is export competition defined in equation (6). The coefficients and standard errors for log markups, RTFP, R&D and new sales are multiplied by 100. Standard errors appear in parentheses and are clustered at the firm level.

$Competition_{ft}$. Firms that face stiffer competition in foreign markets from Chinese firms in other provinces respond by increasing quality at the cost of lower RTFP, by reducing markups, and by reducing all measures of innovation.

3.4 Scale and Competition: Econometric Conclusions

We have documented very significant effects of scale and competition, the latter including competition both within China and into foreign markets. Scale and competition impacted quality, RTFP, markups, and three measures of innovation. Further, we documented very substantial degrees of response heterogeneity across firms of different sizes. These econometric results are in need of a model for two reasons. First, not all of the results are obvious: a model is needed for thinking about the form of the heterogeneity. Second and perhaps more importantly, the econometric estimates are difference-in-difference estimates of *relative effects*. A calibrated model is needed to investigate *level effects*.

4 Closed Economy Model

Motivated by the econometric results discussed above, we now develop and study a structural model of trade and innovation. This structural approach will offer several additional insights. First, it formalizes a theory of how firms might innovate to escape changes in competition induced by trade. Second, the calibrated model offers a quantification of the effects of trade on innovation, and of how the levels of these effects vary across firms. Third, the model serves to formally decompose the effects of trade on innovation into the scale and competition channels highlighted by our econometric results.

4.1 Model environment

4.1.1 Demand

There are two types of goods in the economy: a homogeneous numeraire good and a differentiated good. The differentiated good is available in multiple *grades* indexed by $g \in \{1, \cdots, G\}$ with $1 < G < \infty$, and each grade of the differentiated good is produced by a continuum of firms of endogenous measure. There is a measure L of households each endowed with one unit of labor, which is supplied inelastically. Time is continuous, and each household has preferences given by:

$$\bar{U}_t = \int_0^\infty e^{-\beta s} U_{t+s} ds \tag{7}$$

$$U_t = Q_t^0 + \sum_{g=1}^{G} Q_t^g \tag{8}$$

$$Q_t^g = \alpha_0^g \int_{\Omega_t^g} \left(s^g q_{f,t}^g\right) df - \frac{1}{2}\alpha_1^g \int_{\Omega_t^g} \left(s^g q_{f,t}^g\right)^2 df - \frac{1}{2}\alpha_2^g \left[\int_{\Omega_t^g} \left(s^g q_{f,t}^g\right) df\right]^2 \tag{9}$$

Here, Q_t^0 denotes consumption of the numeraire good, Q_t^g denotes consumption of a grade g bundle of the differentiated good, $q_{f,t}^g$ for $f \in \Omega_t^g$ denotes consumption of firm f's variety of a grade g good, and Ω_t^g denotes the set of firms producing grade g varieties. Note that preferences are: (i) linear across time (7); (ii) linear across the numeraire and grades of the differentiated good (8); and (iii) of the Melitz-Ottaviano (2008) type within a grade of the differentiated good (9). Furthermore, note that despite the linearity of preferences in equation (8), there is love of variety within each grade in equation (9), and hence no two varieties of the differentiated good are perfect substitutes.

The weights $\{s^g\}_{g=1}^G$ in (9) reflect exogenous quality differences across grades, while the preference parameters $\{\alpha_0^g, \alpha_1^g, \alpha_2^g\}_{g=1}^G$ are also heterogeneous across grades and can be interpreted as follows. First, α_0^g is a demand shifter capturing the extent to which the household prefers grade g varieties of the differentiated good relative to the homogeneous good or other grades of the differentiated product. Second, α_1^g indexes the degree of product differentiation: as $\alpha_1^g \to 0$, all varieties within the grade become perfect substitutes. Third, α_2^g captures the extent of competition effects: as $\alpha_2^g \to 0$, utility is additively separable across varieties within the grade, and hence demand for one variety does not depend on demand for any other variety.

This preference structure implies linear demand functions given by:

$$\bar{q}_{f,t}^g \equiv Ls^g q_{f,t}^g = \frac{\alpha_0^g L}{\alpha_1^g + \alpha_2^g N_t^g} + \left(\frac{\alpha_2 N_t^g}{\alpha_1 + \alpha_2 N_t^g}\right)\left(\frac{L}{\alpha_1}\right)\bar{p}_t^g - \left(\frac{L}{\alpha_1}\right)\hat{p}_{f,t}^g \tag{10}$$

where $\bar{q}_{f,t}^g$ is aggregate quality-adjusted demand for firm f's output, $\hat{p}_{f,t}^g \equiv p_{f,t}^g / s^g$ is the quality-adjusted price charged by firm f, $\bar{p}_t^g \equiv \frac{1}{N_t^g}\int_{\Omega_t^g} \hat{p}_{f,t}^g df$ is the average quality-adjusted price charged by firms producing grade g varieties, and N_t^g is the measure of firms producing grade g varieties. Note that because of the quasi-linear demand structure (equation (8)), the demand for a firm's output depends only on the mass of competitors *within its own grade* as well as the prices charged by these firms. In other words, there is competition within grades but not across grades. This allows us to focus on the effects of competition within the firm's most immediate market, while abstracting from other sources of competition in order to reduce the computational complexity of the model.

4.1.2 Production and pricing

The production structure of the model is as follows. First, the numeraire good is produced one-for-one using labor under perfect competition, which implies a unit wage.[20] Next, all firms producing grade g varieties have access to the same production technology, which enables production of one unit of the good using c^g units of labor.

Each firm $f \in \Omega_t^g$ then takes N_t^g and $\bar{p}_t^g$ as given, and chooses $\hat{p}_{f,t}^g$ to maximize profits subject to the demand function (10). Since all firms within the grade have the same marginal cost, each firm chooses the same profit-maximizing price, and we henceforth omit the firm subscript. We also now omit time subscripts for brevity. The solution to the firm profit-maximization problem implies that markups (price over marginal cost), sales,

[20]We assume that the total labor supply L is large enough such that the numeraire is indeed produced and consumed in equilibrium.

and profits for each firm operating in grade g can be expressed as:

$$\mu^g = \frac{\phi^g + \rho^g}{\rho^g} \tag{11}$$

$$y^g = L\kappa^g \phi^g \left(\phi^g + \rho^g\right) \tag{12}$$

$$\pi^g = L\kappa^g \left(\phi^g\right)^2 \tag{13}$$

Here, the role of technology (cost per unit of quality demanded) is captured by the composite parameter:

$$\rho^g \equiv \frac{c^g}{\alpha_0^g s^g}, \tag{14}$$

the role of scale is captured by the composite parameter:

$$\kappa^g \equiv \frac{\left(\alpha_0^g\right)^2}{\alpha_1^g}, \tag{15}$$

the strength of competition is captured by the composite parameter:

$$\alpha^g \equiv \frac{\alpha_2^g}{\alpha_1^g}, \tag{16}$$

and ϕ^g is an endogenous variable that declines with the mass of firms operating within the grade:

$$\phi^g \equiv \frac{1 - \rho^g}{2 + \alpha^g N^g} \tag{17}$$

Note that for these variables, the identifiable parameter set collapses from five parameters $\{\alpha_0^g, \alpha_1^g, \alpha_2^g, s^g, c^g\}$ to three parameters $\{\rho^g, \kappa^g, \alpha^g\}$.

Importantly, observe that the measure of active firms N^g determines the level of competition within the grade, and as N^g increases, markups, sales, and profits all decline. In equilibrium, $\{N_g\}_{g=1}^G$ is determined endogenously by firms' innovation decisions, as described below in section 4.1.4. Furthermore, since the household's marginal utility of consuming any given variety is bounded, there is a maximum price above which the demand specified by (10) is equal to zero. We therefore assume that grade qualities and costs vary in such a way that output and profits are positive for any finite N^g.

Assumption 1. *For each grade $g \in \{1, \cdots, G\}$, the cost per unit of quality demanded satisfies $\rho^g < 1$.*

4.1.3 Innovation

In addition to making production decisions, firms also engage in innovative activities. We assume that each firm has access to the production technology for at most one grade at a time, but can move up the grade ladder by investing in R&D.[21] Specifically, a grade g firm that hires $RD^g(a^g)$ units of labor for R&D activities advances to grade $g+1$ according to a Poisson process with rate a^g. Upon a successful innovation, the firm adopts the grade $g+1$ technology and discards the grade g technology; there is hence creative destruction within the firm. The function RD^g captures grade-specific innovation costs, and is assumed to satisfy the following regularity properties.

Assumption 2. *For each grade $g \in \{1, \cdots, G-1\}$, the innovation cost function $RD^g : \mathbb{R}^+ \to \mathbb{R}^+$ satisfies (i) $RD^{g'} > 0$, (ii) $RD^{g''} > 0$, $RD^g(0) = 0$, (iii) $\lim_{a\to 0} RD^{g'}(a) = 0$, and (iv) $\lim_{a\to\infty} RD^{g'}(a) = \infty$.*

In what follows, we will study steady-states of the model in which the value of being a grade g producer, V^g, is constant over time. Given the innovation process, the values $\{V^g\}_{g=1}^{G-1}$ must then satisfy the following Bellman equation:

$$(\beta + \epsilon) V^g = \max_{a^g} \left\{ \pi^g - RD^g(a^g) + a^g \left(V^{g+1} - V^g \right) \right\} \tag{18}$$

where ϵ is the rate of exogenous firm exit. Evidently, if $V^{g+1} \leq V^g$, a grade g firm will not choose to innovate and therefore will optimally choose $a^g = 0$. On the other hand if $V^{g+1} > V^g$, the firm would like to move up the grade ladder. The optimal innovation decision for the firm can then be generally characterized by:

$$RD^{g'}(a^g) = \max\left\{ V^{g+1} - V^g, 0 \right\} \tag{19}$$

which admits a unique solution for a^g given assumption 2.

The first-order condition (19) captures the essence of the theory of innovation in this model. In particular, note that the incentive to innovate depends on the difference between pre- and post-innovation values. The effects of scale on innovation can thus heuristically be interpreted as a proportional change in both V^g and V^{g+1}: if the market for the differentiated product becomes larger overall, firms have greater incentive to innovate. On the other hand, the effects of competition may affect V^g and V^{g+1} differentially. If a

[21]The assumption that firms produce a single product at a time is without loss of generality if firms make innovation decisions at the product level and there are no innovation spillovers across products within a firm. For issues of tractability, we abstract from the more general case of multiproduct firms with innovation that spills over across products.

firm faces more competitors in its current grade (larger N^g), the value of failing to successfully innovate falls (smaller V^g), and hence innovation is incentivized. Conversely, if competition is tougher in the post-innovation market (larger N^{g+1}), the value of successfully innovating falls (smaller V^{g+1}), and hence innovation is disincentivized.

To complete the description of innovation in the model, it remains to specify what occurs at the boundaries of the grade ladder: how firms at the frontier grade G innovate, and how firms enter the market. First, since the number of grades is assumed to be finite, we model innovation by firms at the frontier by assuming that these firms innovate to reduce the hazard rate of exit. The value of operating as a grade G firm then satisfies:

$$\beta V^G = \max_{a^G} \left\{ \pi^G - RD^G \left(a^G \right) - \epsilon^G \left(a^G \right) V^G \right\} \tag{20}$$

where ϵ^G is a function specifying the exit rate at the frontier, and is assumed to satisfy the following regularity properties.[22]

Assumption 3. *The frontier exit rate function* $\epsilon^G : \mathbb{R}^+ \to \mathbb{R}^+$ *satisfies (i)* $\epsilon^{G'} < 0$, *(ii)* $\epsilon^{G''} > 0$, *and (iii)* $\lim_{a \to 0} \epsilon^{G'}(a) < 0$.

The first-order condition for innovation at the frontier is then:

$$RD^{G'} \left(a^G \right) = -\epsilon^{G'} \left(a^G \right) V^G \tag{21}$$

which admits a unique solution for a^G given assumption 3.

Finally, any firm wishing to enter the market for the differentiated good can hire f^E units of labor as an entry cost, following which it obtains the technology for producing a new variety of grade $g = 1$. Free entry then requires:

$$V^1 = f^E \tag{22}$$

4.1.4 Mass of firms in steady-state

To determine the measure of firms producing each grade of the differentiated good, we consider the implications of a steady-state equilibrium. First, note that at each point in time, a measure N^E of firms enter and become grade 1 producers, a measureϵN^1 of grade 1 producers exit the economy, and a measure $a^1 N^1$ of grade 1 producers successfully innovate and become grade 2 producers. Therefore in steady-state, the measure of grade 1 producers must satisfy:

$$N^E = \left(\epsilon + a^1 \right) N^1 \tag{23}$$

[22] In practice, we use a simple functional form $\epsilon^G(a) = 1/a$.

Next, consider the inflow and outflow of grade g producers for $g \in \{2, \cdots G-1\}$. At each point in time, a measure $a^{g-1}N^{g-1}$ of grade $g-1$ producers successfully innovate to become grade g producers. Simultaneously, a measure ϵN^g of grade g producers exit and a measure $a^g N^g$ successfully innovate to become grade $g+1$ producers. Therefore, in steady-state, the masses of firms must satisfy:

$$a^{g-1}N^{g-1} = (\epsilon + a^g) N^g \tag{24}$$

Finally, at the frontier grade G, a measure $a^{G-1}N^{G-1}$ of grade $G-1$ producers successfully innovate to become grade G producers, while a measure $\epsilon^G \left(a^G\right) N^G$ of grade G producers exit. Hence:

$$a^{G-1}N^{G-1} = \epsilon \left(a^G\right) N^G \tag{25}$$

4.1.5 Labor market

With the quasi-linear demand structure, labor market clearing simply requires that the numeraire good is indeed produced in equilibrium. Specifically, the total amount of labor employed by firms in the differentiated goods sector for production, innovation, and entry costs must be less than the total labor endowment:

$$\bar{Q}^0 \equiv LQ^0 = L - N^E f^E - \sum_{g=1}^{G} N^g \left[l^g + RD^g\left(a^g\right)\right] > 0 \tag{26}$$

where production labor hired by each grade g firm is:

$$l^g = L\kappa^g \rho^g \phi^g \tag{27}$$

4.1.6 Welfare

Household utility at each point in time is given by (8). Consumption of the numeraire is given by (26), while consumption of the grade g bundle can be expressed as:

$$Q^g = \kappa^g N^g \phi^g - \frac{1}{2}\kappa^g N^g \left(\phi^g\right)^2 - \frac{1}{2}\alpha^g \kappa^g \left(N^g \phi^g\right)^2 \tag{28}$$

Note again from equations (26)-(28) that of the set of original model parameters $\{\alpha_0^g, \alpha_1^g, \alpha_2^g, s^g, c^g\}$, only the composites $\{\rho^g, \kappa^g, \alpha^g\}$ matter for welfare.[23]

[23] While the size of the labor endowment L does not matter independently from the scale composite parameters $\{\kappa^g\}_{g=1}^{G}$ for outcomes within the differentiated sector, it does matter for welfare because L determines the size of the differentiated sector relative to the numeraire sector under quasi-linear preferences.

4.2 Equilibrium definition and solution

4.2.1 Equilibrium definition

A steady-state equilibrium of the model is a mass of entry N^E, a sequence of firm masses $\{N^g\}_{g=1}^G$, a sequence of profits $\{\pi^g\}_{g=1}^G$, a sequence of product innovation rates $\{a^g\}_{g=1}^G$, and a sequence of firm values $\{V^g\}_{g=1}^G$, all of which satisfy the profit equation (13), Bellman equations (18) and (20), R&D optimality conditions (19) and (21), free-entry condition (22), steady-state entry conditions (23)-(25), and the labor market condition (26).

4.2.2 Solution algorithm

To solve the model, we iterate backwards on the Bellman equation (18) according to the following algorithm, which takes several seconds on a standard personal computer:

1. Guess N^G and compute V^G and a^G from (20) and (21).
2. For $g \in \{1, \cdots, G-1\}$, given V^{g+1}, N^{g+1}, and a^{g+1}, numerically solve (18), (19), and (24) to obtain V^g, N^g, and a^g.
3. Compute N^E from (23).
4. Repeat steps 1-3, adjusting the guess of N^G until the free-entry condition (22) is satisfied.
5. Check that the labor market condition (26) holds.

5 Open Economy Model

We now extend the model developed above to incorporate multiple (potentially asymmetric) locations and trade costs. Here,we develop the general case with an arbitrary number of countries, and then focus on a North-South model in the calibration and counterfactual simulations of the model.

5.1 Model environment

There are $J \geq 2$ locations. These locations are potentially asymmetric in terms of labor endowments $\{L_i\}_{i=1}^J$, wages $\{w_i\}_{i=1}^J$, entry costs $\{f_i^E\}_{i=1}^N$, production technologies $\{s_i^g, c_i^g\}_{i=1}^N$, and R&D costs $\{RD_i^g\}_{i=1}^N$. Given the available data, we assume that households in every location have identical preferences given by (7)-(9). We also assume that there are iceberg trade costs $\{\tau_{ij}\}_{i,j=1}^{J}$ between locations, where $\tau_{ij} \geq 1$ denotes the cost of shipping goods from j to i. Finally, we assume that the numeraire good is freely traded and produced using $\frac{1}{w_i}$ units of labor in location i, so that w_i is also the wage in country i.[24] As in the closed economy model, we focus on steady-state equilibria in which the masses of firms producing each grade are time-invariant.

5.1.1 Demand

Consider a firm from location j with access to grade g technology. The demand function faced by this firm in the location i market takes the same form as (10), and is given by:

$$\bar{q}_{ij}^{g} \equiv L_i s_j^g q_{ij}^g = \frac{\alpha_0^g L_i}{\alpha_1^g + \alpha_2^g N_i^g} + \left(\frac{\alpha_2^g N_i^g}{\alpha_1^g + \alpha_2^g N_i^g}\right)\left(\frac{L_i}{\alpha_1^g}\right)\bar{p}_i^g - \left(\frac{L_i}{\alpha_1^g}\right)\hat{p}_{ij}^g \tag{29}$$

where $\hat{p}_{ij}^g$ is the quality-adjusted price charged by the firm. Here, $N_i^g = \sum_{j=1}^{J} N_{ij}^g$ denotes the measure of firms supplying location i with grade g varieties, where N_{ij}^g is the measure of these firms that supply i from j. The average quality-adjusted price charged by firms supplying grade g varieties in location i is given by $\bar{p}_i^g = \frac{1}{N_i^g}\sum_{j=1}^{J} N_{ij}^g \hat{p}_{ij}^g$.

5.1.2 Production

With positive trade costs, it is possible that some grades will not be traded across locations. In contrast with the closed economy model, we therefore now have to differentiate between the measure of firms that supply location i with grade g varieties, N_i^g, and the measure of firms that have access to the technology for producing grade g varieties in location i, M_i^g .

As shown in the appendix, grade g producers in j will find it profitable to sell in i if and only if their *supply cost* $\rho_{ij}^g \equiv \frac{w_j c_j^g \tau_{ij}}{\alpha_0^g s_j^g}$ is below a certain threshold value $\rho_{i,max}^g$:

$$\rho_{ij}^g \leq \rho_{i,max}^g \equiv \frac{2 + \bar{\rho}_i^g \alpha^g N_i^g}{2 + \alpha^g N_i^g} \tag{30}$$

Observe that the right-hand side of (30) does not vary with the source country j. This implies that in any equilibrium of the model, grade g varieties are supplied to location i from source countries with the lowest supply costs. Condition (30) can also be written as:

$$\rho_{ij}^g \leq 1 - \frac{1}{2}\alpha^g \tilde{\rho}_{ij}^g \tag{31}$$

[24] As before, we assume that labor endowments in each location are large enough such that the numeraire is produced and consumed in positive amounts in all locations.

where $\tilde{\rho}_{ij}^{g}$ summarizes the state of technology for all suppliers in the grade g market in location i relative to technology for firms from j:

$$\tilde{\rho}_{ij}^{g} \equiv \sum_{k=1}^{J} N_{ik}^{g} \left(\rho_{ij}^{g} - \rho_{ik}^{g} \right) \tag{32}$$

Note that given the supply cost ρ_{ij}^{g} for firms from j, the term $\tilde{\rho}_{ij}^{g}$ is greater when there are more firms selling in i with supply costs lower than ρ_{ij}^{g}, or when there are fewer firms selling in i with supply costs greater than ρ_{ij}^{g}. Furthermore, observe that $\tilde{\rho}_{ij}^{g}$ depends only on the measure of suppliers in i from all locations other than j, and in particular does not depend on N_{ij}^{g}. This implies that in any equilibrium of the model, either all firms from j with access to grade g technology sell in i or none of them do:

$$N_{ij}^{g} = \begin{cases} M_{j}^{g} & , \text{ if } \rho_{ij}^{g} \leq \rho_{i,max}^{g} \\ 0 & , \text{ o/w} \end{cases} \tag{33}$$

Assuming condition (31) holds, the solution for firm prices then implies the following expressions for equilibrium markups, sales, and profits:

$$\mu_{ij}^{g} = \frac{\phi_{ij}^{g} + \rho_{ij}^{g}}{\rho_{ij}^{g}} \tag{34}$$

$$y_{ij}^{g} = L_i \kappa^{g} \phi_{ij}^{g} \left(\phi_{ij}^{g} + \rho_{ij}^{g} \right) \tag{35}$$

$$\pi_{ij}^{g} = L_i \kappa^{g} \left(\phi_{ij}^{g} \right)^{2} \tag{36}$$

As in the closed economy, ϕ_{ij}^{g} is an endogenous variable that summarizes the effect of competition from other firms:

$$\phi_{ij}^{g} \equiv \frac{1 - \rho_{ij}^{g} - \frac{1}{2}\alpha^{g} \tilde{\rho}_{ij}^{g}}{2 + \alpha^{g} N_{i}^{g}} \tag{37}$$

Note that because firms from different locations have potentially different production technologies, what matters for the degree of competition is not only the measure of competitors (N_{i}^{g}), but also the supply costs of a firm's competitors relative to its own supply cost ($\tilde{\rho}_{ij}^{g}$). Observe that if supply costs are identical across all source locations, then $\tilde{\rho}_{ij}^{g} = 0$ and expressions (34), (35), and (36) reduce to their closed-economy counterparts (11), (12), and (13). In this special case, the market for grade g varieties in i resembles a closed economy.

Finally, to ensure that all locations can be profitably supplied with each grade of the differentiated good by firms from at least one location, we make the following assumption.

Assumption 4. *For each location $i \in \{1, \cdots, J\}$ and each grade $g \in \{1, \cdots, G\}$, the minimum marginal supply cost satisfies* $\min_{j \in \{1, \cdots, J\}} \rho_{ij}^{g} < 1$.

5.1.3 Innovation

The innovation process in the open economy is the same as in the closed economy. Firms in each location choose investments in R&D and move up the grade ladder stochastically subject to an exit shock that arrives with identical Poisson rate ϵ in all locations. The value of being a producer in location i of a variety of grade $g < G$ then satisfies:

$$(\beta + \epsilon) V_i^g = \max_{a_i^g} \left\{ \pi_i^g - w_i RD_i^g \left(a_i^g\right) + a_i^g \left(V_i^{g+1} - V_i^g\right) \right\} \tag{38}$$

As before, the optimal innovation decision is characterized in general by:

$$RD_i^{g'} \left(a_i^g\right) = \max \left\{ V_i^{g+1} - V_i^g, 0 \right\} \tag{39}$$

Similarly, at the frontier grade G where innovation reduces the hazard rate of exit, the value of a firm satisfies:

$$\beta V_i^G = \max_{a_i^G} \left\{ \pi_i^G - w_i RD_i^G \left(a_i^G\right) + \epsilon \left(a_i^G\right) V_i^G \right\} \tag{40}$$

while the optimal innovation decision is characterized by:

$$w_i RD_i^G \left(a_i^G\right) = -\epsilon' \left(a_i^G\right) V_i^G$$

The key difference here relative to the closed economy is that now total profits in (38) and (40) are the sum of domestic profits and profits across all potential export destinations:

$$\pi_i^g = \sum_{j=1}^{J} \pi_{ji}^g \tag{41}$$

5.1.4 Entry

Firms in location i can enter the economy by paying an entry cost of f_i^E units of labor to obtain technology for producing a grade 1 variety. Free entry in each location therefore imposes:

$$V_i^1 = f_i^E \tag{42}$$

5.1.5 Steady-state firm grade distribution

Following the same logic as for the closed-economy, the steady-state mass of producers for each grade in location i satisfies the following:

$$M_i^E = \left(\epsilon + a_i^1\right) M_i^1 \tag{43}$$

where M_i^E denotes the mass of entrants. For $g \in \{2, \cdots, G-1\}$, the equilibrium relation is:

$$a_i^{g-1} M_i^{g-1} = \left(\epsilon + a_i^g\right) M_i^g \tag{44}$$

while at the frontier grade:

$$a_i^{G-1} M_i^{G-1} = \epsilon \left(a_i^G\right) M_i^G \tag{45}$$

5.1.6 Labor market

In order for the numeraire good to be produced in all locations in equilibrium, the total amount of labor employed by firms in the differentiated goods sector for production, innovation, and entry costs must be less than the total labor endowment:

$$L_i - M_i^E f_i^E - \sum_{g=1}^{G} \left[\sum_{j=1}^{J} N_{ji}^g l_{ij}^g + M_i^g w_i RD_i^g \left(a_i^g\right) \right] > 0 \tag{46}$$

where total production labor hired by grade g producers in j that sell in i is given by:

$$l_{ij}^g = L_i \kappa^g \rho_{ij}^g \phi_{ij}^g \tag{47}$$

5.1.7 Welfare

Household utility in location i at each point in time is given by:

$$U_i = Q_i^0 + \sum_{g=1}^{G} s^g Q_i^g \tag{48}$$

Since the numeraire is assumed to be freely traded, production and consumption of the numeraire in each location need not be equal. Each household's consumption of the numeraire in location i is given by its income net of expenditures on differentiated products.

Household income is the sum of wages and firm profits (we assume that aggregate profits in each location are disbursed equally to all households in that location). Hence:

$$L_i Q_i^0 = w_i L_i + \Pi_i - \sum_{g=1}^{G} \sum_{j=1}^{J} N_{ij}^g y_{ij}^g \tag{49}$$

$$\Pi_i = \sum_{g=1}^{G} \left[\sum_{j=1}^{J} N_{ji}^g \pi_{ji}^g - M_i^g w_i RD_i^g \left(a_i^g \right) \right] - M_i^E f_i^E \tag{50}$$

Finally, consumption of the grade g bundle can be expressed as:

$$Q_i^g = \kappa^g \sum_{j=1}^{G} N_{ij}^g \phi_{ij}^g - \frac{1}{2}\kappa^g \sum_{j=1}^{J} N_{ij}^g \left(\phi_{ij}^g \right)^2 - \frac{1}{2}\alpha^g \kappa^g \left(\sum_{j=1}^{J} N_{ij}^g \phi_{ij}^g \right)^2 \tag{51}$$

As in the closed economy, note that of the set of original model parameters $\left\{ \alpha_0^g, \alpha_1^g, \alpha_2^g, s_i^g, c_{ij}^g \right\}$, only the composites ρ_{ij}^g, α^g, and κ^g matter for welfare.

5.2 Equilibrium definition and solution

5.2.1 Equilibrium definition

A steady-state equilibrium of the open-economy model is a list of sequences for entry $\left\{ M_i^E \right\}_{i=1}^{J}$, production $\left\{ \left\{ M_i^g \right\}_{g=1}^{G} \right\}_{i=1}^{J}$, supply $\left\{ \left\{ N_{ij}^g \right\}_{g=1}^{G} \right\}_{i,j=1}^{J}$, profits $\left\{ \left\{ \pi_{ij}^g \right\}_{g=1}^{G} \right\}_{i,j=1}^{J}$, innovation probabilities $\left\{ \left\{ a_i^g \right\}_{g=1}^{G} \right\}_{i=1}^{J}$, and firm values $\left\{ \left\{ V_i^g \right\}_{g=1}^{G} \right\}_{i=1}^{J}$, all of which satisfy the supply conditions (33), profit equations (36), Bellman equations (38) and (40), R&D optimality conditions (39), free-entry conditions (42), steady-state entry conditions (43)-(45), and the labor market condition (46).

5.2.2 Solution algorithm

To solve the model, we employ a computational algorithm similar to that used to solve the closed-economy model (described in section 4.2.2), which as before takes several seconds on a standard personal computer. Again, this involves guessing the measure of firms operating at the frontier grade in each market, iterating backwards on the Bellman equation (38), and then checking the free-entry condition (42) and labor market condition (46). Interested readers are referred to the online appendix for a detailed description of the solution algorithm for the open economy model.

6 Model Calibration

To calibrate the model, we focus on an economy with $J = 2$ countries and use data for Chinese firms as described above, as well as data for Canadian firms obtained from Statistics Canada. The latter data are used to represent OECD firms, so as to capture trade between China and developed countries as a whole. As such, we scale the number of Canadian firms by the ratio of OECD to Canadian gross domestic product, and weight the export data by the relevant China-OECD trade shares. The implicit assumption made here is that Canadian firms are representative of OECD firms in terms of firm-level characteristics such as sales, exports, profits, and R&D.

6.1 Model parameters

The only functional form assumptions that need to be imposed before proceeding with the model calibration are for the innovation costs $\left\{\{RD_i^g\}_{g=1}^G\right\}_{i=1}^J$. Here, we assume constant elasticity functions:

$$RD_i^g(a) = b_i^g a^{\eta_i^g} \tag{52}$$

where b_i^g measures the level of R&D costs and η_i^g captures the sensitivity of R&D costs to the innovation rate. This parameterization allows us to solve the R&D first-order conditions analytically and speeds up the computation significantly.

The model thus has the following sets of parameters (parameter count in parentheses) - demand parameters ($2G$):

$$\Theta_D \equiv \{\kappa^g, \alpha^g\}_{g=1}^G, \tag{53}$$

technology parameters (J^2G):

$$\Theta_P \equiv \left\{\{\rho_{ij}^g\}_{g=1}^G\right\}_{i,j=1}^J, \tag{54}$$

innovation parameters ($J(2G+1)$):

$$\Theta_I \equiv \left\{f_i^E, \{b_i^g, \eta_i^g\}_{g=1}^G\right\}_{i=1}^J, \tag{55}$$

country-level macro parameters ($2J$):

$$\Theta_M = \{w_i, L_i\}_{i=1}^J, \tag{56}$$

and a set of parameters $\{\beta, \epsilon, G\}$ that will not be calibrated to data.[25]

6.2 Assignment of grades

To calibrate the model's parameters, we first need to take a stand on what defines a "grade" in the data. In the model, grades are technically defined by segmentation of the differentiated goods sector along various dimensions: demand characteristics, production technologies, and innovation costs. Furthermore, the model's assumptions imply that all firms within a grade are identical in every one of these aspects. Hence, assigning firms in the data to grades is not a straightforward task for two reasons. First, it is not immediately obvious what the most relevant variables for clustering firms should be. For example, should we consider firms that have similar production costs or innovation costs as belonging to the same grade? Second, many of the primitive grade-level characteristics are not directly observable.

To deal with this issue, we explore several avenues. First, the notion of firms operating in segmented markets is in essence a demand-side construct: consumers perceive different grades as having different characteristics that matter for utility. From this perspective, the most natural way to define grades would be by the quality of products that firms produce. As discussed in section 2.3, however, we are able to obtain quality estimates only for Chinese firms that are matched to the customs data. Hence, while assigning grades based on quality would be economically appealing, this would preclude us from studying firms in the full CME sample.

A second option for defining grades would be to focus on the interaction between competition and innovation. From this perspective, the most salient characteristic of a grade is its competitive environment. It is precisely because grades differ along this dimension that the model allows us to study how trade affects innovation through not only the standard scale channel but through competitive effects as well. As such, a potential solution is to group firms based on the markup estimates described in section 2.2. Given the data available to us at the moment, this is a feasible option for Chinese firms, but not for the Canadian firms that we use to calibrate the open economy model, although we are currently working with Statistics Canada to estimate markups for Canadian firms as well, which would relax this data constraint.

Finally, the most straightforward option for defining grades is to group firms based on size. In the model calibration and simulation described below, we adopt this ap-

[25] We set the discount factor and exit rates to typical values ($e^{-\beta} = .95$, $e^{-\epsilon} = .9$). In the baseline analysis, we set the number of grades at $G = 10$, and explore how different values of G affect our quantitative results.

proach for the time being due to data availability. Specifically, the assignment of firms to grades in the data is determined by implementing a k-means clustering algorithm on firm sales, which finds the grouping of firms that minimizes the within-group variance of sales. To account for differences across industries and time, the clustering algorithm is implemented separately for each industry-year.

To investigate the robustness of this clustering approach, Figure 4 compares the characteristics of grades defined by both sales and markups for Chinese firms in our data. The results are reassuring: clustering on either firm sales or markups yields grades that are generally increasing in sales, exports, profits, R&D expenditures, markups, and quality, and decreasing in the number of operating firms and the level of competition from other Chinese firms (equation (6)). In particular, the average estimated markups within each grade are almost identical whether grades are defined by markups or sales. As such, we are confident that the results shown below will be similar to results we hope to eventually obtain based on markup clustering for both Chinese and Canadian firms.

6.3 Calibration algorithm

Having assigned firms in the data to grades, we then calibrate the demand, technology, and innovation parameters $\{\Theta_D, \Theta_P, \Theta_I,\}$ by targeting average sales, exports, profits, and R&D for firms within each grade, as well as the number of firms operating within each grade:

$$y_i^g = \tilde{y}_i^g \tag{57}$$

$$y_{xi}^g = \tilde{y}_{xi}^g \tag{58}$$

$$\pi_i^g = \tilde{\pi}_i^g \tag{59}$$

$$RD_i^g = \tilde{RD}_i^g \tag{60}$$

$$N_i^g = \tilde{N}_i^g \tag{61}$$

where $\tilde{x}$ denotes the value of a variable x in the data. The calibration algorithm that achieves this proceeds in three steps.

First, suppose that the model exactly matches the number of firms in each grade-market, $\left\{\{N_i^g\}_{g=1}^G\right\}_{i=1}^2$. Then for each grade $g \in \{1, \cdots, G\}$, the two demand parameters $\{\kappa^g, \alpha^g\}$ and four production technology parameters $\left\{\rho_{ij}^g\right\}_{i,j=1}^2$ can be chosen to exactly match the six moment conditions (57)-(59), where sales, exports, and profits are given

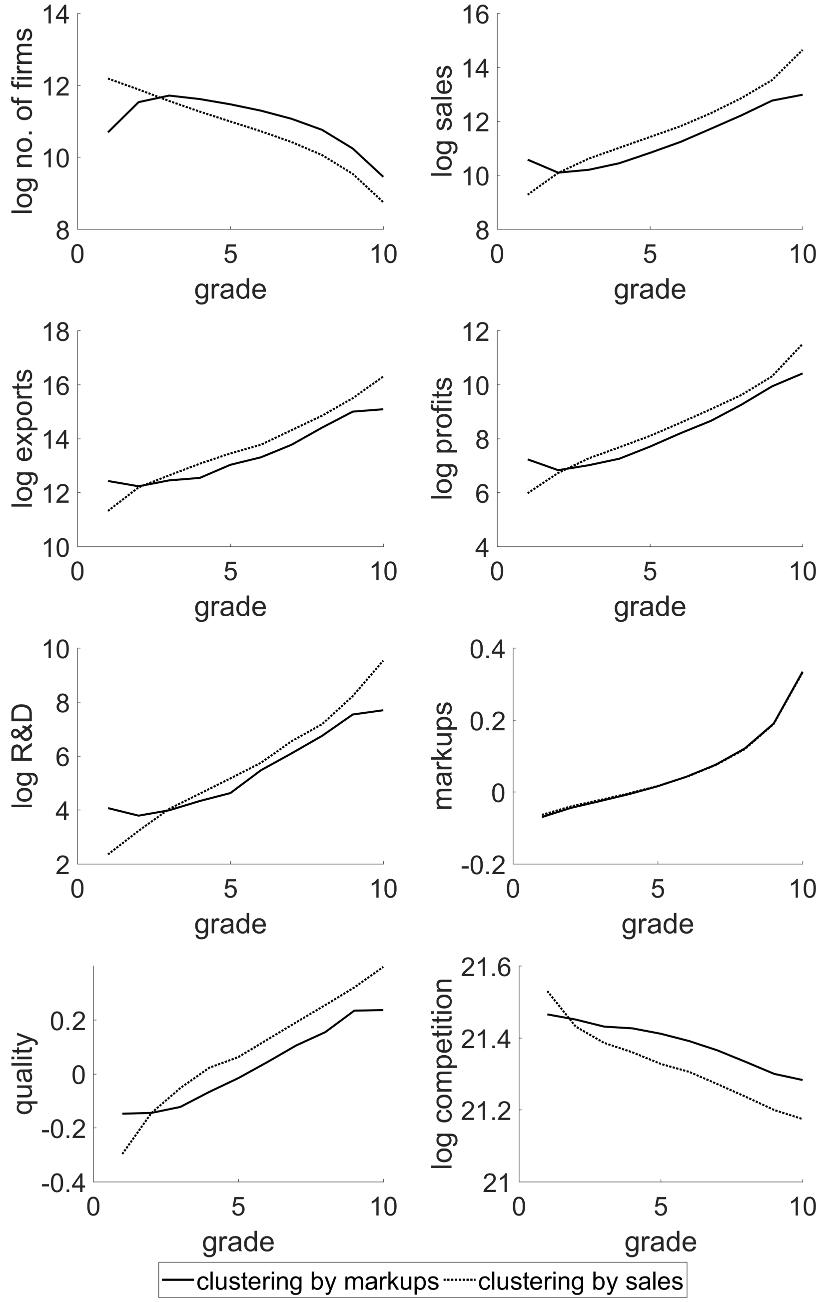


Figure 4: Clustering of Chinese firms by sales and markups

by equations (35)-(36). Intuitively, sales, exports, and profits identify demand, cost, and quality parameters.

Second, suppose that the model exactly matches the number of firms operating at the frontier grade in each market, $\left\{N_i^G\right\}_{i=1}^2$. Then for grade G, the four innovation parameters $\left\{b_i^G, \eta_i^G\right\}_{i=1}^2$ can be chosen to exactly match the two moment conditions (60) and the targeted exit rates in each location, $\epsilon^G\left(a_i^G\right) = \epsilon$. Similarly, for each grade $g \in \{1, \cdots, G-1\}$, the four innovation parameters $\left\{b_i^G, \eta_i^G\right\}_{i=1}^2$ can be chosen to exactly match the four moment conditions (60)-(61). Intuitively, R&D expenses and the relative distribution of firms across grades identify the innovation parameters.

Third, to ensure that the model does in fact match moment condition (61) for the frontier grade G, a simple two-dimensional search over the space of values for the entry costs $\left\{f_i^E\right\}_{i=1}^2$ is implemented. Intuitively, the overall level of entry identifies the entry costs. This last step of the calibration algorithm ensures that the first two steps yield parameter values that allow the model to match all targeted moments exactly.

Finally, the country-level macro parameters Θ_M are calibrated as follows. First, we interpret the differentiated goods sector as representing manufacturing, and choose the labor endowments $\{L_i\}_{i=1}^G$ to match manufacturing shares of 40% and 20% in China and the OECD respectively. Second, the wage in China is chosen as the numeraire, and the OECD wage is chosen to match relative hourly compensation costs based on data from the Bureau of Labor Statistics' International Labor Comparisons dataset.

6.4 Calibration results

6.4.1 Model fit

The fit of the model to the targeted moments is shown in Figure 5. The model is exactly identified by construction, and hence the model matches the targeted moments exactly. Note that the k-means clustering algorithm on sales within industry-year generates groupings of firms such that higher grades generally have larger average firm sales, exports, profits, and R&D, as well as fewer operating firms.

6.4.2 Equilibrium innovation rates

The equilibrium innovation rates implied by the model for Chinese and OECD firms are shown in Figure 6. Note that for both locations, the innovation rates exhibit an inverted-U pattern: firms that are either very far from the frontier grade or very close to it

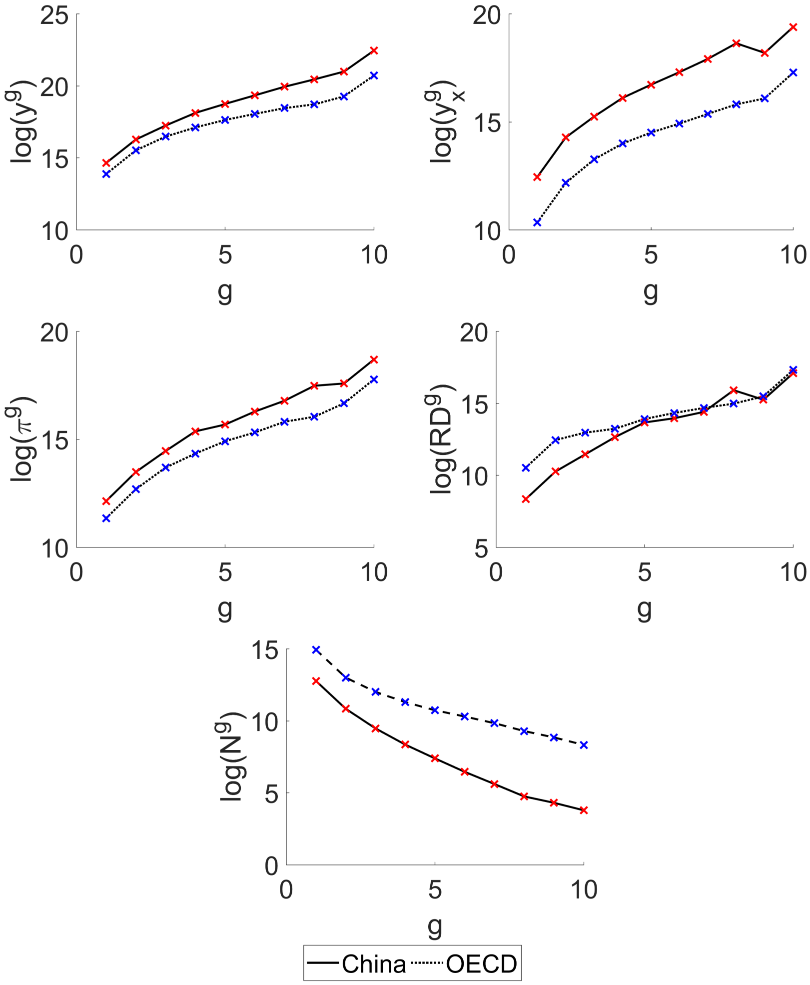


Figure 5: Model fit

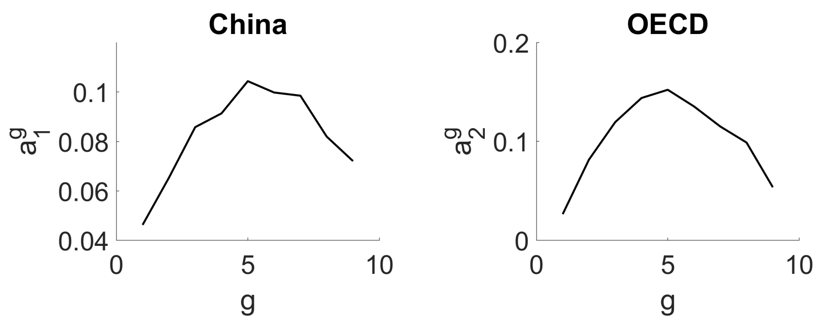


Figure 6: Equilibrium innovation rates

choose lower innovation rates than firms in the middle of the grade ladder. We interpret this result as being consistent with the "escape-the-competition" motive for innovation, as firms in the middle of the grade ladder face the greatest incentive to invest in R&D, so as to escape from tougher competition at lower grades and to move towards the frontier grade where competition is less intense and profits are higher.

7 Counterfactual Exercises

To study the effects of trade on innovation, we employ the model developed above to simulate two counterfactual exercises: a 5% reduction in the cost of exporting (for all grades) from China to the OECD, and a 5% reduction in the cost of exporting (for all grades) from the OECD to China. In each case, we examine the responses of six variables: (i) firm sales, (ii) exports, (iii) profits, (iv) R&D, (v) equilibrium innovation rates, and (vi) the measure of firms operating in each grade.

7.1 Lower export costs for Chinese firms

First, consider the effects of lower export costs from China to the OECD. The results of this counterfactual simulation are summarized in Figure 7. Here, we see that sales, exports, profits, R&D, equilibrium innovation rates, and the measure of firms in each grade increase for Chinese firms in all grades. This is intuitive: lower export costs lead to greater firm size and profitability, which raises the incentives for R&D, leading to higher innovation rates. Since firm values increase, more Chinese firms enter the market. For OECD firms, on the other hand, firm size and exports fall, and the measure of operating firms decreases in all grades. Again, this is intuitive: more competition from Chinese firms reduces OECD firm size and leads some firms to exit the market.

One might conclude from these results that lower Chinese export costs must therefore reduce R&D and innovation by OECD firms. Note from Figure 7, however, that the converse is true: R&D and equilibrium innovation rates *increase* for OECD firms in almost all grades. The reason for this is that the reduction in Chinese export costs affects OECD firms operating in different grades differentially. In particular, it reduces profits for firms operating in the lowest grades, but *increases* profits for firms operating in the highest grades. This latter effect is possible only because the reduction in Chinese export costs induces *exit* by some OECD firms, which reduces domestic OECD competition. Hence, the reduction in Chinese export costs raises the value of operating in high grades versus low grades for OECD firms, which from the R&D first-order condition (39) leads to an increase in innovation.

The result that lower Chinese export costs can induce an increase in innovation by OECD firms stems from the differential effects of the scale and competition channels embedded in the model. In particular, these simulations suggest the importance of accounting for both mechanisms when evaluating the effect of trade on firm-level innovation.

7.2 Lower export costs for OECD firms

Next, consider the effects of lower export costs from the OECD to China. The results of this counterfactual simulation are summarized in Figure 8, and are in essence the immediate opposite of those observed in the first simulation discussed above. In particular, sales, exports, profits, and the measure of firms in each grade increase for OECD firms and fall for Chinese firms in all grades. These responses are again intuitive for the reasons discussed above.

However, in response to lower OECD export costs, R&D and equilibrium innovation rates fall for *both* Chinese and OECD firms. For Chinese firms, the increase in competition from OECD firms reduces the incentive to innovate, hence leading to lower R&D in equilibrium. For OECD firms, the fall in export costs induces more OECD firms to enter the market. Hence, although profits increase because of the reduction in trade costs, the increase in entry causes profits for firms operating in the highest grades to fall relative to profits for firms in the lowest grades, which also reduces the incentives for innovation by OECD firms.

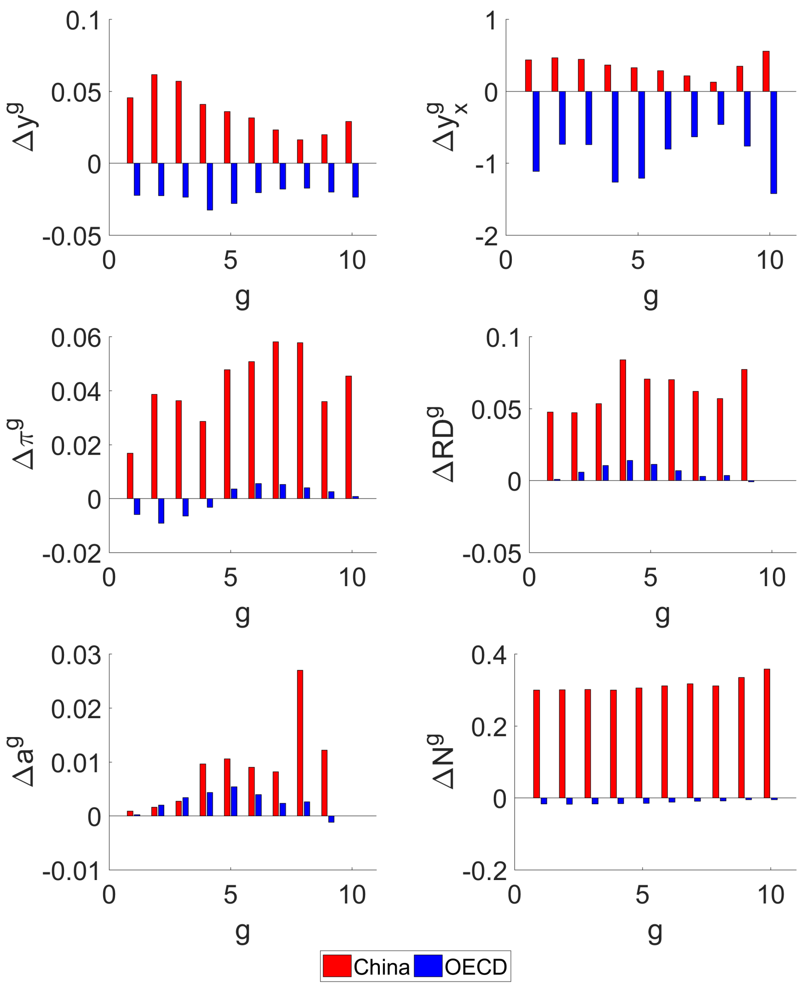


Figure 7: Counterfactual: 5% reduction in export costs from China to OECD

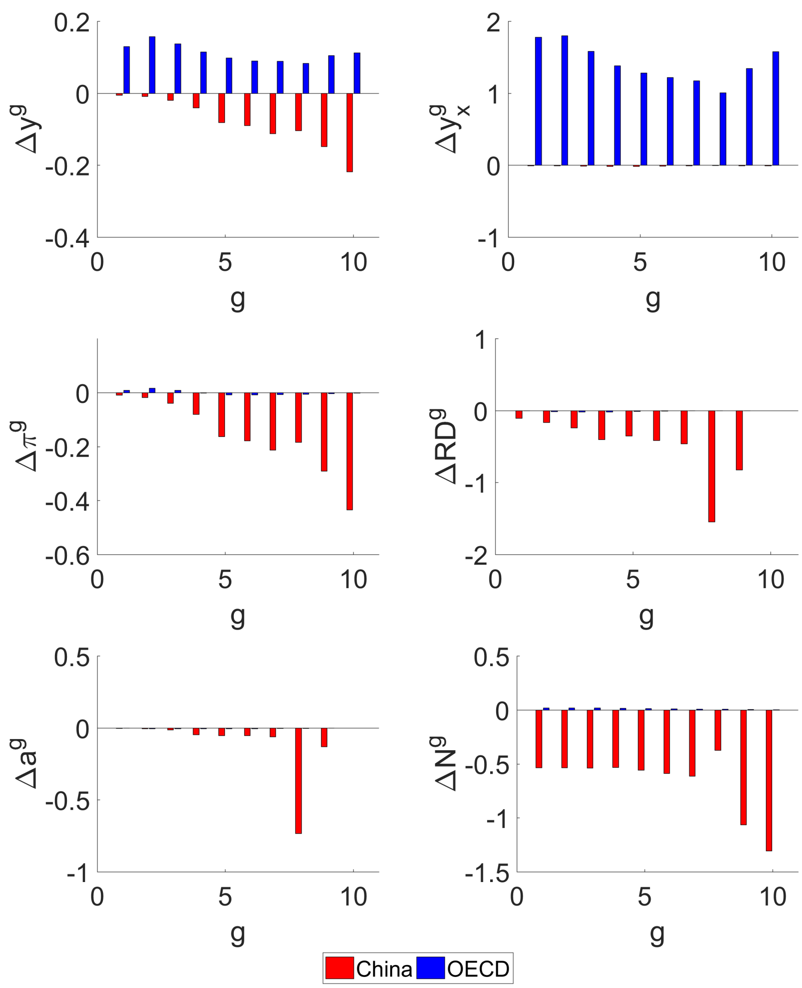


Figure 8: Counterfactual: 5% reduction in export costs from OECD to China

8 Conclusion

In this paper, we have established empirically that increased market size promotes innovation as measured by innovation inputs (R&D and patents). Further, increased innovation leads to increased TFP, higher markups, a larger number of exported products (increased variety), and an increase in the share of revenues generated by new products. We also found that increased competition within China has important effects on firm-level innovation, reducing innovation in the aggregate. The calibrated structural model formalizes these economic mechanisms, and simulations of the model indicate that accounting for the heterogeneous effects of trade on innovation by different firms through both the scale and competition channels is essential for understanding firm-level innovation in China. In particular, when firms can innovate to escape the competition, greater competition induced by lower trade barriers can lead firms to increase innovation rather than reduce it.

References

Acemoglu, Daron and Joshua Linn, "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry," *Quarterly Journal of Economics*, 2004, *119* (3), 1049–1090.

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer, "Identification properties of recent production function estimators," *Econometrica*, November 2015, *83* (6), 2411–2451.

Ackerberg, Daniel, Benkard Lanier, Steven Berry, and Ariel Pakes, "Econometric Tools for Analyzing Market Outcomes," in J. J. Heckman and E.E. Leamer, eds., *Handbook of Econometrics*, Vol. 6, Elsevier, 2007.

Aghion, Philippe, Antonin Bergeaud, Matthieu Lequien, and Marc Melitz, "The Impact of Exports on Innovation: Theory and Evidence," Working paper. 2017.

_ , Christopher Harris, Peter Howitt, and John Vickers, "Competition, Imitation and Growth with Step-by-Step Innovation," *Review of Economic Studies*, 2001, *68* (3), 467–492.

_ , Nicholas Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, "Competition and Innovation: An Inverted-U Relationship," *Quarterly Journal of Economics*, 2005, *120* (2), 701–728.

Akcigit, Ufuk, Sina Ates, and Giammario Impullitti, "Innovation and Trade Policy in a Globalized World," Working paper. 2017.

Atkeson, Andrew and Ariel Burstein, "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, 2010, *118* (3), 433–484.

_ and _ , "Aggregate Implications of Innovation Policy," Working paper. 2018.

Autor, David, David Dorn, Gordon H. Hanson, Gary Pisano, and Pian Shu, "Foreign Competition and Domestic Innovation: Evidence from U.S. Patents," Working paper. 2017.

Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu, "R&D Investment, Exporting, and Productivity Dynamics," *American Economic Review*, 2011, *101* (4), 1312–1344.

Beerli, Andreas, Franziska J. Weiss, Fabrizio Zilibotti, and Josef Zweimller, "Demand Forces of Technical Change: Evidence from the Chinese Manufacturing Industry," Working paper. 2018.

Berry, Steven T., "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, Summer 1994, *25* (2), 242–262.

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

Bombardini, Matilde, Bingjing Li, and Ruoying Wang, "Import Competition and Innovation: Evidence from China," Working paper. 2018.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang, "Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing," *Journal of Development Economics*, March 2012, *97* (2), 339–351.

_ , _ , and _ , "Challenges of working with the Chinese NBS firm-level data," *China Economic Review*, 2014, *30* (C), 339–352.

_ , Johannes Van Biesebroeck, Luhang Wang, and Yifan Zhang, "WTO Accession and Performance of Chinese Manufacturing Firms," *American Economic Review*, September 2017, *107* (9), 2784–2820.

Bustos, Paula, "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms," *American Economic Review*, 2011, *101* (1), 304–340.

Coelli, Federica, Andreas Moxnes, and Karen Helene Ulltveit-Moe, "Better, Faster, Stronger: Global Innovation and Trade Liberalization," Working paper. 2018.

De Loecker, Jan, "Detecting Learning by Exporting," *American Economic Journal: Microeconomics*, August 2013, *5* (3), 1–21.

_ and Frederic Warzynski, "Markups and Firm-Level Export Status," *American Economic Review*, October 2012, *102* (6), 2437–71.

Feenstra, Robert C., Zhiyuan Li, and Miaojie Yu, "Exports and Credit Constraints under Incomplete Information: Theory and Evidence from China," *Review of Economics and Statistics*, October 2014, *96* (4), 729–744.

Fieler, Ana Cecilia and Ann Harrison, "Escaping Import Competition and Downstream Tariffs in China," Working paper. 2018.

Impullitti, Giammario and Omar Licandro, "Trade, Firm Selection, and Innovation: the Competition Channel," *Economic Journal*, 2018, *128* (608), 189–229.

Khandelwal, Amit, "The Long and Short (of) Quality Ladders," *Review of Economic Studies*, October 2010, *77* (4), 1450–1476.

Lileeva, Alla and Daniel Trefler, "Improved Access to Foreign Markets Raises Plant-Level Productivity ... for Some Plants," *Quarterly Journal of Economics*, August 2010, *CXXV* (3), 1051–1100.

Liu, Dan, Liugang Sheng, and Miaojie Yu, "Highways and Firms" Exports: Evidence from China," Working Paper, CCER 2018.

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

_ **and Giancarlo I. P. Ottaviano**, "Market Size, Trade, and Productivity," *Review of Economic Studies*, 2008, *75* (1).

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, November 1996, *64* (6), 1263–1297.

Orr, Scott, Daniel Trefler, and Miaojie Yu, "Estimating Productivity Using Chinese Data: Methods, Challenges and Results," in Lili Yan Ing and Miaojie Yu, eds., *World Trade Evolution: Growth, Productivity and Employment*, Routledge, forthcoming.

Roberts, Mark J. and James R. Tybout, "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs," *American Economic Review*, September 1997, *87* (4), 545–564.

Schott, Peter, "The Relative Sophistication of Chinese Exports," *Economic Policy*, 2008, *53*, 5–49.

Sutton, John and Daniel Trefler, "Deductions from the Export Basket: Capabilities, Wealth and Trade," *Journal of Political Economy*, June 2016, *124* (3), 826–878.

Yu, Miaojie, "Processing Trade, Tariff Reductions and Firm Productivity: Evidence from Chinese Firms," *Economic Journal*, June 2015, *125* (585), 943–988.

A Pricing in the Open Economy

Assuming that it is profitable for grade g producers in j to sell in i, the demand function (29) implies the following optimal quality-adjusted price:

$$\hat{p}_{ij}^{g} = \frac{1}{2}\left(\hat{p}_{i,max}^{g} + \alpha_0^{g}\rho_{ij}^{g}\right) \tag{62}$$

where $\rho_{ij}^{g} \equiv \frac{w_j c_j^{g} \tau_{ij}}{\alpha_0^{g} s_j^{g}}$ is the marginal cost of supplying a unit of quality demanded in location i from location j for a grade g variety (henceforth, the "supply cost"), and $\hat{p}_{i,max}^{g}$ is the maximum quality-adjusted price for grade g varieties in i at which demand is non-negative :

$$\hat{p}_{i,max}^{g} = \frac{\alpha_0^{g}\alpha_1^{g} + \alpha_2^{g} N_i^{g} \bar{p}_i^{g}}{\alpha_1^{g} + \alpha_2^{g} N_i^{g}} \tag{63}$$

The average quality-adjusted price in i can then be written as:

$$\bar{p}_i^{g} = \frac{1}{2}\left(p_{i,max}^{g} + \alpha_0^{g}\bar{\rho}_i^{g}\right) \tag{64}$$

where $\bar{\rho}_i^{g}$ is the average supply cost across producers from all locations supplying grade g varieties in i:

$$\bar{\rho}_i^{g} \equiv \frac{1}{N_i^{g}} \sum_{j=1}^{J} N_{ij}^{g} \rho_{ij}^{g} \tag{65}$$

Solving equations (62)-(64), we obtain the following expressions for the optimal, maximum, and average prices as functions of the masses of suppliers in i, $\left\{N_{ij}^{g}\right\}_{j=1}^{J}$:

$$\hat{p}_{ij}^{g} = \alpha_0^{g}\left(\phi_{ij}^{g} + \rho_{ij}^{g}\right) \tag{66}$$

$$\hat{p}_{i,max}^{g} = \alpha_0^{g}\left(\frac{2 + \alpha^{g}\bar{\rho}_i^{g} N_i^{g}}{2 + \alpha^{g} N_i^{g}}\right) \tag{67}$$

$$\bar{p}_i^{g} = \alpha_0^{g}\left(\frac{1 + \bar{\rho}_i^{g} + \bar{\rho}_i^{g}\alpha^{g} N_i^{g}}{2 + \alpha^{g} N_i^{g}}\right) \tag{68}$$

Note that grade g producers in j will find it profitable to sell in i if and only if $\hat{p}_{ij}^{g} \leq \hat{p}_{i,max}^{g}$. From (66) and (67), this is equivalent to the following condition:

$$\rho_{ij}^{g} \leq \frac{2 + \bar{\rho}_i^{g}\alpha^{g} N_i^{g}}{2 + \alpha^{g} N_i^{g}} \equiv \rho_{i,max}^{g} \tag{69}$$