Markup Premia of Exporters: Because of Exporting, or In Spite of It?*

Alexander McQuoid Loris Rubini

February 1, 2016

Abstract

We study the effect of exporting on markups. Our analysis rests on three stylized facts: (1) exporters charge higher markups than non-exporters; (2) entering into the export market is associated with larger markups; and (3) domestic and foreign firm sales are negatively correlated. The first two facts have sparked arguments that suggest exporting increases markups, but the causal relationship has not been studied directly. To do so, we build a model consistent with all three stylized facts. Our model is based on Melitz and Ottaviano (2008), with decreasing returns to scale technologies to allow for negative correlations between domestic and foreign sales. We calibrate the model to match key moments for Chilean firms, and then simulate counter-factual reductions in trade costs. Our results suggest that markup responses are quite heterogeneous overall. Along the intensive margin, lower trade costs increase markups on average and for most firms. Along the extensive margin, firm markups unambiguously decline. Three mechanisms are at work: first, the reduction in trade costs is not fully passed through to prices, increasing markups. Second, reduced trade costs drive firms to expand output, which increases marginal costs under decreasing returns to scale, reducing markups. Third, foreign demand is more elastic than domestic demand, and therefore greater trade exposure implies lower markups. While the first effect is more prevalent along the intensive margin, the second and third prevail in the extensive margin. So while exporters have higher markups than non-exporters, and firms increase their markups when starting to export under constant trade costs, when lower trade costs induce firms to start exporting, markups decrease.

JEL classification: F12; F17

Keywords: Exporting and decreasing returns to scale. Correlation between exports and domestic sales. Exporting and markups.

*We thank Bernardo Blum, Amit Khandelwal, Giammario Impullitti, Sam Kortum and Peter Schott for helpful comments. We benefitted from talks at the Kansas Fed, ILADES, NASME 2014, SED 2014, WEAI 2014 and RIDGE 2014. Loris Rubini gratefully acknowledges financial support from the Spanish Ministry of Science and Innovation (ECO2011-27014). Alexander McQuoid: Florida International University and The United States Naval Academy, mcquoid@usna.edu. Loris Rubini: Universidad Católica de Chile and Universidad Carlos III de Madrid, lrubini@uc.cl.

1 Introduction

The relationship between exporting and markups is poorly understood. Theoretical investigations have often sacrificed empirical regularities for the sake of tractability, while empirical investigations have been hampered by limitations in measurement. Recent innovations in measuring markups in standard firm level datasets, however, call for a reassessment of the mechanisms behind firm markup adjustment, particularly as it relates to the impact of declining trade costs on exporting markups. As De Loecker and Goldberg (2014) have argued, markups contain valuable information about the performance of the firm and can be used to identify the competitive effects of trade liberalizations.

Our approach starts with two stylized facts, first documented by De Loecker and Warzynski (2012): (1) exporters charge higher markups than non-exporters, and (2) entering into the export market is associated with larger markups. Taken together, these two observations have been used to suggest that exporting increases markups, but the underlying causal relationship has not been properly addressed. To study this causality, we add a third relevant stylized fact for understanding markup setting behavior: domestic and foreign sales are negatively correlated at the firm level. This suggests marginal costs are increasing, and therefore domestic and foreign markups are intrinsically connected. Blum et al. (2013) observe this negative correlation and note that the presence of a fixed factor of production and decreasing returns to scale in the mobile factor can account for it.

From this foundation, we build a new model of international trade with endogenous, heterogeneous markups, which is consistent with all three stylized facts. We then calibrate the model, and study markups adjustments in response to a counterfactual reduction in trade costs.[1]

Our results show that the markup adjustment process is quite heterogeneous in general. Still, we do isolate key firm characteristics and mechanisms that affect the direction in which markups change. Broadly speaking, when lower trade costs increase the exports of firms that were already exporting before the change, markups increase. But when the change drives firms to start exporting, markups drop.

All told, this implies that exporters charge higher markups not because they export, but rather in spite of it. Productivity and demand characteristics drive firms to jointly export, *and* to charge higher markups, but it does not follow that exporting increases markups. This is consistent with recent findings by Jamandreu and Yin (2014) that Chinese exporters charge a higher markup at home than abroad. Thus, we challenge the naive conclusion that since exporters charge higher markups, lowering trade costs should result in increased markups as more firms become exporters.

[1]We purposely avoid targeting our motivating facts in the calibration, which allows us to check the validity of the model by testing its implications against the stylized facts.

Our motivating empirical regularities were first uncovered by De Loecker and Warzynski (2012). Using data from Slovenian firms, they find markups are higher for exporters than for non-exporters and markups tend to increase when firms enter the export market. We confirm these findings using firm level data from Chile. To this, we add a third empirical regularity that firm sales domestically and abroad are negatively correlated. This observation is important because it suggests the prevalence of decreasing returns to scale in production, and consequently foreign and domestic markups are jointly, rather than independently, determined.

Our model is based on Melitz and Ottaviano (2008), who develop a setting with linear demand functions, monopolistic competition, and iceberg trade costs. To this we add a fixed factor of production, and add shocks to productivity, domestic demand, and foreign demand. Pairing demand shocks with a fixed factor in production can account for the obsevred negative correlation between domestic and foreign sales, as suggested by Blum et al. (2013). The productivity shock on the other hand generates a positive correlation between domestic and foreign sales. Thus, our model does not by construction deliver a negative correlation, and the calibration ultimately determines its sign.

To understand the implications from the interaction between shocks and technology, consider the following thought experiment. The demand shocks, in conjunction with decreasing returns in the mobile factor, imply that the decision to export is directly related to the decision to sell domestically. A positive shock to foreign demand leads to an increase in foreign sales, but the expansion of output leads to higher costs in both markets, and therefore higher prices domesitcally, where demand was unchanged. Shocks originating in the foreign market affect decisions in the domestic market and vice versa. Under constant marginal costs, shocks only affect local markets, and therefore can be analyzed in isolation. The productivity shock on the other hand reduces the marginal cost of selling both domestically and abroad, thus lowering both prices, increasing sales in both markets, and thus generating a positive correlation.

The presence of fixed inputs of production is quite prominent in the data. Asker et al. (2014) find that inputs adjust less frequently in more volatile industries. Repeated changes in export status, therefore, could imply that many inputs adjust slowly. To simplify the analysis, we model this rigidity by assuming that firms cannot change the quantity of the fixed input. Thus, the model works as if technologies had decreasing returns to scale, even when the returns could be constant or increasing.

Our main results are that the reaction of markups is heterogeneous, but follow systematic patterns. The average markup increases in response to falling trade costs, but focusing on the aggregate average hides important distinctions across firms. Almost half of all firms decrease markups, and there is significant skewness in the distribution of markup adjustments. To better understand this heterogeneity, we focus on the decision to export before and after the decline in trade costs.

Our counterfactual reduces an iceberg trade cost from 50% to 10%. Along the intensive margin, most markups increase. That is, firms that export under both trade regimes increase their markups by 11% on average, although 30% of these firms reduce their markups. Conversely, along the extensive margin, all markups fall (on average by 5%). Thus, when a decline in trade costs *drives* firms to start exporting, these firms reduce their markups. This suggests that an exporter's markup is larger than a non-exporter's in spite of exporting, not because of it.

The intuition behind these changes is as follows. Since trade costs are iceberg costs, reducing them amounts to reducing marginal costs. Some of this reduction, but not all, is passed on to prices, resulting in an increase in markups.[2] At the same time, output expands, leading to an increase in marginal costs given decreasing returns to scale in production. Again, some but not all of this increase is passed on to prices, driving firms to reduce markups. Finally, the foreign elasticity of demand is different than the domestic elasticity. When the foreign elasticity is lower than the domestic one, as we estimate in the data, a greater trade exposure implies lower markups. The ultimate effect for a firm depends upon the relative strengths of each of these three forces.

Along the intensive margin, the first effect tends to dominate, so that for most firms that were already exporting before the change in trade costs, markups increase. Along the extensive margin, a reduction in trade costs does not imply a reduction in marginal costs. As a result, there is no decline in prices, and hence no increase in markups. On the other hand, the scale effect is still present. A large expansion in output due to exporting implies a large increase in marginal costs, and consequently a reduction in markups since, once again, not all of the increase in costs is passed on to prices. Finally, there is an additional increase in revenues that comes from exports. But the high foreign elasticity implies this effect is small. Hence, for firms that begin exporting as trade costs decline, the reduction in trade costs leads to a decrease in markups.

Our calibration works as follows. First, we leverage the theoretical framework to identify and extract the realization of each of the three shocks from the data. The model provides a nonlinear mapping from domestic sales, foreign sales, and markups to the unobservable shocks to productivity and demand (domestic and foreign). We directly observe information on domestic sales and exports in the data, and we estimate firm markups using available input information following De Loecker and Warzynski (2012). After recovering the shock realizations, we estimate the distribution of domestic demand and productivity shocks via maximum likelihood.

Calibrating the foreign demand shock process is more complicated since we observe a biased sample of foreign demand realizations. Only firms that have sufficiently high foreign demand shocks relative to their domestic and productivity shocks are observed exporting. To work around this, we calibrate the parameters by matching total export volume and the share of firms that export

[2] Only in the extreme case of constant elasticity of demand and monopolistic competition, as with Dixit-Stiglitz preferences, would marginal cost savings be completely passed on to prices.

in the data.

Lastly, to match the remaining parameters (which relate to time-series components of the model), we use a simulated method of moments approach, and target firm sales autocorrelations for different types of firms.

While we purposely do not target our motivating stylized facts in the calibration, the calibrated model nonetheless delivers them. In the data for Chilean manufacturers, exporters charge a markup that is 26 percent larger than non-exporters and entering the export market is associated with a 2.5 percent increase in the markup. In the calibrated model, the markup premium for exporters is 37 percent, and entering the export market under constant trade costs is associated with an average increase in markups of 1 percent.

The correlation between domestic and foreign sales in the data is -0.19. In the model, this correlation is -0.15. Notice that the model is not forced to deliver a negative correlation. The introduction of fixed factor and decreasing returns to a mobile factor generates effective decreasing returns to scale, and thus allows for the possibility of a negative correlation, which would be absent in a constant marginal cost environment. However, productivity shocks generate a positive correlation. The finding of an aggregate negative correlation confirms the importance of introducing effective decreasing returns to scale in production.

We investigate this correlation further by analyzing the correlation between exports and domestic sales for different types of firms. We group firms according to the frequency with which they export. We find in both the data and the model that the correlation between domestic and foreign sales increases with export frequency. For example, the correlation for firms that export in every period in the data is +0.19, and for those firms that only export in around half of the periods, -0.37. The respective numbers in the model are +0.04 and -0.18.

For all of our groupings, we produce correlations of the same sign as the data, which switches with exporting frequency. The reason why the correlations increase with export frequency has to do with firm size: large firms tend to be frequent exporters, and tend to have flatter marginal cost curves, since our calibrated marginal cost function is strictly concave. Thus, large firms' marginal costs are closer to constant, and this amplifies the importance of productivity shocks, which account for the positive correlation.

Another observation from comparing the model and data for the different firm groupings is that the model delivers smaller absolute values than the data. One explanation for this result concerns the number of partner countries, which we do not observe directly in our data. Our model features only two countries, but an observed empirical regularity is that more frequent exporters sell to more destinations. If each foreign demand is associated with a different, independent demand shock, total foreign demand is much less volatile for frequent exporters. Therefore, for these firms, the effect of the productivity shock becomes more important, which accounts for the positive correlation. The opposite happens to infrequent exporters, with a smaller number of partners, and more volatile foreign demand shocks.[3]

To assess the importance of our assumption of decreasing returns to scale, we simulate the decline in trade costs under the more typical assumption of constant returns to scale. We find that nearly all firms on the intensive margin increase markups, with only a tiny fraction decreasing markups (about 2 percent). For extensive margin firms, still no firms increase markups, but now the decline in markups is essentially zero (less than 0.5 percent on average). Constant returns to scale therefore overstate the predicted effect of increased markups, missing both intensive and extensive margin adjustments.

As an additional test of the robustness of our model, we use an alternative approach developed by Jamandreu and Yin (2014) to estimate markups for exporters in both domestic and foreign markets. We find that domestic markups tend to be larger than foreign markups for Chilean exporters, consistent with Jamandreu and Yin (2014)'s findings for Chinese exporters.[4] These findings support the thesis that the exporter markup premia is in fact in spite of exporting rather than because of exporting.

An attractive feature of the equilibrium in our model is that some large firms are unproductive. This is absent in traditional trade models, since productivity usually determines firm size. Furthermore, the share of output exported varies greatly within firms, consistent with the data, but unlike traditional models which feature homogeneous, isoelastic demands across firms.

The rest of the paper is organized as follows. Section 2 reviews related literature. Section 3 describes the empirical evidence that motivates the model. Section 4 describes the model while Section 5 discusses estimation techniques. The ability of the model to match stylized facts is discussed in Section 6. Section 7 presents the main findings. Section 8 analyzes the sensitivity of the results, and Section 9 concludes.

2 Related Literature

We estimate the effect of trade costs on markups by performing a theoretical, counterfactual exercise. The ideal approach would be to estimate the effect of a change in trade costs in the data, but lack of available data is a substantial hurdle. To the best of our knowledge, the only paper able to identify the effect of changes in trade costs is De Loecker et al. (2012), who use a unique dataset containing information on quantities and prices. Their focus is on importers during a

[3]While the minimum number of partners is one as in the model, the model is calibrated to averages, thus producing average correlations.

[4]The technique is only valid for exporters, and as such, cannot be used in our main investigation of markup differences between exporters and non-exporters.

multidimensional trade liberalization episode in India. They find, as we do in the case of exporters, that the response of markups to declining trade costs is heterogeneous.

Our results suggest that the distribution of markup responses is heterogeneous and driven by firm-specific characteristics related to demand elasticity and production scale. Previous literature has considered the response of markups to trade liberalization, but these approaches have focused on the aggregate effect rather than explaining observed heterogeneity. Arkolakis et al. (2012) show that for a particular class of models, trade costs do not affect the distribution of markups. Two relevant assumptions are a particular shape for the distribution of firm productivities (Pareto) and constant returns to scale in production, neither of which are present in our model.

Edmond et al. (2013) study the behavior of markups in a setting where trade is driven by comparative advantage based on Atkeson and Burstein (2008). They find that the effect of trade costs on markups depends on how "close" productivities are between firms that produce the same good, both within and across countries. This follows because this distance determines market shares, and consequently markups. For a common good, when productivity between countries is very different or when producers within a country have very different productivities, markups increase when trade costs decline. Thus, markups tend to increase by less when the environment is more competitive. We are consistent with their result, in the sense that markups in goods that face more elastic demands tend to increase by less (or decrease by more) when trade costs fall. A low elasticity is our proxy for a more competitive sector. Our work complements Edmond et al. (2013) by highlighting the alternative role of love for variety driving trade as in Melitz (2003). A difference is that while Edmond et al. (2013) focus on the comparative advantage gains from trade as in Eaton and Kortum (2002), we focus on love of variety as in Melitz (2003), and we explore the determinants of firm heterogeneity in markup adjustment.

The challenge of dropping the assumption of constant marginal cost is to provide an alternative framework that captures essential characteristics of the data while remaining useful to investigate the complex interactions between interdependent markets. We are not the first to notice the need for decreasing marginal returns. Blum et al. (2013) account for the negative correlation between domestic and export sales growth by developing a framework with physical capacity constraints, which is isomorphic to decreasing returns to scale. Ahn and McQuoid (2013) document similar substitution patterns in both Indonesia and Chile, and find that both financial and physical constraints play an important role in accounting for these observations.

Soderbery (2011) uncovers a similar pattern using firm-level data from Thailand and uses a self-reported measure of firm capacity utilization to study the importance of physical capacity constraints in rationalizing the observed behavior. By using a similar modeling approach of linear demand combined with random and idiosyncratic capacity constraints, he derives conditions under which domestic welfare may decline with the introduction of trade. While his focus is on the qualitative implications of the model, we are interested in using the data to estimate key parameters of the model and then perform counterfactual policy experiments. Our results suggest that substitution patterns are more systematic than would be expected based on random capacity constraint draws.

There is also evidence of decreasing returns to scale from richer economies. Vannoorenberghe (2012) explores output volatility for French firms to conclude that the assumption of constant marginal cost may be unwarranted, while Nguyen and Schaur (2011) use Danish firm data to consider the impact of increasing marginal cost on firm output volatility. Berman et al. (2011) conjecture that capacity constraints might make foreign and domestic market sales substitutes whereas unconstrained firms might see foreign and domestic sales as complements.

The assumption of decreasing returns to scale has been used in theoretical approaches that have considered the dynamics of new exporters (see Ruhl and Willis (2008), Kohn et al. (2012), and Rho and Rodrigue (2012) for example) or in patterns of foreign acquisitions (see Spearot (2012)).

Other studies have departed from the assumption of constant returns to scale by exploring the implications of increasing returns to scale via innovation. Atkeson and Burstein (2010) introduce a costly productivity choice into the Melitz (2003) framework, which effectively introduces technologies with increasing returns to scale. Rubini (2014) shows that this assumption is of particular importance when studying the effects of large macroeconomic changes in trade policy, such as the Free Trade Agreement between the United States and Canada. Our study concentrates on the Chilean economy between 1995 and 2005, a period of relative stability in terms of aggregate imports and exports, suggesting the absence of large changes that require the modeling of innovation.

3 Data

We focus on a panel of Chilean manufacturing firms from 1995 through 2006. This dataset includes all manufacturing firms with 10 or more employees. Standard measures of firm activity are recorded, including information on inputs, outputs, ownership, assets, exporting, and a variety of other measures that provide a complete portrait of the firm. The data has been widely used in empirical studies of firm behavior, most notably in Liu (1993) and Pavcnik (2002). A thorough description of the data can be found in Blum et al. (2013).[5]

Focusing on the sample from 1995-2006, there are 61,548 total observations and 10,163 unique firms. Of these observations, 19,433 belong to firms classified as exporters, meaning that these firms export at some point in the sample. 32% of the sample observations belong to a firm that will export at least once during the sample, or roughly 26% of all firms (2,701 unique firms).

[5] All measures of sales, materials, and capital used in the analysis were deflated using an industry-level price index found in Almeida and Fernandes (2013).

There is a significant amount of switching in to and out of the export market during the sample. In a given year, 2.5% of firms are starting exporting (meaning they did not export in the previous year, but are exporting in the current year) while another 2.5% of firms have ceased exporting. Furthermore, in a given year, 17% of firms are continuing exporting, meaning that they exported in both the previous year and the current year, while 68% of firms are continuing non-exporters (meaning these firms did not export in the last year or in the current year).[6]

The amount of churning at the extensive margin is quite notable. 85% of firms are staying in the market (or markets) that they operated in the previous period, but 15% of firms are operating in a new market (or markets).

Finally, it is important to emphasize that exporting is itself a rare phenomena. If 12% of all observations belong to firms that switch more than once, among the class of all exporters, this accounts for 38% of all observations associated with exporters, while 17% of all exporter observations belong to firms that experience 3 or more changes in their exporting behavior.

We start by documenting significant differences between exporters and non-exporters, which is well attested in the heterogeneous firm literature already. There are statistically and economically significant exporter premia in the data. Summary statistics are reported in Table I.

Exporter Type	Total Sales	Domestic Sales	Employees	Value-added	Investment	Capital	Productivity
Exporter	10643193	6097976	125.2	6246539	464895	6560982	0.5998
	(325034)***	(248487)***	(1.27)***	(227183)***	(35489)***	(303989)***	(0.011)***
Non-Exporter	957116.7	957117	34.86	585302	38096	426037	5.629171
	(182573)	(139673)	(0.71)	(127690)	(19944)	(170828)	(0.006)
N	61,548	61,548	61,548	61,548	56,479	61,548	57,773

Notes: Coefficients from regression of column variable on exporter indicator function. Standard errors are reported in parentheses. (***) indicates coefficient on exporter significant at 0.1 percent.

Table I: Summary Statistics (by Exporter Status)

To identify and quantify patterns of substitution between domestic and foreign sales at the firm level, we calculate correlations between export and domestic sales for each firm. Furthermore, we investigate whether substitution patterns differ significantly across types of firms.

When we consider the correlation between domestic and foreign sales across all firms, we find a raw correlation of 0.16 overall. This might be taken as evidence that exports and domestic sales are complements, but in fact the relationship captures differences between types of firms. By looking across firms, the relationship identified in the data is not a within-firm experience, but

[6] 10% are firms that are new to the sample, or are returning to the sample having been absent in the previous year.

rather captures the fact that larger firms tend to sell more domestically and tend to sell more abroad, which generates the observed positive relationship.

If we focus instead on within-firm behavior, we find a very different story. The aggregate within-firm correlation is -0.18, which is of similar magnitude but the opposite sign when compared to the correlation across all firms. The within-firm correlation is indicative of the fundamental tradeoff firms face when choosing between supplying the domestic market and supplying the foreign market. This result is consistent with previous literature that has identified patterns of substitution between domestic and foreign sales.

The correlation observed in the data might be driven by latent variables and not reflect a direct relationship between domestic and export sales. As will become explicit in our theoretical exposition, one needs to be careful to distinguish between productivity shocks and demand shocks when observing sales across borders for an individual firm since a productivity shock will tend to create a positive correlation between domestic and export sales while individual market demand shocks will generate patterns of substitution. To better get at the direct relationship between domestic and foreign sales, consider the partial correlation after controlling for firm fixed effects as well as year and industry effects found in column (3) of Table II. After partialling out these effects, the overall correlation is -0.19. After calibrating and simulating the model, evaluation of the model will be based on matching this standard, which cannot be matched with constant returns technology assumptions.

	(1)	(2)	(3)	N
Exporters	0.16	-0.18	-0.19	19,443
Firm Fixed Effects	No	Yes	Yes	
Sector Fixed Effects	No	No	Yes	
Year Dummies	No	No	Yes	

Table II: (Partial) Correlation of domestic and foreign sales

Lastly, to motivate our demand side assumptions, we estimate and analyze firm level markups, following the method suggested by De Loecker and Warzynski (2012).[7] The major innovations of this estimation approach are the ability to account for simultaneity in input decisions and the use of a flexible production structure. In particular, the method relies on cost minimization behavior, and is able to identify markups from the elasticity of output with respect to a variable input, after controlling for unobserved productivity. This procedure is consistent with our production and

[7] De Loecker and Warzynski (2012) argue that this methodology may underestimate markups, because of the use the industry price deflators to correct for changes in individual prices.

demand modeling assumptions.

Year	Mean	Median	p5	p25	p75	p95	N
1996	2.298	1.936	0.724	1.338	2.827	5.065	4,369
1997	2.279	1.901	0.699	1.304	2.804	4.958	4,212
1998	2.420	1.813	0.691	1.243	2.716	5.137	4,237
1999	2.180	1.754	0.633	1.184	2.648	4.948	4,039
2000	3.007	1.781	0.635	1.195	2.770	6.197	3,877
2001	2.134	1.665	0.669	1.155	2.511	4.884	3,429
2002	2.223	1.739	0.647	1.153	2.625	5.205	3,847
2003	2.435	1.700	0.541	1.115	2.602	5.189	4,026
2004	2.254	1.786	0.671	1.199	2.676	5.169	3,968
2005	2.240	1.770	0.651	1.184	2.660	5.236	3,953
2006	2.432	1.786	0.658	1.179	2.729	5.694	4,009
Aggregate	2.356	1.787	0.655	1.207	2.697	5.219	43,966

Table III: Distribution of Markups across years

There is overwhelming evidence in the data of heterogeneity of markups at the level of the firm, and these markups change significantly over time as well. Across all observations, the mean markup is larger than the median markup, and this observation holds in each individual year and within each sector (not reported). The average markup for the entire sample is 2.36 while the median markup is 1.79. The skewness is the data is driven by two forces. On the lower bound, firms with markups much below 1 are likely to exit the market since they are not sufficiently covering costs. For the top 5% of firms, markups exceed 5, suggesting a few firms are able to price well above costs.

When looking at the relationship between export status and markups, we find a similar result to De Loecker and Warzynski (2012) in that exporters tend to have larger markups than non-exporters, and this is robust to the inclusion of observable characteristics such as input usage, productivity, industry and year controls. Exporters charge 26% higher markups than non-exporters when looking across firms, which drops to 2.5% when looking at within firm adjustments.

While this evidence is suggestive and worthy of further investigation, given that exporting behavior is not randomly assigned, there should be caution in interpreting these results causally. We will return to these issues when we conduct counterfactual experiments on the simulated data. We now turn to building the theoretical model with these facts and relationships in mind.

ln(markup)	1	2
Export Status	0.259	
	(34.72)***	
Starter		0.025
		(1.92)+
Stopper		-0.009
		(-0.70)
Continuer		0.032
		(2.75)**
Sector FE	yes	yes
Year Dummies	yes	yes
Firm FE	no	yes
Observations	43,975	43,975

Notes: The dependent variable is firm-level markup. Export Status is a 1 when a firm is exporting in that period, and 0 otherwise. The Starter indicator is a 1 when a firm has positive exports in a given year and no export sales in the previous year, and 0 otherwise. The Stopper indicator is a 1 when a firm has no export sales in a given year but had positive exports in the previous year, and 0 otherwise. The Continuer indicator variable is a 1 when a firm has positive exports this period and had positive exports in the previous period, and 0 otherwise. A constant term, capital and labor usage, and firm productivity are included in each regression and omitted in the table. T-statistics are provided in parentheses based on robust standard errors. Significance: + 10 percent; * 5 percent; ** 1 percent, *** 0.1 percent.

Table IV: Markups and Exporting Behavior

4 Model

We use the Melitz and Ottaviano (2008) framework as the building block for our analysis. This has the advantage of generating heterogeneous, endogenous markups in equilibrium while keeping the environment relatively tractable. It does so by assuming preferences that generate linear demands. We extend the model by introducing a fixed factor of production and decreasing returns in the mobile factor to account for the negative correlation between domestic and foreign sales. There are three distinct firm shocks (shock to productivity, domestic demand, and export demand). There is a fixed mass M of firms able to produce differentiated goods. In equilibrium, not all firms will produce because demands may be too low given the productivity shocks. Notice that with linear demands we can generate entry and exit into the export market without fixed (or sunk) costs, so we assume there are none.

Time is discrete. There are two symmetric countries, populated by a continuum of consumers of mass 1. Country H is the Home country and country F is the foreign country.

Consumers. Consumers have within period preferences given by

$$
\begin{aligned}
U =& q_0 + \int_{\Omega_H} \exp(x(\omega))q(\omega)d\omega + \int_{\Omega_F} \exp(y(\omega))q(\omega)d(\omega) - \\
& 1/4\gamma\left(\int_{\Omega_H} q(\omega)^2 d\omega + \int_{\Omega_F} q(\omega)^2 d\omega\right) - 1/2\eta\left(\int_{\Omega_H} q(\omega)d\omega + \int_{\Omega_F} q(\omega)d\omega\right)^2 \qquad (1)
\end{aligned}
$$

where Ω_i is the set of goods produced in $i, i = H, F, q(\omega)$ is the quantity consumed of good ω, $x(\omega)$ is the domestic demand shock for good ω, and $y(\omega)$ is the foreign demand shock for good ω. Given the symmetry of the model across countries, we also use $y(\omega)$ to denote the shocks to foreign demand received by domestic producers. q_0 is a non-traded, numeraire good produced by a stand-in representative firm with linear technology. $\gamma > 0$ and $\eta > 0$ are preference parameters that govern the elasticity of demand and the elasticity of substitution between varieties, respectively. Intuitively, a larger γ reduces the substitutability between tradable goods. Take for instance $\gamma = 0$. In this case, utility is linear, and there is perfect substitution between goods. As γ increases, the substitutability between goods falls. A larger η implies a stronger preference for q_0.

The shocks $x(\omega), y(\omega)$ follow AR(1) processes, given by

$$
\begin{aligned}
x_{t+1}(\omega) &= (1-\rho_x)\bar{x} + \rho_x x_t(\omega) + \varepsilon_{xt}(\omega) \\
y_{t+1}(\omega) &= (1-\rho_y)\bar{y} + \rho_y y_t(\omega) + \varepsilon_{yt}(\omega)
\end{aligned}
$$

where $\varepsilon_{xt}(\omega) \sim N(0, \sigma_x^2), \varepsilon_{yt}(\omega) \sim N(0, \sigma_y^2), 0 < \rho_x < 1, 0 < \rho_y < 1$.

Each consumer has one unit of labor each period which is supplied inelastically. Given prices $p(\omega), p_0$, a wage w and profits π, the budget constraint is

$$
\int_{\Omega_H} p(\omega)q(\omega)d\omega + \int_{\Omega_F} p(\omega)q(\omega)d\omega + p_0 q_0 = w + \pi \qquad (2)
$$

Maximizing the utility function with respect to the budget constraint delivers a demand function that firms take as given when maximizing profits. The inverse demand functions are:

$$
p_H(\omega, q_H) = \exp(x(\omega)) - \eta Q - \frac{\gamma}{2} q_H \qquad (3)
$$

$$
p_F(\omega, q_F) = \exp(y(\omega)) - \eta Q - \frac{\gamma}{2} q_F \qquad (4)
$$

$p_i, i = H, F$ is the price of the good depending on the market where it is sold, and $Q = \int_{\Omega_H} q(\omega)d\omega + \int_{\Omega_F} q(\omega)d\omega$. Notice that the demand for a particular good may be negative, which implies the existence of a choke price above which no quantity will be sold in equilibrium.

Firms. There is one representative firm in the non-tradable sector with technology $q_0(n) = n$.

This sector is perfectly competitive, which implies that in equilibrium, $p_0 = w = 1$.

In the tradable sector, there is one firm per good, acting as a monopolist. There are M firms willing to produce, although not all choose to produce, since linear demands imply that profits could be negative. Firms have constant returns to scale with two inputs to production, capital and labor, and firms own their capital stock. Blum et al. (2013) argue that rigidities in the choice of capital by the firm in the short run are needed to account for the negative correlation between domestic and foreign sales. Since this is a model to study short run effects, we assume that firms cannot modify their capital stock. This implies that firms face effectively decreasing returns to scale. Asker et al. (2014) find that although technologies may be constant returns, there are many rigidities present in the short run such that technologies effectively look as if they have decreasing returns to scale, justifying our assumption.

In addition, firms face idiosyncratic productivity shocks. These are notable for two reasons. First, an improvement in productivity would drive a firm to increase sales both abroad and at home, generating a positive correlation between foreign and domestic sales. Thus, we are not forcing the model to deliver a negative correlation: in principle, the observed correlation can be either positive or negative depending on the relative strength of productivity shocks compared to input rigidities. Second, by bundling together the firm's fixed capital stock and the productivity shock, we can think of the firm as having decreasing returns to scale and a shock to productivity.

The technology for a firm ω with productivity $A(\omega)$, capital stock $\bar{k}(\omega)$, and labor $l(\omega)$ is

$$q(\omega) = A(\omega)\bar{k}(\omega)^{\theta} l(\omega)^{1/\alpha}$$

where $\alpha > 1$ and $\theta > 0$. Since capital is fixed, we can rewrite this as

$$q(\omega) = \tilde{A}(\omega) l(\omega)^{1/\alpha}$$

where $\tilde{A}(\omega) = A(\omega)\bar{k}(\omega)^{\theta}$, so that the firm's choice is only the labor choice.[8]

It is convenient to find the cost function associated with this production function. Taking the wage rate as the numeraire and setting it equal to 1, the cost function of the firm ω is

$$c(q; \omega) = \exp(z(\omega)) q^{\alpha}$$

where $\exp(z(\omega)) = \tilde{A}(\omega)^{-\alpha}$.

Any firm can sell domestically or export. The export cost is a variable iceberg cost, so that if q_F units are to be exported, the producer must produce τq_F units, where $\tau > 1$. Labelling q_H the

[8]Notice that since $\bar{k}(\omega)$ is fixed, we do not impose a market clearing condition on it, as Blum et al. (2013).

units sold domestically, the cost function can be rewritten as

$$c(q_H, q_F; \omega) = \exp(z(\omega))(q_H + \tau q_F)^\alpha$$

$z(\omega)$ follows an AR(1) process:

$$z_{t+1}(\omega) = (1 - \rho_z)\bar{z} + \rho_z z_t(\omega) + \varepsilon_{zt}(\omega)$$

where $\varepsilon_{zt}(\omega) \sim N(0, \sigma_z^2), 0 < \rho_z < 1$. Each period, firms observe their productivity and the demand shocks and solve

$$\max\left\{0, \max_{p_H, q_H, p_F, q_F} p_H q_H + p_F q_F - \exp(z(\omega))(q_H + \tau q_F)^\alpha\right\}$$
$$\text{s.t. equations (3) and (4).} \tag{5}$$

Market Clearing. In equilibrium, all firms producing tradable goods with positive demands ($x > \eta Q$ or $y > \eta Q$) will demand labor units. The representative firm producing non-tradable goods also demands labor. The quasilinear nature of preferences implies that all labor in excess of that demanded by the tradable sector is absorbed by the non-tradable sector. Thus, $\int_{\Omega_H} n(\omega) d\omega + n_0 = 1$, where $n(\omega)$ solves problem (5) and n_0 is the labor demand of the non-tradable sector.

4.1 Equilibrium

While the setup is dynamic, the decisions of the firm are static, since there is no endogenous state variable. An equilibrium is a list of quantities $q(\omega)$ and q_0, labor inputs $n(\omega)$ and n_0 and prices $p(\omega)$ such that consumers maximize (1) subject to equation (2), firms solve (5), and markets clear in every period.

In what follows, it is convenient to drop the name of the good $\omega \in \Omega_H \cup \Omega_F$ and refer to firms by their type, i.e., a triplet (x, y, z). In equilibrium, the solution to problem (5) allows for several corners. In particular, when $\exp(x) < \eta Q$, the good will not be sold domestically, and when $\exp(y) < \eta Q$, it will not be exported. Still, when neither of these conditions are met, it will be sometimes optimal to sell in only one market. The next proposition shows all the possible cases.

Proposition 1. *Let*

$$\tilde{x}(y,z) = \log\left(\gamma\left(\frac{\exp(y)-\eta Q}{\tau\alpha\exp(z)}\right)^{\frac{1}{\alpha-1}} + \frac{\exp(y)-\eta Q}{\tau} + \eta Q\right) \tag{6}$$

$$\tilde{y}(x,z) = \log\left(\frac{\gamma}{\tau}\left(\frac{\exp(x)-\eta Q}{\alpha\exp(z)}\right)^{\frac{1}{\alpha-1}} + \tau\left(\exp(x)-\eta Q\right) + \eta Q\right) \tag{7}$$

A firm x, y, z sells domestically and abroad when $\exp(x) > \eta Q, \exp(y) > \eta Q, x \geq \tilde{y}(x,z)$ and $y \geq \tilde{y}(x,z)$. It only sells domestically when $\exp(x) > \eta Q$ and $y < \tilde{y}(x,z)$, and only exports when $\exp(y) > \eta Q$ and $x < \tilde{x}(y,z)$.

Proof. The proof is detailed in Appendix A. Intuitively, when x is too large relative to y, the firm will not export, since exporting increases its marginal cost given decreasing returns to scale, and it may be optimal to keep these costs low. The opposite happens if y is large relative to x, in which case the firm will choose not to sell domestically and export all its output. □

Proposition 2. *The solution described by proposition 1 is unique.*

Proof. See Appendix A. □

Proposition 1 fully describes the behavior of a firm in equilibrium. Each firm observes its demand functions, which are determined by their demand shocks x and y, and determines whether to sell to both markets, to one, or to none. A firm will not operate in any market when both shocks x and y are too low. It will sell only domestically when x is very large relative to y, it will sell in both markets when x and y are relatively close, and it will only export when y is large relative to x.

5 Calibration

We set $\gamma = 2$, $\eta = 1$ and $\tau = 1.5$. These are normalizations that do not affect the results. The reason is as follows. Consider first η. This determines the degree to which consumers like the tradable good relative to the non-tradable. Since we are not focusing on the non-tradable good, this plays no role.

The parameters γ and τ only affect the value of the estimated parameters for the distribution of shocks in the economy, but not the results or the counterfactuals. To see the intuition behind this, consider for example the role of τ, and the way we calibrate the parameters governing the distribution of foreign demand shocks (which we detail later). These parameters are calibrated to match the share of output exported and the share of firms that export. If one would pick a larger τ, then when it comes to matching the targets one would simply pick a larger mean or variance for the y shock. Similarly, γ affects the substitutability between varieties, which in turn determines the markup. Since we calibrate parameters to match the distribution of markups in the data, this again would simply affect the estimates of the distribution, not the results.

The reason why the counterfactuals are not affected is that how markups react to changes in trade costs depends on the elasticity of demand, both domestic and foreign, a calibration target. The model delivers by construction the markups we observe in the data, and this pins down the elasticities of demand. So choosing a different τ, for example, and recalibrating everything to match the moments we match, in particular the distribution of markups, would yield the same effect of a change in trade costs on markups. We have done sensitivity experiments to confirm that the results do not depend on these parameters, which are available on request.

To determine α, we rely on the estimate in Coşar et al. (2010), and set $\alpha = 1.69$. We later perform sensitivity experiments to show how this choice affects the results. Coşar et al. (2010) estimate this parameter via Generalized Method of Moments in a model where firms producing tradable goods have decreasing returns to scale in a perfectly competitive environment. Since our models differ in structure and competitive environment, we use this estimate as a starting point before conducting sensitivity analysis in Section 8 to explore how our results depend on α.

For the parameters governing the distribution of firms we use firm level data on domestic revenues, exports and markups to back out the unobserved triplet (x, y, z) consistent with the observed data. While we observe exports and domestic sales directly from the data, we rely on De Loecker and Warzynski (2012) to estimate markups. This procedure forces us to eliminate some data that exhibit features not consistent with our model, such as negative markups. We treat each year as a different cross-section, which gives us a total of 21,441 observations to calibrate the model.[9] Also, we discard firms that the model suggests they should export but not sell domestically, on the basis that this does not happen in the data (there are only 18 firm-year cases in the entire sample, and only 3 firms that do this every period).

The calibration strategy works in two steps. The first step calibrates the cross-section parameters, and the second deals with the time-series components. Our theory predicts that as $t \to \infty$, the economy converges to the following invariant distributions of shock realizations:

$$x \sim N\left(\hat{x}, \frac{\sigma_x^2}{1-\rho_x^2}\right), y \sim N\left(\hat{y}, \frac{\sigma_y^2}{1-\rho_y^2}\right), z \sim N\left(\hat{z}, \frac{\sigma_z^2}{1-\rho_z^2}\right)$$

We use the theory to back out the shock realizations in the data and then estimate μ_i and $\hat{\sigma}_i$, where $\hat{\sigma}_i = \frac{\sigma_i^2}{1-\rho_i}$, for $i = x, y, z$ via maximum likelihood in the cross section. The time-series calibration then identifies ρ_i and σ_i from $\hat{\sigma}_i$.

The way to back out the shock realizations is the following. The model implies that the triplet

[9] See Appendix B for further details.

(x, y, z) determines domestic sales, exports, and markups for each firm. Using data on domestic sales, exports, and markups, we can thus reverse engineer the decision process and identify the shocks.

Given the realization of the shocks, we compute the parameters of interest via maximum likelihood. This introduces a problem in the estimation of the y shocks, since the fact that we observe exports means that the shocks were sufficiently high, and therefore our sample is biased and not reliable for maximum likelihood.[10] We deal with this by calibrating the parameters μ_y and $\hat{\sigma}_y$ to match the share of output exported and the share of firms that export. These are 37% and 29%, respectively.

This procedure assumes that we know the value of Q. Fortunately, the free parameter M (the exogenous mass of firms) determines Q. So we set $Q = 1$ and back out the M that is consistent with this equilibrium value. We do this for all firms in all years from 1996 through 2005 (the estimation of markups requires us to drop 1995). Figures 1 and 2 show the histograms of the shock realizations that we backed out from the data. At first sight, the assumption of a normal distribution seems to be reasonable.

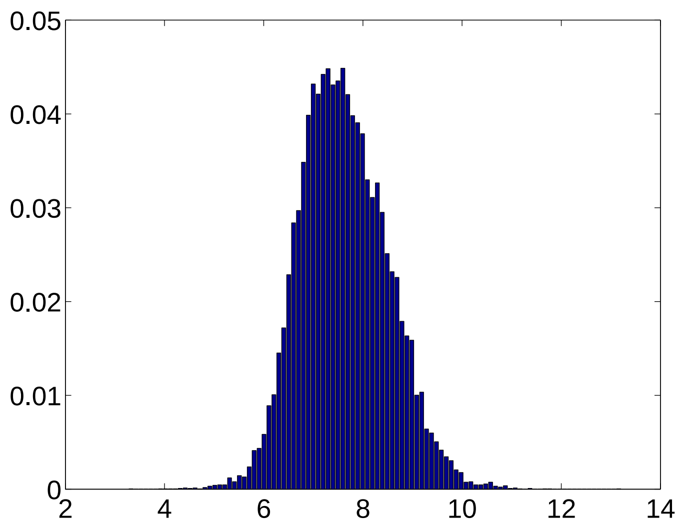


Figure 1: Distribution of domestic demand shocks backed out from data.

The last step involves separating σ_i from ρ_i, for $i = x, y, z$. Ideally, we would compute them performing regressions on each variable on its lags. The problem is that the observed data for x and y is biased, and as such the errors would not be zero mean, so we choose an alternative

[10]One can also argue that the observed distribution is biased, since a firm will sell domestically only when $\exp(x) > \eta Q$. This bias is easy to correct. We found that in general this restriction is not binding, and the results of correcting or not correcting are very similar, so we ignore this bias.

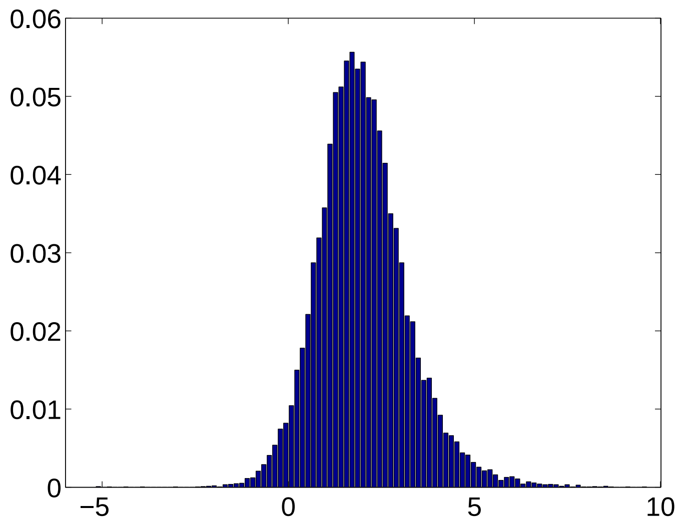


Figure 2: Distribution of productivity shocks backed out from data.

approach. We perform a simulated method of moments that works as follows. We first simulate the behavior of 160,000 firms for 1,000 periods, and keep only the last 10. Then we keep only firms with positive exports every period or zero exports every period.[11] Then we compute three autocorrelation coefficients: domestic sales for non-exporters, domestic sales for exporters, and exports for exporters. These autocorrelations in the data are 0.43, 0.39 and 0.42, respectively. We compare these to the same autocorrelation coefficients in the data. The calibration changes ρ so that the distance between data and model is as close as possible. Table V shows all the parameter values.

Notice the differences in the distributions of the x and y shocks. While on average the x's are larger, the y's have a larger standard deviation. This is important, because low numbers for y do not matter (the firm will be a non exporter), so a large standard deviation can generate large trade volumes in spite of small means.

5.1 Fit of the Model

Before we move on to the findings, it is interesting to compare the simulations of the calibrated model with the data along calibrated dimensions, mainly, the cross-section of domestic sales, foreign sales, and markups.

Figures 3 through 5 compare these distributions in the model and the data. In all cases, the model distribution is quite close to the data distribution, indicating that our calibration strategy is

[11]We discard firms that enter and exit the export market because these will exhibit changes in domestic sales that are too abrupt, and the autocorrelation coefficient will be less informative of the random shock processes.

Parameter	Value	Target
α	1.69	Coşar et al. (2010)
η	1	Normalization
γ	2	Normalization
τ	1.5	Normalization
M	8×10^{-4}	Sets $Q = 1$
$\bar{x}$	7.60	Maximum Likelihood
$\bar{z}$	1.89	Maximum Likelihood
$\bar{\sigma}_x$	0.89	Maximum Likelihood
$\bar{\sigma}_z$	1.26	Maximum Likelihood
$\bar{y}$	6.53	Exports to sales ratio = 37%
$\bar{\sigma}_y$	1.61	Share of exporters = 29%
ρ_x	0.86	Method of Simulated Moments
ρ_y	0.96	Method of Simulated Moments
ρ_z	0.94	Method of Simulated Moments
σ_x	0.45	From ρ_x and $\bar{\sigma}_x$
σ_y	0.48	From ρ_y and $\bar{\sigma}_y$
σ_z	0.41	From ρ_z and $\bar{\sigma}_z$

Table V: Calibrated Parameters

quite successful at matching the intended targets.

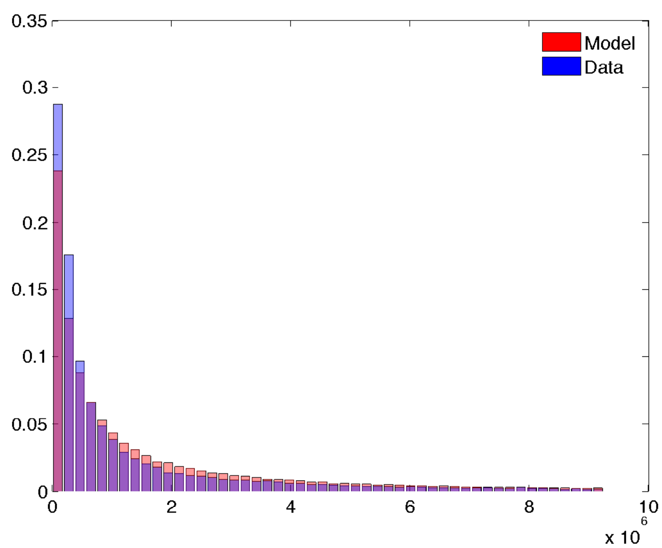
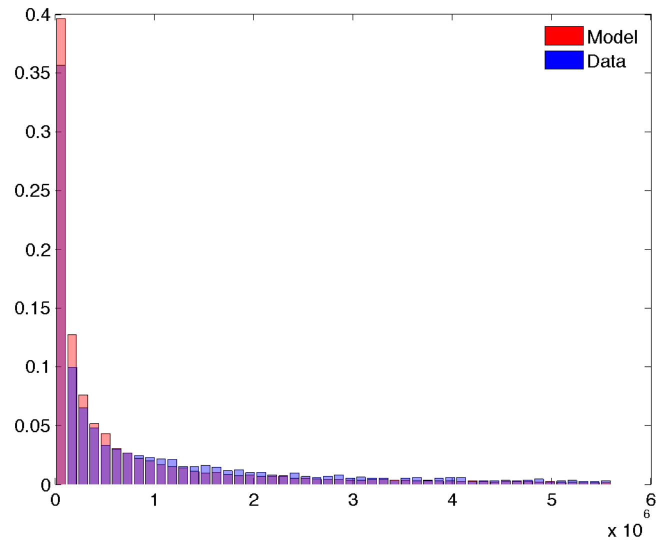


Figure 3: Distribution of domestic sales: model vs. data.

It is not obvious that we should be able to reproduce the distributions in the data, since we



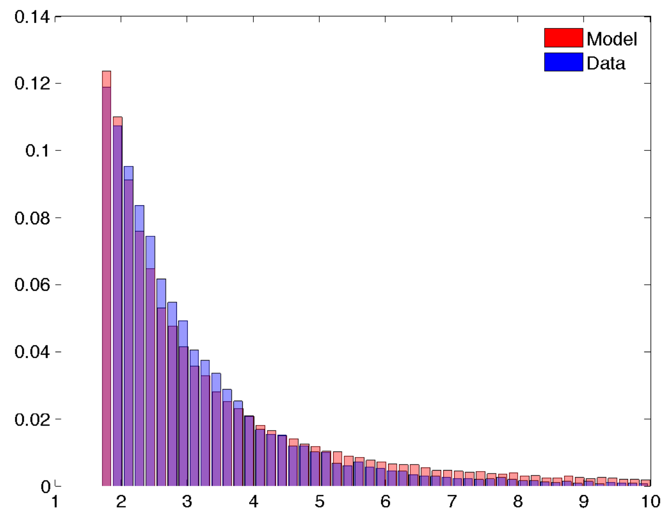


Figure 5: Distribution of markups: model vs. data.

are assuming that the shock processes are not correlated. If they are, the simulated model might deliver different distributions than the data.

For example, if the correlation between x and y in the data is positive, then we would expect firms with large x shocks to have lower domestic sales than in the model, since these firms also have large exports. In the model, given the zero correlation between x and y, the firm might not export, and therefore allocate all its resources to the domestic market, producing larger domestic sales. The fact that the simulated distribution is similar to the data's implies that these correlations are not that strong.

6 Stylized Facts

This section focuses on the ability of the model to match the stylized facts previously discussed. These are: exporters charge markups that are on average 26% larger than non-exporters; entering the export market is associated with a markup increase of 2.5%; and the correlation between domestic and foreign sales is -0.19.

To test the model we perform simulations of the calibrated model. The exercise consists of simulating the behavior of 160,000 firms[12] for 1,000 periods, and keeping only the last 10 periods, to be consistent with data we are working with.

6.1 Markups

A key variable of interest in the trade literature is the effect of export entry on the markup of the firm. De Loecker and Warzynski (2012) find that: (i) exporters charge higher markups than non-exporters in the cross section; and (ii) entering the export market increases the markup.

Given the assumption of decreasing returns to scale, one cannot separately identify domestic and foreign markups. Both empirically and in the model, we measure markups as the ratio of total revenues to total costs. That is,

$$Markup_{it} = \frac{p_{d,it}q_{d,it} + p_{x,it}q_{x,it}}{\exp(z_{it})(q_{d,it} + \tau q_{x,it})^{\alpha}}$$

where $p_{d,it}$ is the price at which firm i sells domestically in period t, $q_{d,it}$ is the domestic quantity sold, and replacing d with x is analogous for exports.

Simulating the model, the cross-section shows that exporters charge, on average, a markup that is 37 percent larger than non-exporters. Markups increase by 1 percent when firms start to export.

Being so close to the data is remarkable, since these observations were not targeted in the calibration. Notice that this takes place under constant trade costs: other things change, namely productivity and demand, which drive a firm to export.

6.2 Prices

We next ask how entering the export market affects the average price a firm sets for their goods. The aim of this section is to compare ourselves with García Marin and Voigtländer (2013), who find that average prices drop by 11% when entering the export market. They compute average prices using data on physical quantities, data that we do not have. That is, using only single product firms, they compute $p_{av} = \frac{total-revenues}{units\ sold}$.

[12]Increasing the number of firms does not change the results in any considerable way.

We compute the average price in our simulated data as $p_{av} = \frac{p_d q_d + p_x q_x}{q_d + \tau q_x}$. We find that average prices drop by 10.1 percent, which is quite close to García Marin and Voigtländer (2013)'s findings.

6.3 Correlation between Exports and Domestic Sales

An important question in the paper is whether the model can account for the correlations between domestic and foreign sales in the data. In the data, the correlation is -0.19, while the model produces an aggregate correlation of -0.15. That is, the model accounts for 79% of the correlation observed in the data.

Given the amount of heterogeneity in the data, we analyze this correlation for different groups of firms, depending on the frequency of exports. We disaggregate firms between those that are always observed exporting, those that export between 90 and 100% of the time, 75-90%, and 50-75% of the time. Table VI summarizes our findings for different types of firms.

In the data, the correlation increases with exporting frequency: those firms that export more frequently show a larger correlation. In fact, the correlation is positive for firms that export more than 90% of the time.

The model can account for this pattern well. As in the data, the correlation increases with export frequency, and firms that export over 90% of the time exhibit a positive correlation. However, the model cannot generate changes as large as in the data across different groups of firms.

Firm Type	Data	Model
All exporters	-0.19	-0.15
Export 100% of periods	+0.19	+0.04
Export 90%-100% of periods	+0.13	+0.01
Export 75%-90% of periods	-0.31	-0.10
Export 50%-75% of periods	-0.37	-0.18

Table VI: Correlations and Export Frequency

Using the model, we can ask the reason for the observed relationship between exporting frequency and the correlation. We conjecture that this is due to the shape of the cost curve. Marginal costs are $\exp(z)\alpha Q^{\alpha-1}$, where Q is units produced. Since $1 < \alpha < 2$, this function is increasing and concave, so that marginal costs increase at a decreasing rate. Thus, marginal costs are relatively flatter (closer to constant returns) for larger firms. When marginal costs are flatter, the effect of decreasing returns becomes less important, and the positive effect of productivity on the correlation dominates, generating a positive correlation.

If this is the reason for the positive relation between correlation and exporting frequency, we should expect size and exporting frequency to be positively correlated. We verify this by regressing the following:

$$\log(Q_{it}) = \beta_0 + \beta_X N_X + \epsilon_{it}$$

where N_X is the number of periods with positive exports.

Our estimates confirm that exporting frequency is positively related to size. We estimate $\beta_X = 0.1432$, which implies that exporting for one additional period increases production by about 14%. This is significant at the 1 percent level. If we replace physical quantities with sales, we still get the same effect, with one additional year of exporting increasing total sales by 11%.

One reason why our model delivers correlations that are less extreme than the data may have to do with the correlation between size and the number of export destinations. While our data does not have the number of export destinations, it is usually the case that larger exporters also export to more countries (see, for example, Bernard et al. (2007)). Assuming that the demand shock from each country is independent, a firm exporting to a larger number of countries faces less aggregate demand fluctuations, and therefore the effect of changes in productivity have a larger weight on the correlation between domestic and foreign sales. Similarly, small firms, by exporting to fewer countries, have higher demand volatility, so demand shocks are key in driving the correlation. While there is only one country in which firms can export to in the model, by calibrating our model to aggregate statistics (export volume and fraction of firms exporting) we are by construction targeting averages, which is why we perform better in the aggregate than when disaggregating.

6.4 Export vs. Domestic Markups

The methodology we employ to measure markups in the data follows De Loecker and Warzynski (2012), and this prevents us from studying domestic and foreign markups separately. Note that the assumption of decreasing returns to scale also suggests that these two markups are not easy to separate, since the marginal cost is the same one. However, prices are not the same, and in this sense markups are not either.

An alternative approach to markup estimation outlined by Jamandreu and Yin (2014) estimates, for a given firm, the foreign markup relative to the domestic markup. Their methodology makes minimal assumptions on cost minimizing behavior. They rely on data on exports, total sales, and variable costs, all of which we have. This provides a second, independent measure to test the validity of our model. An important limitation is that it is only valid for exporting firms, and therefore it cannot be used to compare the markup of an exporter with that of a non-exporter, on of the main points of this paper.

Jamandreu and Yin (2014) use their methodology on a dataset of Chinese firms. They find that

Chinese exporters charge lower markups on exports than on domestic sales. Using their approach, we find this is also true for Chilean exporters.

When we restrict our sample to only exporters, we estimate that markups on exports are approximately 9 percent lower than in the domestic market, consistent with the results in Jamandreu and Yin (2014). When year specific effects are removed from the data, the estimated difference is reduced somewhat, implying that foreign markups are 6.4 percent lower than domestic markups. However, as discussed above, firm heterogeneity is significant in a multitude of dimensions, including markups. As such, this aggregate estimate (based on pooled firm observations) may be hiding substantial heterogeneity in markup setting behavior at more disaggregated levels.

This method can be applied at lower levels of aggregation as well. We separately estimate markups for 75 distinct 4-digit industries with sufficient observations for estimation. On average, estimated markups across these 75 industries are 15 percent lower in the export market compared to the domestic market, with the median industry estimate being 9.5 percent lower. There is significant heterogeneity across industries. Figure 6 shows a histogram of the foreign markup relative to the domestic one, where a positive 1 means that the foreign markup is 100 percent larger than the domestic one. When we include time fixed effects, these estimates are slightly larger in absolute value, with the average industry charging a foreign markup that is 21 percent lower than the domestic markup (the median is 12 percent lower markup abroad).

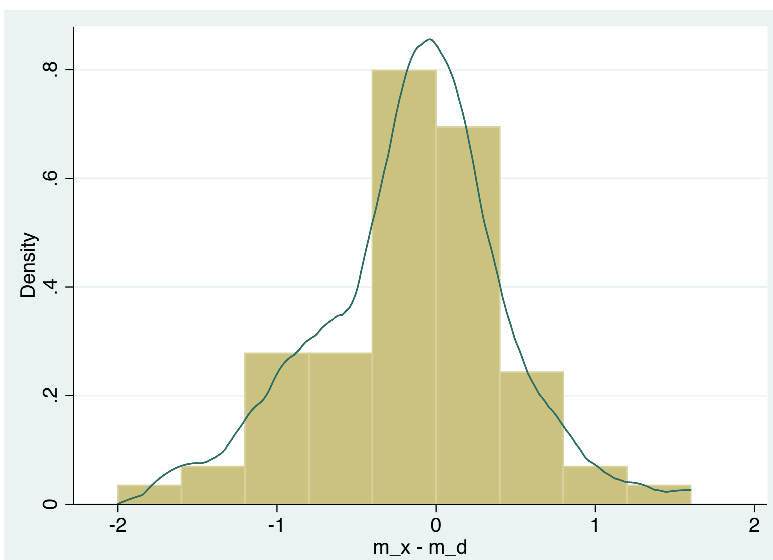


Figure 6: Foreign Minus Domestic Markups Across 75 4-digit Industries

To best match our quantitative exercise, we next estimate markup differences firm by firm, resulting in markup estimates for over 1,700 individual firms. There is significant heterogeneity and skewness in these estimates, but the median suggests markups abroad are 5 percent lower (the average is contaminated with very high extremes, so we do not report it, but removing the top and bottom 5 percent of firms yields an average markup that is 3 percent lower abroad). While there is substantial heterogeneity across firms, most firms charge lower markups in foreign markets compared to domestic markups. We show the distribution of these markup premia in Figure 7.

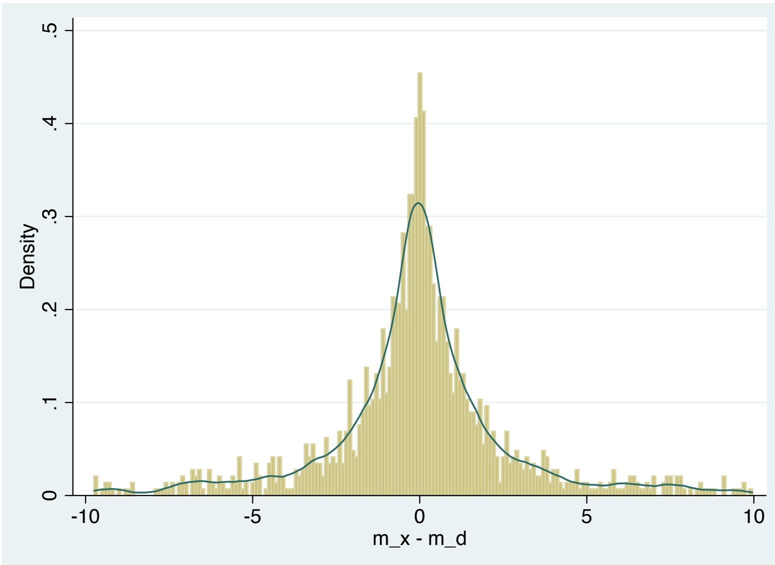


Figure 7: Foreign Minus Domestic Markups Across All Exporters

If we focus on a balanced panel of perennial exporters we end up with 170 firms that are observed to export in all 11 years. The median percentage difference in estimated markups is 10 percent lower abroad, and the average percentage difference in estimated markups is 13 percent lower abroad. We show this in Figure 8.

In the model exporters on average charge higher markups domestically than abroad. Figure 9 shows the distribution of relative markups at home and abroad. The x-axis measures foreign minus domestic markups, and the y-axis measures the frequency of each observation. It is easy to see how the distribution is skewed to the left, that is, on average the numbers are negative. In fact, on average domestic markups are about twice as large as foreign markups (2.15 times), and 63 percent of exporters charge markups that are higher domestically than abroad. These numbers are much higher than the numbers found in the data, but they are qualitatively similar: both in the data and the model, exporters on average charge higher markups at home than abroad.

Average markups are also larger in the domestic market than in the foreign market. Figure 10 shows the distribution of foreign and domestic markups separately when considering all exporters. The average domestic market is about twice (2.3 times) the foreign markup.

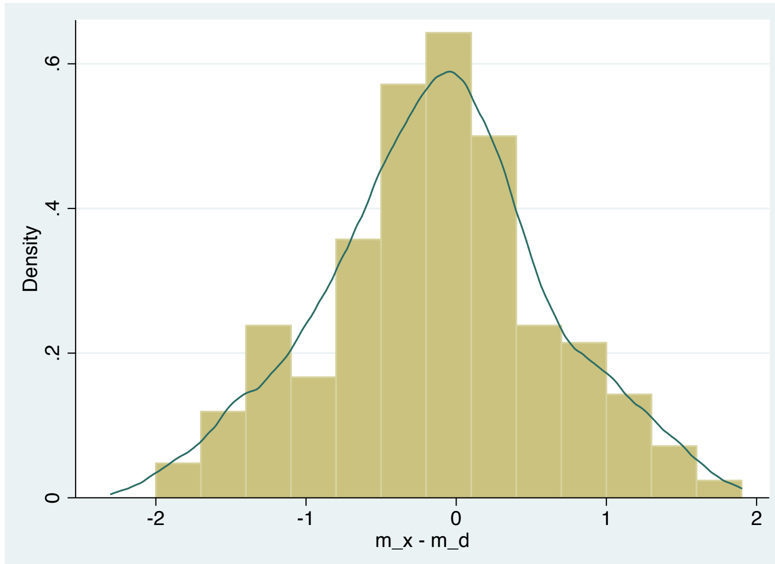


Figure 8: Foreign Minus Domestic Markups Across Firms that Always Export

This is also consistent with Impullitti and Licandro (2010), who find by using a calibrated model that exporters charge higher markups at home than abroad.

This suggests that foreign demands are, on average, more elastic, and therefore a change that shifts output toward the export market should reduce markups. However, the fact that reducing trade costs reduces the marginal cost would lead to an increase in markups. The next section explores quantitatively how markups and trade costs interact.

The fact that our model matches these aspects of the data, even though they were not directly included in the calibration and estimation process, provides additional evidence in support of the approach. Furthermore, the robustness of our analysis using multiple independent approaches to markup estimation at the firm level minimizes concerns over the fragility of our markup results.

7 Main Findings

The previous section shows that the model can match well the stylized facts, and this suggests that the model is reliable to determine the effects of trade costs on markups. To determine these effects, we drop trade costs from our benchmark value of 1.5 to 1.1.[13]

We find that, among exporters, the response of markups to trade costs is quite heterogeneous. On average, they increase by 7.7%, and the median increase is almost 3%, although there is a lot of heterogeneity. In fact, 57% of firms increase their markups and 43% reduce it. Figure 11 is a

[13] We perform both general equilibrium counterfactuals (that is, the aggregate Q in equations (3) and (4) change) and in partial equilibrium (with Q unchanged). The results under both specifications are very similar, so we report the results under general equilibrium only.

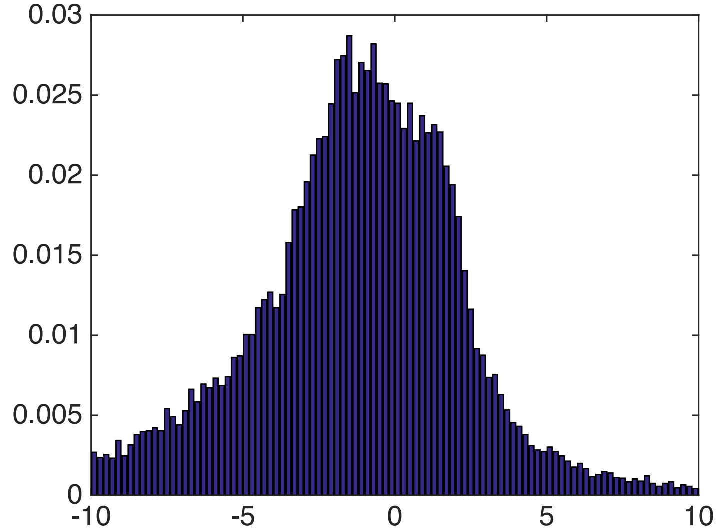


Figure 9: Distribution of foreign minus domestic markups for each exporter

histogram of the change in markups including all firms that export when $\tau = 1.1$.

To better understand what drives the change in markups, we split our sample of firms into two groups. The "intensive margin" group includes firms that were already exporting before the change in trade costs. The "extensive margin" group includes firms that started exporting only after the reduction.

7.1 The Intensive Margin

Markups usually increase along the intensive margin, although not always. The median increase is 7 percent, and the average increase is 11 percent. About 29 percent of firms reduce their markups. Figure 12 shows the change across different firms. The extent to which markups increase is related to how elastic domestic and foreign demands are. To explore this further, we run the following regression:

$$\Delta Markup = \beta_0 + \beta_1 |\eta_d| + \beta_2 |\eta_x| + \epsilon$$

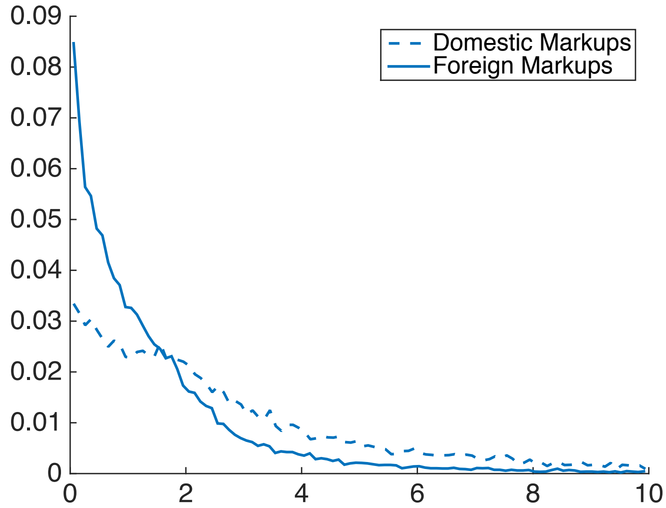


Figure 10: Distribution of foreign and domestic markups

where $\Delta Markup$ is the percentage change in markup, and $|\eta_d|$ and $|\eta_x|$ are the absolute value of the elasticity of demand to prices, domestic and foreign, respectively.[14]

What truly matters in determining markups is the elasticity of foreign demand. The point estimate for β_2 is $-1.5e^{-4}$, significant at the 1% level ($\beta_0 = 1.1$ and β_1 is not significant at the 1% level). When foreign elasticity is very high, firms find it optimal to lower their prices more, expanding their output by more, and generating a smaller increase in the markup.

The reason why some firms increase their markups is that the reduction in trade costs is a reduction in marginal costs, and these firms do not fully pass on this decline in costs to prices. In fact, only under constant elasticities of demand, where markups are constant, will firms pass the reduction entirely on to the consumer. In this case, the reduction in price is less than the reduction in cost, resulting in an increased markup.

Similarly, the reason why some firms reduce their markup is because of the increasing marginal cost. The drop in trade costs produces an increase in output, and this increases marginal costs. Again, firms only partially pass on this increase to the consumers, thus lowering markups. In fact, the correlation between changes in markups and changes in marginal costs among these firms is -0.87, showing that changes in marginal costs drive almost all the changes in markups.

A natural question is whether firms that increase their markups become less efficient, in a Pareto

[14] We evaluate elasticities as the average of the elasticity before and after the change in trade costs. We compute the elasticities as follows

$$\eta_x = \frac{\partial q_x}{\partial p_x}\frac{p_x}{q_x} = -\frac{e^y - \eta Q - \gamma/2q_x}{\gamma/2q_x}, \eta_d = \frac{\partial q_d}{\partial p_d}\frac{p_d}{q_d} = -\frac{e^x - \eta Q - \gamma/2q_d}{\gamma/2q_d}$$

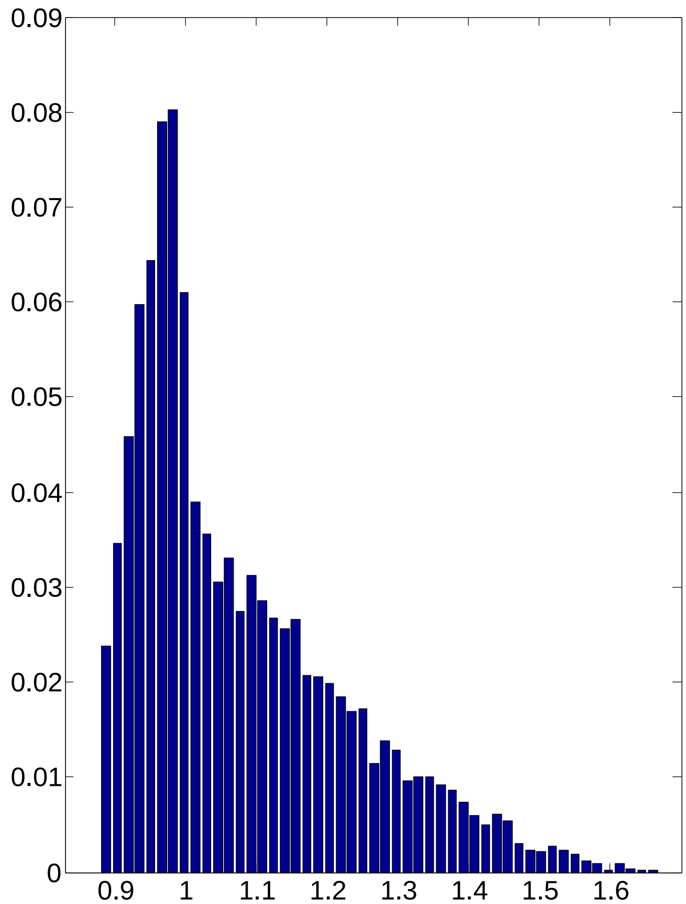


Figure 11: Changes in Markups Among Exporters

sense. Efficiency would require markups to be zero, since price should equal marginal cost. Thus, an increase in markups looks like an efficiency loss. But the reduction in trade cost, which is waste in this model, constitutes a gain in efficiency. It turns out that the increase in markups never fully offsets the reduction in trade costs, since the price of exports drops for these firms, as we show in Appendix C. Thus, intensive firms become more efficient following the drop in trade costs.

7.2 The Extensive Margin

The behavior of markups is very different among firms that only export under the low trade cost regimes. These firms lower their markup. The median and average markup falls by 5 percent. No firm increases the markup. Figure 13 shows the distribution of the changes in markups for these firms.

This reveals that firms entering the export market when trade costs drop reduces the markup. Recall that the model can replicate well the fact that markups for exporters are larger in the cross-

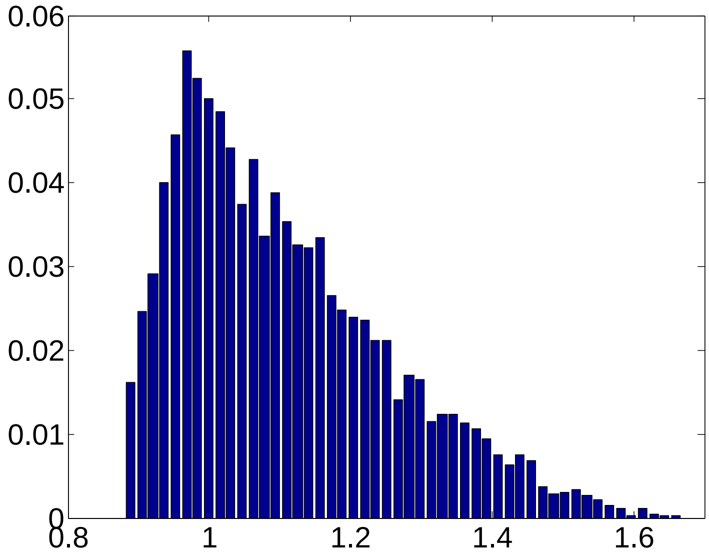


Figure 12: Changes in Markups Along the Intensive Margin

section, and that firms increase their markup when they start exporting under constant trade costs. Nonetheless, when firms enter the export market *because* trade costs decline, markups decrease. Our counterfactuals suggest that exporters share certain characteristics (higher foreign and domestic demand) that imply large markups and exporting. That is, the large markup is not a consequence of a low trade cost.

To explore deeper the drivers of the change in markups we perform a regression similar to the one related to changes along the intensive margin.[15] Thus, we regress

$$\ln(\Delta Markup) = \beta_0 + \beta_1|\eta_d| + \beta_2|\eta_x| + \epsilon$$

Our results are similar to those for the intensive margin. The key estimate is $\beta_2 = -3e^{-4}$, significant at the 1% level ($\beta_0 = 0.97$ and β_1 is not significant at the 1% level). Thus, larger foreign elasticities lead to lower markups.

The drop in markups makes these firms more efficient. Thus, both firms along the intensive and extensive margins become more efficient.

Appendix C explores the effect of the reduction in trade costs on other margins, mainly sales and prices.

[15]The difference is that η_x is evaluated only at the point with low trade costs.

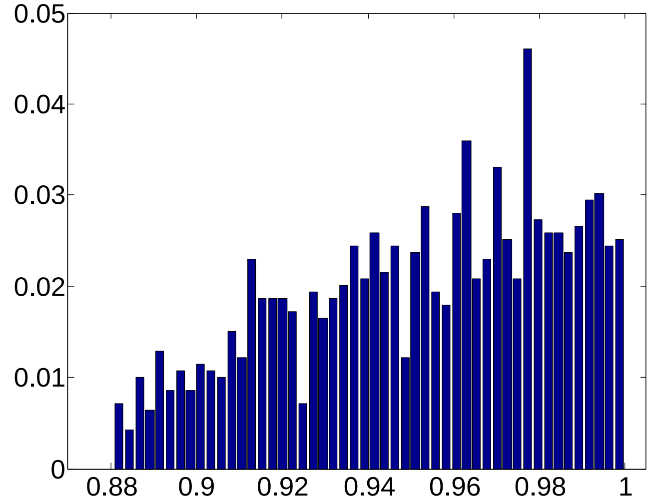


Figure 13: Change in Markups Along the Extensive Margin

8 Sensitivity Analysis

A key parameter that we calibrate by following other related studies is α, which determines the curvature of the cost function. In this section, we show the results of changing α, while keeping all other parameter values unchanged. Intuitively, a larger α implies a greater degree of decreasing returns to scale, so the correlations between domestic and foreign sales should be decreasing in α. We confirm this in our exercises when we change α and focus on the aggregate correlation between domestic and foreign sales. The upper panel of table VII shows that reducing α from 1.69 to 1.5 increases the aggregate correlation from -0.15 to -0.07, and increasing α to 1.95[16] reduces the correlation to -0.25.

When we disaggregate these correlations dividing firms into their export frequency, we note that this change does not translate smoothly into each subgroup. In fact, while the correlation among infrequent exporters (firms exporting less than 75% of the time) shows similar changes as the aggregate correlation, the correlation among frequent exporters does not. Contrary to the intuition previously described, the correlation for firms that export 100% of the time actually decreases when we move from $\alpha = 1.69$ to $\alpha = 1.5$.[17]

What explains this change is that more firms export 100% of the time. Panel 2 of Table VII shows the fraction of firms in each export category. Under $\alpha = 1.69$, 14% of exporters export every period. This number increases to 39% when $\alpha = 1.5$. This implies that smaller firms enter this

[16]A larger value of α implies less firms export all the time. When α is larger than 2, some simulations show no firm exporting all the time which prevents us from computing the correlations.

[17]It increases when moving from $\alpha = 1.69$ to $\alpha = 1.95$, although this isn't apparent given our choice to round to 2 decimal points.

Correlations and Export Frequency			
Firm Type	Benchmark (α = 1.69)	α = 1.5	α = 1.95
All exporters	-0.15	-0.07	-0.25
Export 100% of periods	+0.04	+0.02	+0.04
Export 90%-100% of periods	+0.01	+0.01	-0.02
Export 75%-90% of periods	-0.10	-0.07	-0.14
Export 50%-75% of periods	-0.18	-0.13	-0.23

Share of Firms and their Export Frequency			
Share of Exporters	Benchmark (α = 1.69)	α = 1.5	α = 1.95
Export 100% of periods	14%	39%	1%
Export 90%-100% of periods	20%	47%	2%
Export 75%-90% of periods	12%	15%	2%
Export 50%-75% of periods	21%	18%	10%

Table VII: Correlations under different values of α

group, and these firms have steeper marginal cost curves, thus producing lower correlations and driving averages down.[18]

The reason why firms export more often when α is lower is that exporting is more attractive, since expanding output does not carry such a large increase in marginal costs. In fact, entering the export market (under constant trade costs) is associated with an increase in markup of 10 percent when $\alpha = 1.5$, compared to 1 percent in the benchmark case. When $\alpha = 1.95$, markups tend to drop by 3 percent when entering the export market.

Also, under $\alpha = 1.5$, exporters change a markup that is 55 percent larger than non-exporters (against 16 percent in the benchmark economy). When $\alpha = 1.95$, this premium is only 14 percent.

The counterfactuals also change in the expected direction. When dropping trade costs from $\tau = 1.5$ to $\tau = 1.1$, the average increase in markups along the intensive margin is 16 percent, compared to 11 percent in the benchmark case. Also, about 18 percent of firms reduce their markups, compared with 29 percent in the benchmark.

The intuition behind this is simple. Recall that the reason why some firms reduce their markups is related to an increase in their marginal costs, produced by the expansion in output. Under a lower α, the expansion in output affects marginal costs less, and therefore they increase less. As a result,

[18] Note that the percentages may add up to more than 100. This is because some firms are in more than one group. For example, the 90-100 group includes firms in the 100 group.

the effect of the fall in trade costs becomes more important, on average firms increase their markups by more, and fewer firms reduce their markups.

The effects along the extensive margin are similar. When $\alpha = 1.5$, the average reduction in markups is of 3.5 percent, compared to 4.5 percent in the benchmark. Again, the reason is that the increase in output does not increase marginal costs as much, therefore dampening the reduction in markups.[19]

8.1 Constant Returns to Scale

The last case we explore is that of constant returns to scale, that is, $\alpha = 1$. We only note the effects of trade costs on markups. Again, we do not recalibrate the entire model, we simply change the parameter α.

Along the intensive margin, 98 percent of firms increase their markups, that is, almost no firm reduces their markup. This is because the expansion in output is not increasing marginal costs, and therefore marginal costs decrease because of the reduction in trade costs. The 2 percent that actually reduce markups are responding to a very elastic foreign demand curve that affects them more under lower trade costs. On average, the increase in markups for this group of firms is 13 percent, larger than with decreasing returns to scale.

Along the extensive margin, there is very little change in markups. No firm increases its markup as before, but the reduction is now very mild. On average, it is less than 0.5 percent (compared to 10 times more in the benchmark case). Actually, the firm that reduces its markup the most reduces it by less than 5 percent, that is, less than the average change in the benchmark case.

The fact that markups along the intensive margin drop under constant returns shows that the foreign elasticity of demand is larger than the domestic one. On the other hand, the fact that the reduction is much lower than under decreasing returns to scale shows that quantitatively, what matters most is the convexity of the cost function to determine the drop in markups.

9 Conclusion

Understanding the effects of trade costs on markups is key for determining the effects of trade liberalization. We find very heterogeneous responses which depend on key firm characteristics. Along the intensive margin, a reduction in trade costs tends to increase markups, although this is not true for firms with very elastic foreign demands. Along the extensive margin, markups decrease.

[19] We omit the counterfactuals when $\alpha = 1.95$, but all the changes are in the opposite direction, as expected.

A crucial assumption for understanding these two results is the presence of a fixed factor of production, and decreasing returns to scale in the mobile factor. Along the intensive margin, a decline in trade costs represents a decline in marginal costs, which firms only partially pass on to prices, but the expansion of output associated with the increase in exports increases marginal costs, which again are only partially passed on to prices. Differences in market elasticities also play a role, with the ultimate effect on markups depending upon the relative strength of these three competing forces. Along the extensive margin, only the scale and elasticity effects are operational, unambiguously reducing markups since foreign demand faced by these firms is more elastic than domestic demand.

The concluding message that can be extracted from this paper is that exporters charge higher markups than non-exporters not because they export, but in spite of it. The same reasons that drive these firms to export (high productivity, or high foreign demand) also drive them to set relatively high markups. In any case, exporting reduces their markups: foreign elasticity is higher, and exporting firms tend to be larger, which together with increasing marginal costs, implies that they face higher marginal costs.

Our study has strong implications in terms of the efficiency effects of trade liberalizations. In an efficient allocation, markups are zero, so a reduction in markups implies a gain in efficiency. In this sense, firms that enter the export market following a reduction in trade costs become more efficient, as their markups decline. On the other hand, markups increase for most firms that were exporting before the change in trade costs. But these firms actually gain in efficiency since trade costs fall, which is only partially offset by an increase in markups. Thus, both extensive and intensive margin exporters become more efficient as trade costs fall. Allowing markups to be endogenous and heterogeneous therefore provides an additional channel by which trade liberalization provides a gain in efficiency.

References

Ahn, JaeBin and Alexander McQuoid, "Capacity constrained exporters: micro evidence and macro implications," 2013.

Almeida, Rita and Ana M. Fernandes, "Explaining local manufacturing growth in Chile: the advantages of sectoral diversity," *Applied Economics*, June 2013, *45* (16), 2201–2213.

Arkolakis, Costas, Arnaud Costinot, Dave Donaldson, and Andres Rodriguez Clare, "The Elusive Pro-Competitive Effects of Trade," 2012.

Asker, John, Allan Collard-Wexler, and Jan De Loecker, "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 2014, *122* (5), 1013 – 1063.

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

_ and _ , "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, 06 2010, *118* (3), 433–484.

Berman, Nicolas, Antoine Berthou, and Jerome Hericourt, "Export Dynamics and Sales at Home," August 2011.

Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott, "Firms in International Trade," *Journal of Economic Perspectives*, Summer 2007, *21* (3), 105–130.

Blum, Bernardo, Sebastain Claro, and Ignatius Horstmann, "Occasional vs Perennial Exporters: The Impact of Capacity on Export Mode," *Journal of International Economics*, May 2013, *90, 1*, 65–74.

Coşar, A Kerem, Nezih Guner, and James Tybout, "Firm dynamics, job turnover, and wage distributions in an open economy," Technical Report, National Bureau of Economic Research 2010.

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

_ and Pinelopi Koujianou Goldberg, "Firm Performance in a Global Market," *Annual Review of Economics*, 08 2014, *6* (1), 201–227.

_ , Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik, "Prices, Markups and Trade Reform," NBER Working Papers 17925, National Bureau of Economic Research, Inc March 2012.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, September 2002, *70* (5), 1741–1779.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, "Competition, Markups, and the Gains from International Trade," Economics Series 299, Institute for Advanced Studies July 2013.

García Marin, Alvaro and Nico Voigtländer, "Exporting and Plant-Level Efficiency Gains: Its in the Measure," NBER Working Papers 19033, National Bureau of Economic Research, Inc May 2013.

Impullitti, Giammario and Omar Licandro, "Trade, Firm selection, and innovation: the competition channel," UFAE and IAE Working Papers 841.10, Unitat de Fonaments de l'Anlisi Econmica (UAB) and Institut d'Anlisi Econmica (CSIC) September 2010.

Jamandreu, Jordi and Heng Yin, "Cost and Product Advantages: A Firm-level Model for the Chinese Exports and Industry Growth," Technical Report May 2014.

Kohn, David, Fernando Leibovici, and Michal Szkup, "Financial Frictions and New Exporter Dynamics," January 2012.

Liu, Lili, "Entry-exit, learning, and productivity change Evidence from Chile," *Journal of Development Economics*, 1993, *42* (2), 217 – 242.

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

Melitz, Marc J and Gianmarco IP Ottaviano, "Market size, trade, and productivity," *The review of economic studies*, 2008, *75* (1), 295–316.

Nguyen, Daniel and Georg Schaur, "Cost Linkages Transmit Volatility Across Markets," May 2011.

Pavcnik, Nina, "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants," *The Review of Economic Studies*, 2002, *69* (1), 245–276.

Rho, Young-Woo and Joel Rodrigue, "Firm-Level Investment and Export Dynamics," January 2012.

Rubini, Loris, "Innovation and the Trade Elasticity," *Journal of Monetary Economy*, 2014, *66*, 32–46.

Ruhl, Kim and Jonathan Willis, "New Exporter Dynamics," November 2008.

Soderbery, Anson, "Market Size, Structure, and Access: Trade with Capacity Constraints," October 2011.

Spearot, Alan C., "Firm Heterogeneity, New Investment and Acquisitions," *The Journal of Industrial Economics*, 2012, *60* (1), 1–45.

Vannoorenberghe, G., "Firm-level volatility and exports," *Journal of International Economics*, 2012, *1*.

Appendix

Appendix A Proof of Propositions 1 and 2

A.1 Proof of Proposition 1

The proof proceeds in two steps. It first shows that q_H is increasing in x and q_F is increasing in y. Given this, it finds the threshold $\tilde{x}(y,z)$ as the combination of shocks y and z that generate $q_H = 0$. If $x < \tilde{x}(y,z)$, the firm will not produce for the domestic market. Similarly, it finds the threshold $\tilde{y}(x,z)$ as the combination of shocks x and z that generate $q_F = 0$. If $y < \tilde{y}(x,z)$, the firm will not produce for the export market.

Start with the problem of maximizing profits:

$$\max_{q_H,q_F}(\exp(x) - \eta Q)q_H - \frac{\gamma}{2}q_H^2 + (\exp(y) - \eta Q)q_F - \frac{\gamma}{2}q_F - (q_H + \tau q_F)^\alpha$$

In an interior solution, the first order conditions are:

$$\exp(x) - \eta Q - \gamma q_H = \alpha\left(q_H + \tau q_F\right)^{\alpha-1} \tag{A.1}$$

$$\exp(y) - \eta Q - \gamma q_F = \tau\alpha\left(q_H + \tau q_F\right)^{\alpha-1} \tag{A.2}$$

$$\Rightarrow (\exp(x) - \eta Q - \gamma q_H)\tau = \exp(y) - \eta Q - \gamma q_F \Rightarrow$$

$$q_H = \frac{\exp(x) - \exp(y)/\tau}{\gamma} - \frac{\eta}{\gamma}Q\left(1 - \tau^{-1}\right) + \frac{q_F}{\tau} \tag{A.3}$$

Using equations (A.2) and (A.3), q_F solves

$$\exp(y) - \eta Q - \gamma q_F - \tau\alpha\left(\frac{\exp(x) - \exp(y)/\tau}{\gamma} - \frac{\eta}{\gamma}Q\left(1 - \tau^{-1}\right) + q_H\left(\frac{1}{\tau} + \tau\right)\right)^{\alpha-1} = 0 \tag{A.4}$$

Next, apply the implicit function theorem to equation (A.4) to find $\frac{\partial q_F}{\partial \exp(y)}$. Let

$$F = \exp(y) - \eta Q - \gamma q_F - \tau\alpha\left(\frac{\exp(x) - \exp(y)/\tau}{\gamma} - \frac{\eta}{\gamma}Q\left(1 - \tau^{-1}\right) + q_H\left(\frac{1}{\tau} + \tau\right)\right)^{\alpha-1}$$

The implicit function theorem states

$$\frac{\partial q_F}{\partial \exp(y)} = -\frac{\frac{\partial F}{\partial \exp(y)}}{\frac{\partial F}{\partial q_F}}$$

Thus,

$$\frac{\partial q_F}{\partial \exp(y)} = -\frac{1+\alpha(\alpha-1)/\gamma\left(\frac{\exp(x)-\exp(y)/\tau}{\gamma}-\frac{\eta}{\gamma}Q\left(1-\tau^{-1}\right)+q_H\left(\frac{1}{\tau}+\tau\right)\right)^{\alpha-2}}{-\gamma} =$$
$$1/\gamma+\alpha(\alpha-1)/\gamma^2\left(\frac{\exp(x)-\exp(y)/\tau}{\gamma}-\frac{\eta}{\gamma}Q\left(1-\tau^{-1}\right)+q_H\left(\frac{1}{\tau}+\tau\right)\right)^{\alpha-2} =$$
$$1/\gamma+\alpha(\alpha-1)/\gamma^2(q_H+\tau q_F)^{\alpha-2} > 0$$

The way to prove $\frac{\partial q_H}{\partial \exp(x)}$ is similar, but instead of writing q_H as a function of q_F in equation (A.3), write q_F as a function of q_H and insert this into equation (A.1) to find $\frac{\partial q_H}{\partial \exp(x)} > 0$. In this way, one can also prove $\frac{\partial q_H}{\partial \exp(y)} < 0$ and $\frac{\partial q_F}{\partial \exp(x)} < 0$.

To finish the proof, we just need to show that if $y = \tilde{y}(x,z)$, then $q_F = 0$ and if $x = \tilde{x}(y,z)$, then $q_H = 0$. To do this, consider first the threshold $\tilde{y}(x,z)$. Replacing $q_F = 0$ into equations (A.1) and (A.2),

$$\exp(x) - \eta Q - \gamma q_H = \alpha q_H^{\alpha-1}, \quad \exp(y) - \eta Q = \tau\alpha q_H^{\alpha-1}$$

Thus,

$$\exp(x) - \eta Q - \gamma\left(\frac{\exp(y)-\eta Q}{\tau\alpha}\right)^{\frac{1}{\alpha-1}} = \frac{\exp(y)-\eta Q}{\tau} \tag{A.5}$$

Solving equation (A.5) for y as a function of x and z delivers the threshold $\tilde{y}(x,z)$ in proposition 1. A similar procedure delivers the threshold $\tilde{x}(y,z)$. □

A.2 Proof of Proposition 2

We prove that the solution is unique by showing that, for a given triplet (x,y,z), a firm's decision is unique. Let x_0, y_0, z_0 be such that the firm chooses to sell domestically but not export, that is, $\exp(x_0) > \eta Q, x_0 > \tilde{x}(y_0,z_0)$. Then $y_0 < \tilde{y}(x_0,z_0)$. Similarly, if the firm chooses to export only, that is $\exp(y_0) > \eta Q, y_0 > \tilde{y}(x_0,z_0)$, then $x_0 < \tilde{x}(y_0,z_0)$. The proof shows the first part of the proposition. The second part is straightforward given the first part. Proceed by contradiction, that is, assume that (x_0,y_0,z_0) are such that $x_0 > \tilde{x}(y_0,z_0)$ and $y_0 > \tilde{y}(x_0,z_0)$. The proof shows this leads to a contradiction.

Let $\tilde{x} = \exp(x_0) - \eta Q$ and $\tilde{y} = \exp(y_0) - \eta Q$.

$$x_0 > \tilde{x}(y_0,z_0) \Rightarrow \tilde{x} > \gamma\left(\frac{\tilde{y}}{\tau\alpha\exp(z_0)}\right)^{\frac{1}{\alpha-1}} + \frac{\tilde{y}}{\tau} \tag{A.6}$$

$$y_0 > \tilde{y}(y_0,z_0) \Rightarrow \tilde{y} > \frac{\gamma}{\tau}\left(\frac{\tilde{x}}{\alpha\exp(z_0)}\right)^{\frac{1}{\alpha-1}} + \tilde{x}\tau \tag{A.7}$$

Using equation (A.7) in equation (A.6),

$$\tilde{x} > \gamma^{\frac{\alpha}{\alpha-1}} \tilde{x}^{\frac{1}{\alpha-1}} (\tau\alpha \exp(z_0))^{\frac{-2}{\alpha-1}} + \tau\tilde{x} \Leftrightarrow (1-\tau) > \gamma^{\frac{\alpha}{\alpha-1}} \tilde{x}^{\frac{2-\alpha}{\alpha-1}} (\tau\alpha \exp(z_0))^{\frac{-2}{\alpha-1}}$$

The last line is a contradiction, since the term on the left hand side is negative and the term on the right hand side nonnegative.

Appendix B Extracting Shock Realizations

The following appendix details the cross section calibration. First we identify the shock realizations from the data. Then we use these realizations to estimate the distributions of shocks via maximum likelihood.

The process is as follows. Firms observe shocks x, y, z, unobservable to us, and make production decisions, both for the export and domestic markets, which are available to us. In addition, information on sales plus other information on costs available in the database allows us to estimate markups for each firm, as in De Loecker and Warzynski (2012). The data on domestic sales, exports, and markups allows us to solve a non linear system of three equations and three unknowns that determine the shocks x, y and z.

This process requires information on ηQ and γ. As we argue in the calibration section, η and γ can be normalized, so we fix then equal to 1 and 2, respectively. Q on the other hand is an equilibrium variable. However, another normalization, the mass of firms M, determines the size of Q in equilibrium. Thus, we normalize M so that $Q = 1$.

B.1 Non Exporters

In the case of non exporters, we do not have relevant information on the export demand shock y. Thus, we can only extract the realization of the shocks x and z. We do this using data on markups (m) and sales (r). We first identify total cost c as:

$$m = \frac{r}{c} \Rightarrow c = \frac{r}{m}$$

Recall the first order condition and the price for non exporters,

$$\exp(x) - \eta Q - \gamma q = \beta \exp(z) q^{\beta-1} \tag{B.1}$$

$$\tag{B.2}$$

Multiply equation (B.1) by q to obtain:

$$pq - \frac{\gamma}{2} q^2 = \beta \exp(z) q^{\beta} = \beta c \tag{B.3}$$

Given revenues pq and costs c, we use equation (B.3) to identify q. Using this, we obtain the marginal cost as

$$\beta \frac{c}{q} = \beta \exp(z) q^{\beta-1}$$

Given q, this pins down z. Plugging this into (B.1) pins down the shock realizations x.

B.2 Exporters

In this case we extract all shock realizations x, y, z as follows. Let r_d be the revenues of domestic sales and r_x exports. Multiplying the first order conditions by q_d and q_x delivers

$$r_d - \frac{\gamma}{2} q_d^2 = \beta \exp(z)(q_d + \tau q_x)^{\beta-1} q_d$$
$$r_x - \frac{\gamma}{2} q_x^2 = \beta \exp(z)(q_d + \tau q_x)^{\beta-1} \tau q_x$$

Adding these up

$$r_d + r_x - \frac{\gamma}{2}\left(q_d^2 + q_x^2\right) = \beta \exp(z)(q_d + \tau q_x)^{\beta} = \beta c$$

where $c = \frac{r_d + r_x}{m}$. Rearranging,

$$\tilde{q} = q_d^2 + q_x^2 = \frac{r_d + r_x - \beta c}{\gamma/2}$$

So we have $\left(q_d^2 + q_x^2\right) = \tilde{q}$. We can then find q_d and q_x by solving a system of two equations and two unknowns. The second equation combines the two equations above. The equations are

$$q_d^2 + q_x^2 = \frac{r_d + r_x - \beta c}{\gamma/2}$$
$$\frac{r_d}{q_d} - \frac{\gamma}{2} q_d = \frac{r_x}{\tau q_x} - \frac{\gamma}{2\tau} q_x$$

q_d is therefore the solution to the following non-linear equation:

$$\frac{r_d}{q_d} - \frac{\gamma}{2} q_d = \frac{r_x}{\tau \sqrt{\tilde{q} - q_d^2}} - \frac{\gamma}{2\tau} \sqrt{\tilde{q} - q_d^2}$$

Given these variables, we obtain the marginal cost as

$$c' = \beta \exp(z)(q_d + \tau q_x)^{\beta-1} = \beta \frac{c}{q_d + \tau q_x}$$

Next obtain x, y from

$$\exp(x) - \eta Q - \frac{\gamma}{2} q_d = c'$$
$$\exp(y) - \eta Q - \frac{\gamma}{2} q_x = \tau c'$$

Lastly, obtain $\exp(z)$ from

$$c' = \beta \exp(z)(q_d + \tau q_x)^{\beta-1}$$

Once we have all the data on $x, y, \exp(z)$, we can estimate the parameters in the distributions via Maximum Likelihood. Under the assumption that the processes for the variables are

$$\begin{aligned} x' &= \rho_x x + (1-\rho_x)\mu_x + \epsilon_x, \quad \epsilon_x \sim N(0, \sigma_x^2) \\ y' &= \rho_y y + (1-\rho_y)\mu_y + \epsilon_y, \quad \epsilon_y \sim N(0, \sigma_y^2) \\ \log(\exp(z)') &= \rho_d \log(\exp(z)) + (1-\rho_d)\mu_d + \epsilon_d, \quad \epsilon_d \sim N(0, \sigma_d^2) \end{aligned}$$

the distributions of the cross section in each variable are

$$\begin{aligned} x &\sim N\left(\mu_x, \frac{\sigma_x^2}{1-\rho_x^2}\right) \\ y &\sim N\left(\mu_y, \frac{\sigma_y^2}{1-\rho_y^2}\right) \\ z &\sim N\left(\mu_d, \frac{\sigma_d^2}{1-\rho_d^2}\right) \end{aligned}$$

However, we need to deal with the selection bias. We observe only x such that $\exp(x) \geq \eta Q$ and $\exp(y) \geq \tilde{y}(x, \exp(z))$ where $\tilde{y}(x, \exp(z))$ solves

$$\begin{aligned} \gamma\left(\frac{\tilde{y}(x, \exp(z)) - \eta Q}{\tau \exp(z)\beta}\right)^{\frac{1}{\beta-1}} + \eta Q\left(1-\tau^{-1}\right) + \frac{\tilde{y}(x, \exp(z))}{\tau} - \exp(x) = 0, \\ \tilde{y}(x, \exp(z)) = \max\left\{\eta Q, \tilde{y}(x, \exp(z))\right\} \end{aligned}$$

The densities for the variables x and z are

$$\begin{aligned} f_x(x) &= \frac{normpdf\left(x, \mu_x, \frac{\sigma_x^2}{1-\rho_x^2}\right)}{1 - normcdf\left(\eta Q, \mu_x, \frac{\sigma_x^2}{1-\rho_x^2}\right)} \\ f_d(\exp(z)) &= normpdf\left(\log(\exp(z)), \mu_d, \frac{\sigma_d^2}{1-\rho_d^2}\right) \end{aligned}$$

However, it turns out that the restriction $\exp(x) \geq \eta Q$ hardly binds, so we ignore it. The problem is different in the case of the variable y. In this case, we have a problem of missing data, and it is not missing at random. One option would be to perform a censored Maximum Likelihood Estimation. The problem is that, since most firms are non exporters in the sample, there are too many missing observations, and therefore the estimates are not likely going to be good. Thus, we do not estimate the distribution of y. Instead, we calibrate the relevant parameters μ_y and σ_y so that we match the ratio of total exports to total sales in the economy, and the proportion of firms that export.

Appendix C Effects on Sales and Prices

C.1 Sales

When focusing only on the intensive margin, several features stand out. The biggest change, as expected, is in exports, which increase on average by 47 percent, although the median increase is smaller, at 7 percent. No firm reduces its exports. There is a great degree of heterogeneity in this increase, as we show in Figure 14.[20]

More surprising is the behavior of domestic sales: most firms increase their domestic sales following a drop in trade costs (76 percent). On average, domestic sales increase by 5 percent, although the median increase is only 0.2 percent. Figure 15 shows the distribution of increases in domestic sales by firms that exported before and after the reduction in trade costs.

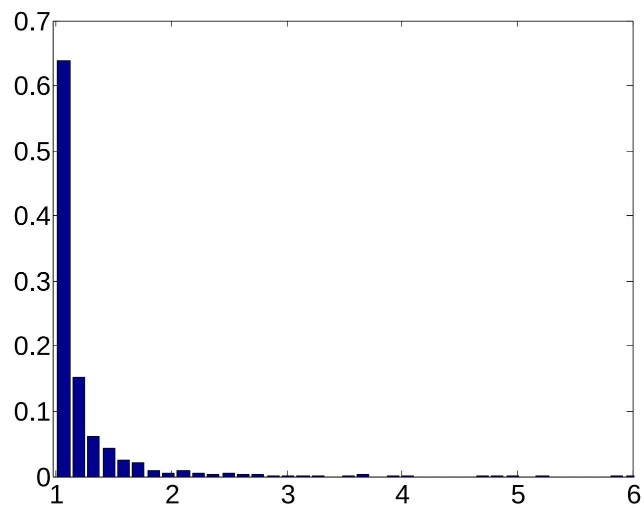


Figure 14: Increase in Exports Along the Intensive Margin

The fact that some firms sell more domestically after a reduction in tariffs is not present in standard international trade models. In our model, the reason is as follows. A reduction in trade costs implies a gain in efficiency and a reduction in costs. While this affects exports more than domestic sales, the decreasing returns to scale technology implies that cost reductions are also present for domestic output. When faced with a reduction in costs, firms tend to increase output. The allocation of this increased output depends on the elasticities of demand across markets. To verify this, we perform two regressions. The first regresses the increase in domestic sales (in logs) on the domestic and foreign elasticities of demand. The second regression does the same, but

[20]For expositional purposes, the figures do not include the top and bottom 1 percent.

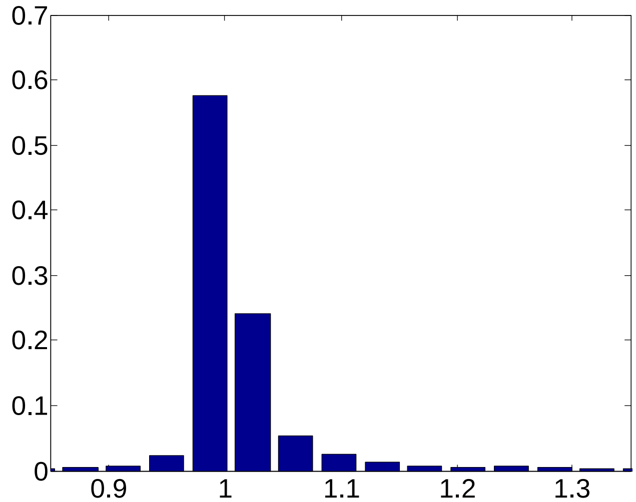


Figure 15: Change in Domestic Sales Along the Intensive Margin

changes the dependent variable to export sales. That is, we regress

$$\log\left(\frac{\text{Domestic sales low }\tau}{\text{Domestic sales high }\tau}\right) = \beta_{0d} + \beta_{1d}|\eta_d| + \beta_{2d}|\eta_x| + \epsilon_d$$
$$\log\left(\frac{\text{Export sales low }\tau}{\text{Export sales high }\tau}\right) = \beta_{0x} + \beta_{1x}|\eta_d| + \beta_{2x}|\eta_x| + \epsilon_x$$

Table VIII reports the results. The estimates are very robust. They show that the elasticity of demand is key to determine the change in exports and domestic sales. Exports increase more when the elasticity of foreign demand is larger, and increase less if domestic demand is elastic.

Parameter	Estimate	99% Confidence Interval	R^2
β_{1x}	-0.0009	$[-0.0017, -0.0001]$	0.5001
β_{2x}	0.2673	$[0.2597, 0.2749]$	
β_{1d}	-0.0048	$[-0.0057, -0.0040]$	0.4613
β_{2d}	-0.0310	$[-0.0389, -0.0231]$	

Table VIII: Elasticities and the Change in Domestic and Foreign Sales

Notice that the key elements for these results are decreasing returns to scale technologies, and heterogeneous elasticities of demand. Models based on Dixit-Stiglitz preferences cannot replicate this, even when paired with decreasing returns to scale technologies. In fact, in Melitz (2003), domestic sales can only be affected via a general equilibrium effect (wages increase after a reduction in trade costs), and unequivocally domestic sales drop in this case.

The logic of decreasing returns to scale is also apparent when considering the impact on domestic sales for firms along the extensive margin. For these firms, the decline in trade costs does not represent an efficiency gain since these firms were not originally selling abroad. Instead, the decline in trade costs encourages these firms to enter the export market, which raises marginal costs, and tends to cause firms to substitute away from the domestic market. As can be seen in Figure 16, about 60 percent of firms reduce domestic sales, with an average reduction of 4.3 percent. The remaining 40 percent of firms hardly change their domestic sales (the maximum change is an increase of 1.2 percent). Again note that with constant marginal cost technology, there would be no impact on domestic sales for these firms.

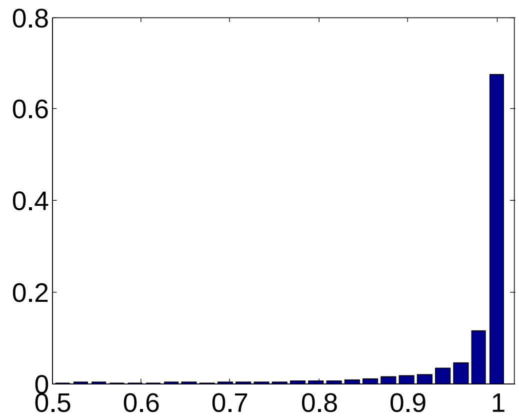


Figure 16: Change in Domestic Sales Along the Intensive Margin

C.2 Prices

Next, we analyze the effect of the reduction in trade costs on prices. The behavior of prices follows closely the behavior of sales, so we do not go into much detail in this section.

We focus first on changes along the intensive margin, that is, firms that were exporting prior to the change in trade costs. As one would expect, export prices drop with lower trade costs. The median drop is 7.9 percent, and the average drop is 7.3 percent. No price increases. Figure 17 shows a histogram with the change in export prices.

The story is somewhat different considering domestic prices. On average, these change very slightly. However, the changes tend to be price drops. The average change is a reduction of 1 percent, and the median a drop of 0.6 percent. Twenty five percent of prices increase. Figure 18 shows a histogram with the change in domestic prices.

The reason why domestic prices can increase or decrease is intuitive. A reduction in trade costs is a reduction in marginal costs. Given decreasing marginal returns, the marginal cost both for domestic and foreign quantities decreases, so domestic prices can go down. However, since trade costs affect exports more, exports increase more, increasing the marginal cost, and potentially increasing the domestic price.

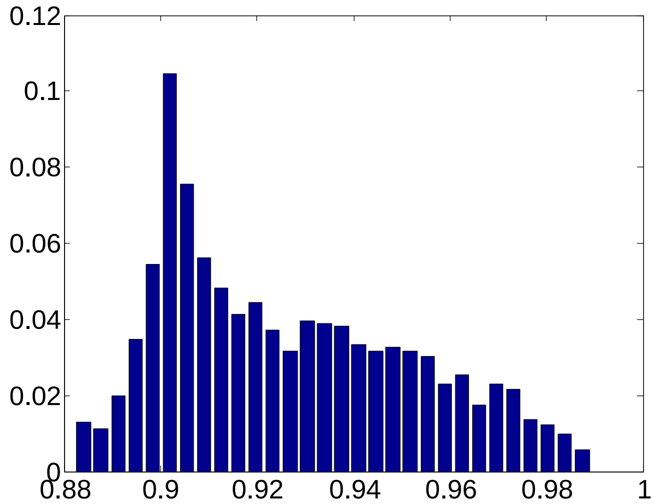


Figure 17: Change in Export Prices Along the Intensive Margin

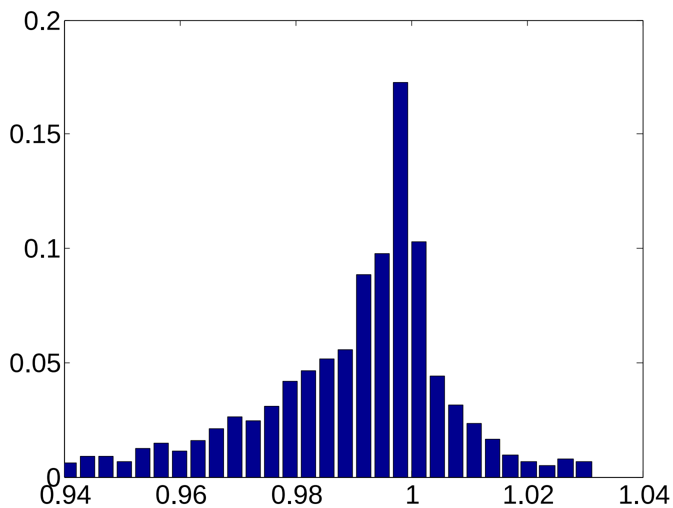


Figure 18: Change in Domestic Prices Along the Intensive Margin