Markups and Misallocation with Evidence from Exchange Rate Shocks*

Ariel Weinberger†
University of Oklahoma

March 2019

Abstract

In a setting with firms that charge variable markups, this paper finds that firm heterogeneity has welfare implications that result exclusively from the differential markup adjustment to global shocks. Changes in allocative efficiency are summarized by a sufficient statistic that can be directly measured with aggregate data. I use Chilean data between 1995-2007 to show that exchange rate shocks can be an important driver of allocative efficiency, as there are large changes in misallocation over time due to the way firms pass-through these shocks into markups. At the firm-level, there is evidence that importing firms pass-through real exchange rate appreciations into their markups, and over time industries that import a larger share of their inputs become more misallocated. In a structural model with productivity gains from importing, where firm market power increases with size, I show how firm reallocation in response to a positive supply shock rationalizes the reduced-form results.

JEL Classification: F12, F14, F43, O47, L11

Keywords: variable mark-ups, non-homothetic preferences, exchange rate volatility, misallocation, international trade.

---

*This work has benefited in numerous ways from discussions and the advice of Robert Feenstra, Katheryn Russ and Ina Simonovska. This paper was previously circulated with the title "Markups and Misallocation with Trade and Heterogeneous Firms." I also appreciate comments by Deborah Swenson, Colin Carter, Paul Bergin, Lucas Herrenbrueck, John Morrow, Swati Dhingra, Luciana Juvenal and other participants at UC Davis seminars, The Midwest International Trade Conference at IUPUI, Rocky Mountain Empirical Trade Conference 2015, and PhD Conference at University of Leicester. I thank Instituto Nacional de Estatdisticas (INE), Chile for graciously providing data.

†Address: Department of Economics, University of Oklahoma. 338 Cate Center Drive, Norman, OK 73019. E-mail: Ariel.Weinberger@ou.edu

1 Introduction

The important conclusion of "new-new" trade theory that entry into the global market raises the average productivity of surviving firms was spearheaded by the breakthrough of the Melitz (2003) heterogeneous firm model, which introduced reallocation as an integral component of the gains from trade. In this model, the nature of the market share reallocation is simplified by using Constant Elasticity of Substitution (CES) preferences (Dixit and Stiglitz, 1977), which results in market outcomes identical to the social optimum.[1] Meanwhile, another literature, the one on growth and productivity, has studied within-industry allocative inefficiency, or the possibility to alter the allocation of production such that real income increases. This research finds that misallocation is an important reason for cross-country income differences (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2008).[2] Motivated by this literature, in this paper I incorporate a possible non-optimal market share reallocation to the Melitz model. I use as a starting point the result of Dhingra and Morrow (2019) (DM) that non- constant markups in a monopolistic competition framework imply a sub-optimal allocation across firms. Given this starting point I find a novel measure, or sufficient statistic, for changes in allocative efficiency, which can be directly measured with data. I connect this measure to the experience of Chile, a commodity exporter with a volatile real exchange rate. I show that real exchange rate shocks can be important drivers of allocative efficiency.

The Melitz model is allocatively efficient due to the CES feature of constant market power. However, when the demand side in the Melitz model is generalized to allow for less restrictive preferences, the market equilibrium is not necessarily efficient, as differences in market power allow for firms to over/under- produce relative to the socially optimal case, with a clear mechanism for a more efficient resource allocation. Therefore, I follow the variable elasticity (VES) framework laid out in DM, where preferences are directly additive as in CES (Dixit and Stiglitz, 1977), but the sub-utility function is allowed to be non-homothetic. It allows for variable markups and implies a production allocation that does not equalize relative marginal utilities with relative production costs, as market power allows highly productive firms to only partially pass-through cost advantages. Higher markups by these firms map onto lower aggregate income relative to the allocatively efficient benchmark and thus creates an aggregate distortion.

In this setting, it is intuitive that a shock to the relative price of domestic versus for-

[1] Feenstra and Kee (2008) showed this to be the case in a setting with firm heterogeneity.

[2] Additionally, Basu and Fernald (2002) (BF) expand on Solow productivity gains – akin to shifting out a country's production possibility frontier – to include welfare-improving movements along this frontier that can be measured using real income.

eign goods will result in a differential adjustment across firms, and thus a reallocation that either increases or decreases this aggregate income. Prior to the formal model, I start by showing motivational evidence for Chile. Chile experienced large real exchange rate shocks throughout the period of 1995-2007, as exemplified by a sharp terms of trade increase in the mid-2000's as a result of a commodity boom. During that same period, I show that importers raise their markup relative to non-importers in response to an appreciation, evidence of lower costs passed through to markups. In conjunction, markup heterogeneity increases significantly, and aggregate income growth is slower than would be expected given the growth in production. This is consistent with the findings in Berman et al. (2012) and DeLoecker et al. (2016) which find that cost shocks are passed on to prices at different rates across firms. I then turn to a theoretical model that connects misallocation at the industry level to firm markup adjustments.

The first theoretical contribution of this paper is to show that the aggregate distortion is captured by the difference between growth rates of aggregate value added and physical production. These measures coincide in a constant markup environment, so I link a deviation in the two measures to the growth rate of allocative efficiency and show that this distortion is present in the welfare decomposition of a representative consumer. This sufficient statistic, which is new to the literature and can be calculated with widely available data, captures the counterfactual deviation in value added relative to a model that ignores the heterogeneity in market power.

Given that this new measure holds for a commonly used class of models in the literature, a second important contribution is then to identify the mechanisms that underlie it. I connect allocative efficiency to changes in the price of foreign goods by identifying the heterogeneous effects of aggregate shocks on firm market power. The shock to the relative price of domestic versus foreign goods has two separate – and possibly simultaneous – effects. One possibility is a competition effect, which works through firm-level demand elasticities. A separate possibility is to change the marginal costs of domestic firms through cheaper intermediate inputs. A real exchange rate appreciation lowers the price of foreign inputs, which allows firms to charge higher markups due to incomplete pass-through into prices. Heterogeneity in the degree of pass-through results in resources being reallocated relatively to low-markup firms, a *compositional* effect that lowers the weighted-average markup and allocative efficiency.[3] I focus on the latter shock in order to explain the observations in the Chilean economy in the mid-2000's.[4]

[3]This connection between the weighted average of markups and misallocation fits in with recent work (Macedoni and Weinberger, 2019; Peters, 2018; Edmond et al., 2018; Baqaee and Farhi, 2017).

[4]Chile also experiences a depreciation between 1997-2003. My results are consistent with this episode,

Resolving the connection between markup heterogeneity and global shocks must tie together theory with data, which this paper does in two parts. First, I provide empirical evidence that changes in aggregate misallocation, measured at the industry level, are due to real exchange rate shocks. The interpretation is that the exposure to exchange rate movements depends on the reliance on imported inputs relative to exports. The main aggregate empirical results are that industries dominated by firms that rely on imported inputs become more misallocated in response to an appreciation in the real effective exchange rate (REER). Comparing the extreme case of an industry that imports 100% of inputs and does not export any of their products with an industry whose share of imports in inputs equals the share of exports to sales[5], a 1 percent increase in the growth rate of the REER leads to about a 2 percentage points smaller growth rate in allocative efficiency in the former industry relative to the latter.

Second, I clarify the mechanisms that drive the reduced form aggregate results by estimating a structural model where a higher share of imported inputs raises productivity and firm market power increases with size. The main innovation is to combine VES preferences with importing on the supply side. The sourcing decision of the firm is based on the sufficient statistic result of Blaum et al. (2018), which connects firm size and markups with the share of inputs that are sourced abroad. I use the Chilean firm data to discipline the parameters of the model, and then investigate a shock to the price of imported inputs relative to domestic inputs. Notice that the distortion I measure in the data is at the industry level, but it is the product of aggregating firm-level responses. In the model, an exchange rate appreciation allows importers to source more from abroad, lower their unit cost, and raise their markups. However, the effects are heterogeneous across firms as their import share and markups depend on their intrinsic productivity. This generates a reallocation of inputs to firms with initially lower markups, and misallocation increases. A quantitative analysis predicts that an industry with an average imported input share as observed in the whole economy would experience a 2.83% reduction in allocative efficiency in response to a 10% appreciation. Therefore, the model allows me to connect the observed movement in the aggregate measures with the simulated firm reallocation that is transparent in the model.

---

however the terms of trade shock in the latter period is sharper and provides a more convincing episode in which to investigate cause and effect. The model in this paper could also be easily applied to a rise in competition due to trade liberalization.

[5]In Section 2.3 I define this "net exposure" (Ekholm et al., 2012) measure.

Related Literature The theoretical and empirical contributions should be viewed relative to numerous recent papers. Dhingra and Morrow (2019) characterize qualitative properties of this misallocation and investigate the case where market size increases. Arkolakis et al. (2019) similarly find that this distortion affects the welfare gains from reducing domestic import tariffs. In relation to these two papers, my contribution is to produce a more direct quantitative measure and include shocks to both the input and output markets that motivate time series variation in allocative efficiency. Although my model follows the VES framework of the former, and the directly additive case of the latter paper, I stress that neither paper — nor any other I am aware of — has identified an empirical estimate that maps to the allocative efficiency measure that these papers identify theoretically. Furthermore, I bring in firm-level data and show how real exchange rate shocks in Chile had distortionary effects.

Edmond et al. (2015) also quantify misallocation in the context of a trade model and measure the welfare gains due to trade liberalization. Their framework imposes nested CES preferences so misallocation is due to supply-side frictions and reduces aggregate TFP. In contrast, I suggest a measure separate from firm TFP in Section 3.2 that is based on the change in the covariance of markups and labor input, an interesting statistic that has not been explored in this literature.[6] This feature also differentiates my measure to that in Holmes et al. (2014) which holds only for homothetic preferences.[7] I view my paper as a complement to that study as I also separate allocative efficiency from productive efficiency, albeit in a different setting that translates nicely to available firm balance sheet data. Importantly, I expand on the limited focus of competition on the output side by adding input side effects (the two papers above concentrate on output tariffs only) in order to apply the predictions to a relevant empirical application.

On the empirics side, my findings are consistent with studies on competition, variable markups, and pass-through, but provide aggregate implications that have not been discussed in this context. Liberalization studies find that tougher competition forces firms to lower prices and raises average productivity, and that pass-through of costs to prices is below one (DeLoecker et al. (2016)). Relatedly, Amiti et al. (2014) find that the most productive firms import the most and also have the lowest pass-through. This is consis-

[6]In future work, this covariance measure could bridge the misallocation measures between models of a supply-side focus and those that rely on non-homothetic demand for variable elasticities. For example, Arkolakis et al. (2019) relate the monopolistic competition distortion to changes in the revenue shares for markup aggregates. Intuitively, the covariance (and allocative efficiency) increases when production is reallocated towards varieties with a high marginal utility (which is proportional to price), and this is possible up to the point that markups are equalized across firms.

[7]I maintain the monopolistic competition environment as in Krugman (1979) and Melitz (2003) but generalize the prefereces within the boundaries of separable utility and log-concave demand.

tent with the terms of trade shock in Chile raising total production but also increasing the degree of misallocation because productive firms raise their markups the most. My empirical results are related to Berman et al. (2012) who show that exchange rate movements tend to affect markups and not export volumes.[8] I focus on how this markup effect relates to reallocation. My structural model with VES preferences and heterogeneous firms that lower their unit cost through imports clarifies the way that incomplete pass-through has interesting aggregate welfare implications.[9]

Theoretical trade models have explored variable markups to generalize welfare gains from trade, though the earlier literature concentrates on the decrease in the average markup in search of a "pro-competitive" effect as in Krugman (1979). With free entry, competition decreases average markups and increases aggregate productivity as firms increase their scale and move down their average cost curves.[10] The distortion present in this paper is the result of the interaction of non-homothetic demand (with separable preferences) and firm heterogeneity, thus capturing separate welfare implications. Within the directly additive non-homothetic preference framework, Behrens et al. (2014) have investigated the pro-competitive effect of trade liberalization, while Simonovska (2015) and Jung et al. (2019) investigate the effect of price discrimination across destinations. Cavallari and Etro (2017) have applied non-homothetic preferences to the study of business cycles using a functional form introduced in Bertoletti et al. (2008). Bertoletti et al. (2018) examine pro-competitive effects when preferences are non-homothetic but not additively separable. Demidova (2017) introduces trade policy to the qualitative statements about gains from trade.[11]

The relation of markups and misallocation has also been a focus outside of the trade literature. In Hsieh and Klenow (2009), firms face output and capital distortions. Although markups are constant due to CES preferences, distortions mean that firms opti-

[8]Chatterjee et al. (2013) is another example of an exchange rate shock that affects markup that is consistent with my empirical findings. Though my firm level results are not new at this point, they provide further evidence for a growing field and do offer a novel connection to the measurement of reallocation.

[9]Pavcnik (2002) and Bartelsman et al. (2013) also attempt to measure productivity growth through reallocation in developing countries, though they focus on different sufficient statistics. In a separate strand of the literature, Amiti and Konings (2007), Kasahara and Lapham (2013) and Goldberg et al. (2010) show how a significant part of the productivity gains are a result of cheaper and more abundant intermediate inputs. I show that this was true also for Chilean firms.

[10](Chamberlin, 1933; Vives, 2001) explore variety and scale trade-offs, but not necessarily misallocation of quantity among existing producers, which requires firm heterogeneity. Feenstra and Weinstein (2017) use a homothetic framework with firm heterogeneity to measure the pro-competitive plus variety effects from increased global competition.

[11]As a note, non-homothetic preferences are not *necessary* for variable markups to introduce distortions with monopolistic competition, as shown in Macedoni and Weinberger (2019) and Behrens et al. (2016).

mally choose non-equal marginal products even though they face identical factor prices. My paper establishes a new way to observe deviations from allocative efficiency, but consistent with the aggregate productivity literature, a distortion inflicts a wedge between total revenues and total output. This Aggregate Productivity Growth (APG) literature decomposes APG into growth in average firm productivity and reallocation. In Basu and Fernald (2002), Petrin and Levinsohn (2012) and Baqaee and Farhi (2017), reallocation increases welfare if inputs are reallocated towards the high markup firms. I incorporate the same type of welfare gain into a trade model that is an extension of Melitz (2003).

The rest of this paper is organized as follows. Section 2 introduces the Chilean data and provides a time series motivation for how the real exchange rate affects markup dispersion and aggregate productivity growth. This motivates the theoretical framework in Section 3, which differentiates between growth in real income in the CES and VES models. Section 4 presents empirical evidence for changes in misallocation in response to exchange rate variation, and clarifies the mechanisms with a model that connects firm imports, unit costs, and markups. Finally, Section 5 concludes and discusses the composition of importers and exporters at the country level in relation to misallocation.

2 Data and Background

2.1 Chile's Open Economy Shocks

The focus of this study is to examine changes in allocative efficiency that reflect micro-level reallocation as a response to a shock in the relative price of domestic versus foreign goods. In this subsection I summarize time series facts about the real exchange rate and output tariffs in Chile, which can be mapped to this type of shock. Chile's trade liberalization continues a trend that starts in the late 1970's, but the real exchange rate provides a larger and more unexpected shock during data range I explore.

Macro and open economy data is taken from a variety of sources. The Central Bank of Chile provides macroeconomic measures that include the open economy. Detailed export and import data at the 4-digit level is provided in the world trade flows database of Feenstra et al. (2005). The real effective exchange rate (REER) is available from the FRED (St. Louis Federal Reserve) and IFS (IMF) databases.[12] Terms of trade plus alternative import and export data can be obtained from World Development Indicators (WDI) at

[12]It is a geometric average of relative prices using trade weights. I have also checked that these data can be computed using trade weights from BIS (http://www.bis.org/statistics/eer/) and output prices provided by the Penn World Tables (PWT) 8.0.

the World Bank. The World Integrated Trade Solutions (WITS) database has detailed tariff data that I aggregate to the 4-digit level. [13]

Figure 6 (Appendix) plots manufacturing exports and imports as a ratio of total manufacturing value added.[14] Exports rise steadily throughout this period, with an acceleration after 2003, while imports boom after 2003. Any story explaining Chilean trade must include its large exposure to commodities. This is especially important in the copper industry, which constitutes almost half of its export value in the mid-2000s, after the price of copper *triples* between 2003 and 2007. Therefore, the export surge can be explained as demand driven as Chile gained from the inflation in commodity prices that was due to the increased demand from emerging countries. The large increase in imports post-2003 is driven especially by intermediate inputs.[15] For the manufacturing firms that I consider, importing is as important as the export side, as firms rely on imported inputs of machinery and capital goods.[16] Given the time series of the aggregate trade data, a reasonable story is that export earnings are the initial impetus, with the demand for intermediate inputs driving imports.

The demand shock for Chile's commodities is reflected in the large terms of trade gain starting in 2003. Figure 1 describes the terms of trade (TOT) along with the REER (both normalized to 100). Chile experienced an appreciation before 1997, a sustained depreciation from 1997-2003, and a sustained appreciation 2003-2006 led by the terms of trade gain. The terms of trade is mostly unaffected by the depreciation, but follows a similar, though magnified, trend in the 2003-2006 period. Although both measures imply a large appreciation post-2003 (with TOT being more extreme), the REER allows us to capture more volatility due to the large depreciation between 1997-2003.

The shock to the terms of trade, reflected by the real exchange rate appreciation, is likely exogenous to the non-copper industries, as it is driven by the copper boom. In terms of the 1997-2003 depreciation, there is no obvious cause, but it occurs in conjunction with a global weakness in developing markets. Still, a shock in any direction affects the cost of imported inputs and/or the output prices relative to foreign competitors. In the empirical

[13] A previous version of this paper showed that input tariffs are almost identical during this period, as there is a mostly uniform reduction in industry tariffs.

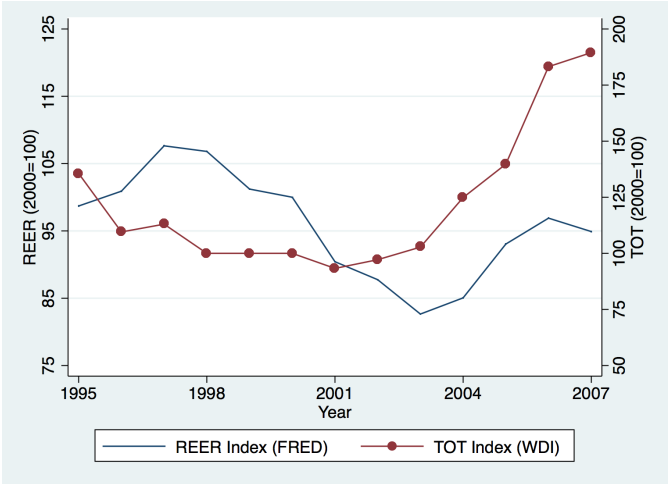
[14] Exports and imports are gross flows (so they can be greater than total manufacturing value added).

[15] Desormeaux et al. (2010) establishes that firms and households import a significant amount of their intermediary inputs. In Weinberger (2015), customs data is used to report large import growth for Chile in *intermediate inputs* during the same time period. An important component of the growth is in new varieties of inputs being imported (the "sub-extensive" margin in Gopinath and Neiman (2014)).

[16] Berthelon (2011) documents that Chilean export performance from $1990 - 2007$, even taking out copper industries, shows growth in the extensive margin and diversification of products as well as partners.

analysis I use the REER as the benchmark measure and provide robustness results with the TOT measure. The REER is a better gauge of relative prices, but the appreciation of post-2003 arguably provides a cleaner identification of a shock to non-copper industries. For that reason I also investigate changes in markups using sub-periods of the data range.

Figure 1: Terms of Trade (2000=100) and Real Effective Exchange Rate (2000=100), 1995-2007



Sources: WDI Indicators, International Financial Statistics (IMF). TOT is an index from WDI. REER, taken from IFS data, is calculated using Penn World Tables to calculate Chile's production price index relative to its top trade partners and taking a geometric average using trade shares (from BIS) as weights. I report the indices with the year 2000 as base year.

The open economy is also affected by changes in the trade regime. The time period examined in this paper is subsequent to the big trade reform in Chile that occurred in the late 1970's (and studied in Pavcnik (2002)). Figure 7 (Appendix) shows the average applied tariff rate from the Comtrade database. In the time span of the data, average applied tariffs in the manufacturing sector decrease from 11% to below 2%. However, this drop is mostly homogeneous across industries and most likely expected by firms as Chile must gradually lower its output tariffs as a condition to joining the WTO in 1995. Its effect on relative prices of foreign and domestic goods is also likely mitigated by concurrent drop in export tariffs. For these reasons, in this paper I explore only the effect of real exchange rate shocks on allocative efficiency (and firm markups).[17]

[17]The Appendix provides theoretical effects of greater competition on markup heterogeneity, but I focus on input cost shocks in the main paper in order to have a more unified story. Relatedly, in the empirical analysis, I control for the level of tariffs.

2.2 Firm and Industry Data

To produce a measure of misallocation and relate it to firm performance requires production data for the universe of Chilean firms. I combine a panel of the Chilean census of firms from 1995-2007 with aggregate statistics from this same period. The firm level data is provided by Encuesta Nacional Industrial Anual (ENIA, National Industrial Survey) and collected by the National Institute of Statistics (INE). It covers a census of manufacturing firms, ISIC (rev. 3) classification 15-37, with more than 10 workers. There are approximately 5,000 firm level observations per year and firms are tracked across time with a unique identification number. Each firm provides detailed economic data such as total sales, number of workers, value of fixed capital, expenditures on intermediate inputs, etc. This data will be used to construct estimates of firm productivity and markups using methods described below. Importantly, firms also report the value of inputs that are imported from abroad and what value of their total sales is exported.

Aggregate measures are motivated by theoretical aggregates in Section 3, but also have counterparts in the literature. The goal is to track aggregate real income growth, and compare it to a measure of physical production. I will make use of industry-level data to construct growth rates of revenue and production of physical quantity. The growth in real revenue is the Aggregate Productivity Growth (APG) measure used by Basu and Fernald (2002) (BF) and Petrin and Levinsohn (2012) (PL) defined by total growth in (deflated) value added within an industry, and corrected for the growth in labor.[18]. For a measure of the aggregate price level I must use the INE's 4-digit ISIC industry deflators.[19] Furthermore, I use a real production index provided by the same agency that conducts the annual firm census. This survey tracks only a subset of the census of firms, but gets data on physical production (divorced from prices). This is used to produce an index of production at the 3-digit ISIC level that allows me to track annual growth in physical production by incumbent firms.[20] This index follows a subset of firms with bases in 1989 (for the 1995-2002 data) and 2002 (used for the 2003-2007 data). I aggregate the quantity index to the 2-digit level and conduct the industry analysis at this level.

Next, I use the firm data to provide evidence on the effect of real exchange rate vari-

[18]By the national revenue accounting identity, the sum of value added is equal to the sum of final demand in an industry. See Appendix A.2 for details on the construction of this measure. As in PL, I correct for total industry wage growth since the theory will assume there is no reallocation of labor across sectors.

[19]These deflators are computed by the INE, which I take as a reasonable approximation of my aggregate price level. The aggregate in my model is a weighted average of individual prices. The INE constructs a Laspeyres index that is aggregated using 7-digit products.

[20]Again, I do not have the micro-level quantities to construct an index identical to my model, but this index is a reasonable approximation.

ation on firm markups and data time series evidence of markup heterogeneity. I also introduce aggregate performance measures of the Chilean economy which motivate the theoretical framework in the next section. I will return to the firm and industry data above in Section 4.1, which identifies real exchange rate variation as an impetus for changes in allocative efficiency.

2.3 Markup Estimation and Firm-Level Response to Exchange Rate Shock

The reallocation story that is the mechanism behind changes in misallocation requires heterogeneous markup adjustment across firms. At the firm-level, markups can be estimated using the method from DeLoecker and Warzynski (2012). The first step is to calculate production function coefficients ala Ackerberg et al. (2015) (ACF) – in itself an extension of the seminal contributions of Olley and Pakes (1996) and Levinsohn and Petrin (2003) (OP and LP) – and then to use these coefficients plus cost shares to estimate firm-level markups. The details on production function estimation and translating this to markups is relegated to Appendix A.3 as this method has been used extensively in the trade literature. Table 4 (in the Appendix) shows the production function coefficients and the median markup in all industries. The median markup across the manufacturing sector as a whole is consistent with past estimates, at 25%.

Firm-Level Response to REER Shocks First, I show that firms in fact respond to changes in the relative price of foreign and domestic goods by adjusting their markup. I measure the differential firm markup responses to exchange rate volatility in a difference-in-difference specification in which groups are treated based on their exposure to the real exchange rate shock. For example, when the cost of imported inputs is lower, importers should raise their markup relative to non-importers.[21]

The Chilean census data has information on the value of exports and imports for each firm. I interact this information with shock to the relative price of imported relative to domestic goods. I run the following specification:

$$\mu_{it} = \alpha_i + \alpha_t + \chi REER_t * Expos_i + REER_t * Z_{it} + e_{it}. \tag{1}$$

I include both firm (α_i) and year (α_t) fixed effects, though in the appendix I also show

[21]The differential adjustment *within importers* will be present in the *aggregate* measures as it implies a compositional change in the allocation of production. In this section I merely confirm that importers adjust their markups in response to the shock.

results with industry (4 digit ISIC)-year interacted fixed effects.[22] The outcome measure is the firm markup, and the coefficient of interest is χ, the differential effect across firm exposure to the exchange rate shock. The last terms include firm time-varying controls interacted with the foreign shock, to control for the fact that relative markups might be affected by changes due to the foreign shock.

$Expos_i$ will take the form of the share of exports in total sales/share relative to imports in intermediate inputs. Following Ekholm et al. (2012), a "Net Exposure" variable is constructed as the difference between export share and import share for a firm. They model firm revenues and costs in order to compute the elasticity of each with respect to the real exchange rate the firm faces. In this partial equilibrium approach, the firms' export share is equal the elasticity of revenues with respect to the real exchange rate and the share of imports in total costs is the elasticity of costs with respect to the real exchange rate. Then the net exposure, the difference between the export share and share of imported inputs, directly affects the elasticity of profits (and therefore markups) with respect to the real exchange rate.[23] The specific derivations used in Ekholm et al. (2012) are provided in Appendix A.5. I fix the exposure to shocks over time to not allow for an endogenous change in exposure.

Regression Results The firm level regressions affirm the predicted markup responses. To start, Appendix Table 8 establishes that the assumption that more productive firms also have higher markups — shown in other papers such as DeLoecker and Warzynski (2012) — holds for this data as well. Capital intensity, imported input share, employment and being a multinational are all positively associated with firm markups. This holds in the case that I measure markups using the DeLoecker and Warzynski (2012) method (I label it "Lerner" because I transform them to a Lerner index), and also 2 alternative markups: a profits to sales ratio ("Profit"), and an inverse labor share measure "Lshare".[24] In all regressions, I control for time-varying TFP, as well as time-varying firm characteristics. Due to space limitations, I omit the latter characteristics from the main Tables (though

[22]In this case time-varying industry shocks are controlled for by the fixed effects, as well as any time-invariant firm characteristic. Furthermore, the main results are almost identical if one restricts the sample to post-2002 *or* pre-2001, which suggest that markups respond to both depreciations and appreciations.

[23]Since equal import and export shares don't necessarily cancel each other out, I also run all regressions with import and export shares as separate regressors.

[24]I construct markups three ways. (1)"Markup (Lerner)" = μ_{it}, where $\frac{1}{1-\mu}$ is the price-cost ratio, or ratio of material output elasticity and material cost share ($\frac{\theta^m}{\alpha^m}$) estimated using (DeLoecker and Warzynski, 2012); (2) "Markup (Profit)" = $\frac{Sales_{it}-wages_{it}-capitalcosts_{it}-inputscosts_{it}}{Sales_{it}}$, constructed using sales and input data; (3) "Markup (Lshare)" = μ_{it}^{Lshare}, where μ^{Lshare} is a Lerner index from the inverse labor share of value added.

they are displayed in the Appendix tables).

Table 1 uses import/export shares, as well as the net exposure, to measure the degree to which firm are exposed to exchange rate variations. I expect firms with negative exposure to be able to raise their markups in response to lower costs. The negative coefficient on the interaction between the real exchange rate and net exposure in the first column means that a real effective exchange rate (REER) increases markups for firms that have negative exposure (input importers) relative to firms with no (or exporting) exposure. A firm that imports all of its inputs and does not export raises its price-cost margin by 8.7 percentage points more than a firm that neither exports nor imports in response to a 1% appreciation. In column (2), it is evident that the higher markups for relatively more importing-intensive firms is driven by the import share. In column (3), I add the industry output tariff interacted with firm exposure, which makes the interaction term on the REER more negative. The positive coefficient on tariffs is also expected, as it implies lower tariffs will result in lower markups for exporters. The Appendix includes results with industry-year interacted, as well as region-year interacted, fixed effects. I also run separate analysis for 1995-2001 and 2002-2007, and the response to the REER holds in both periods.

Finally, the last two columns repeat the specification in the first column, but I replace the main markup measure with two alternatives: a ratio of sales minus costs to total sales ($Profit$) and an inverse of the labor share ($Lshare$). The advantage of these measures is that they don't require an estimation of the production function, which assumes firms face the same material prices.[25] The relationship still holds as in the baseline. The response of the $Lshare$ measure is less precisely estimated, but it likely has the largest measurement error as it implies an average markup of 61%.

Markup Dispersion The majority of the literature on variable markups has focused on average markups due to a "pro-competitive" effect (Feenstra and Weinstein, 2017). That focus does not include possible allocative inefficiencies, which requires that markups vary across firms (Basu and Fernald, 2002; Petrin and Levinsohn, 2012). In a setting that is similar to this paper, Berman et al. (2012) show that French exporters react to real exchange rate shocks differentially depending on their productivity. Next, I show the time series of the dispersion of markups, to motivate the heterogeneity present in the way firms adjust their markups in Chile. Unlike Berman et al. (2012), I will investigate the *compositional* ef-

[25] In running the DeLoecker and Warzynski (2012) procedure, I control for import status (Kasahara and Rodrigue, 2008). I also add TFP as a control in the Columns (1)-(3), since it is estimated jointly in the procedure.

Table 1: Firm Level: Differential Effect on Markup: Fixed Exposure Shares

	Markup (Lerner)			Markup (Profit)	Markup (Lshare)
	(1)	(2)	(3)	(4)	(5)
REER*NetExposure	-0.087***		-0.108***	-0.128**	-0.061
	(0.030)		(0.026)	(0.062)	(0.050)
REER*ImportShare		0.174***			
		(0.039)			
REER*ExportShare		0.058			
		(0.044)			
Tariff*NetExposure			0.007		
			(0.004)		
TFP	0.111***	0.111***	0.111***		
	(0.004)	(0.004)	(0.004)		
Fixed Effects	Firm, Year	Firm, Year	Firm, Year	Firm, Year	Firm, Year
R^2	0.77	0.77	0.77	0.25	0.59
N	32212	32212	32212	32212	32212
Avg Markup	0.23	0.23	0.23	0.39	0.61

This table examines the differential markup responses to foreign shocks depending on firm exposure. "Net Exposure" is the difference between the share of sales that are exported and the share of inputs that are imported. Since it is fixed over time for each firm, it is dropped from the specification. Dependent variable for the first 3 columns is the Lerner index, which the price-cost ratio measured using the procedure outlined in DeLoecker and Warzynski (2012) (DLW). TFP measurement also follows DLW. REER, TOT, and output tariffs are in logs. Column (4) uses a profit share measure of the markup: "Markup (Profit)" = $\frac{Sales_{it}-wages_{it}-capitalcosts_{it}-inputscosts_{it}}{Sales_{it}}$. The last column uses a Lerner index of the inverse labor share: "Markup (Lshare)" = μ^{Lshare}. The last two columns *do not* include TFP as a control as they are not constructed using the productivity estimation procedures. All columns include firm and year fixed effects (for industry-year interacted FEs see Appendix). I interact the following firm characteristics with the foreign shock to use as controls: capital intensity, a dummy if the firm is a multinational, the ratio of skilled to unskilled labor. The table only displays the results for the REER interaction. Standard errors are clustered at the firm level. I drop the basic metal industry (ISIC 27). $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

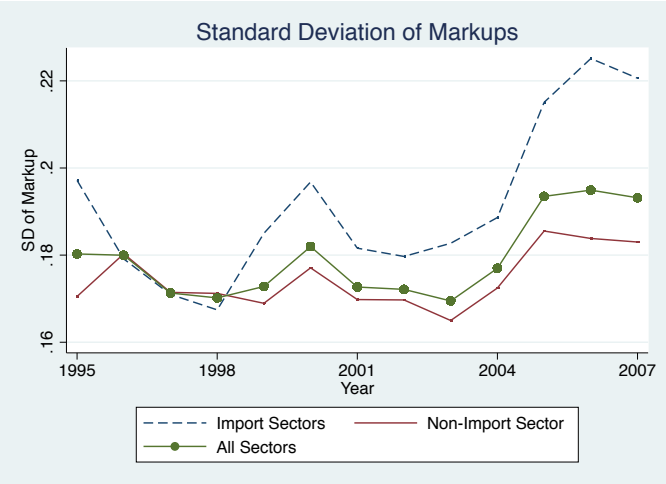
fects of the markup adjustment by measuring aggregate misallocation in response to the real exchange rate shock in Section 4.1. To preview those results, I report the dispersion of markups over time in Chile.

In Figure 2 the dispersion is calculated within each 2-digit sector and averaged (excluding basic metals) to exhibit the manufacturing industry as a whole.[26] There are two main results. First, the middle line shows that there is an increase in markup dispersion in the 2003-2006 period, which is concurrent with the large exchange rate appreciation. Before then, the markup dispersion decreases, which is consistent with the depreciation. The second result is that the increase in dispersion of 2003 is driven entirely by importing industries. The rise in dispersion in importing industries is consistent with the story I build in this paper, that exchange rate movements will reallocate production in a way that a constant markup environment misses.[27]

[26] I use the standard deviation of log markups as my measure, though the results (in terms of dispersion) would be qualitatively similar using the Pareto shape parameter. To estimate the Pareto shape parameter, I use the procedure outlined in Head et al. (2014).

[27] Results for the average across all industries are very similar if I disregard industry classification and conduct the analysis on manufacturing as a whole. This relieves a concern about the reallocation across

Figure 2: Markup Dispersion: Average Across Importing, Non-Importing, and All Sectors



Markup dispersion calculated for each sector by estimating the standard deviation (other methods available upon request). The connected line is the average across all sectors. The solid line takes the average standard deviation for sectors where less than 25% of firms are importers, while the dash line does the same for sector where more than 25% of firms are importers (half of sectors). I eliminate firms in the bottom and top 2.5% of the markup distribution and also the whole Basic Metal industry (ISIC sector 27).

2.4 Measures of the Aggregate Economy

Next, I introduce the implications of the underlying markup heterogeneity by asking: during the same time period, what was the performance of Chile's manufacturing sector? I will show that the time series of real income growth does not mirror that of physical production. Instead, real income growth was much smaller in the 2003-2007 period, and higher from 1997-2001, than would be expected given the rise in production. I compare two industry-level (2-digit ISIC) measures: the aggregate productivity growth (APG) defined in Petrin and Levinsohn (2012) and the growth of total physical production as computed by the ENIA survey of firms (both measures defined above). The next section will provide a formal argument as to why the divergence in these measures is possible and how it reflects allocative efficiency. For now, I point out that physical productivity and revenue productivity must grow at the same rates absent firm-specific distortions that drive a wedge between revenue and output.

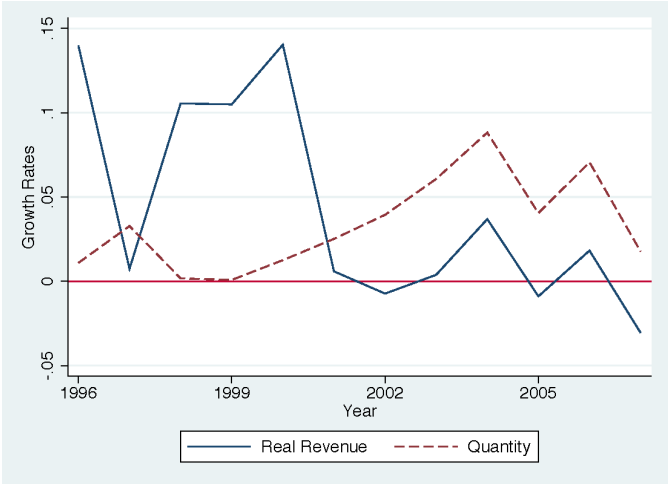
Figure 3 shows real revenue growth and physical production growth at the aggregate manufacturing level, measures defined in the data description above.[28] A complication

industries, ignored in this analysis.

[28]Each is calculated at the 2-digit sector level and I aggregate to the manufacturing level using value added shares by sector. The results above allow for value added weights to change, but I have also used constant shares to eliminate across sector reallocation effects. The growth rates look almost identical, which reflects the fact that there is very little across sector reallocation.

is that the physical production index does not necessarily include all producing firms because it is based on a survey that chooses representative firms in the base year (though these firms make up 80% of manufacturing output). In addition it does not pick up entering firms (most likely small) between the two base periods. In the attempt to make the data as comparable as possible, I produce a revenue growth measure that only includes firms that are in the database for 7 years or longer. As a robustness check, I also compare results when the census of firms are included in the revenue measure. The aggregate data implies that revenue productivity grew much faster than physical productivity up until 2001. This trend was reversed after the terms of trade shock. The structural model in the next section will describe how reallocation explains the difference between these two measures, and why this is entirely consistent with the markup heterogeneity explored above.

Figure 3: Real Income Growth versus Physical Production



Real revenue is the growth in the sum of deflated value added (minus primary input growth) at the 2-digit ISIC level. Economy-wide average taken by weighting each 2-digit group by its value added share. I allow for value added shares to change over time, although constant shares results in almost identical growth rates. Quantity growth is taken from the physical manufacturing index provided by the ENIA at the 2-digit ISIC level with same weighting scheme. Sector 27 is eliminated as in the rest of the analysis since this sector is made up mostly of copper.

I stress that the divergence in these two measures is large in economic terms. In the context of the Chilean economy, I can run the following thought experiment: given a starting point for aggregate value added, what would be the implied real revenue at the end of a period if it is assumed to grow proportionally with physical output versus using the growth rate of the actual revenue growth. Using the respective growth rates aggregated to the manufacturing level, and aggregate value added in manufacturing in 1995 and 2002, I examine two sub-periods: a) Starting from 1995, ignoring the growth rate of misallocation results in revenue that is 41% below actual revenue in 2002 (translates to

2.3 trillion Chilean pesos, or 3.3 billion US dollars); b) Starting in 2002, ignoring misallocation results in revenue that is 22% greater than actual revenue in 2007 (translates to 2.5 trillion Chilean pesos, or 4.8 billion US dollars).[29] Given that these two measures move together absent any misallocation in the economy, these two separate sub-periods provide evidence that changes in allocative efficiency can provide either an amplification or dampening effect on aggregate income depending on whether the economy is becoming more or less resource efficient. [30] [31]

3 Structural Estimation of Allocative Inefficiency

3.1 Setup of Variable Elasticity Model

In this section I explore how misallocation can explain the divergence in the growth rates of real revenue and physical output observed in Chile. I succinctly describe a framework with directly additive preferences that deliver variable demand elasticities, as fully laid out in Dhingra and Morrow (2019) and Zhelobodko et al. (2012). I add an upper tier to describe the interaction across sectors, but focus on the intra-sector action. This sets up an environment in which markup heterogeneity is the driving factor behind allocative inefficiency.

The economy is made up of L workers that supply one unit of labor inelastically. M^e represents the mass of entering varieties, with each firm drawing c, its marginal cost or labor requirement to produce one unit, from a distribution $G(c)$, a continuously differentiable cumulative distribution. Then, c_d is the highest possible cost with positive demand, so that active firms have costs in the range: $c \in (0, c_d]$ and $M^e G(c_d) = N$ represents the mass of varieties supplied. I index firms by their cost, c, with the quantity sold to each (identical) consumer and price being $q(c)$ and $p(c)$ respectively. Preferences are given by the following aggregation across and within sectors, j:

[29] Manufacturing valued added accounts for 20% of the economy in 2002, and 13% of the economy in 2007.

[30] Using the full sample of firms for revenue growth slightly reduce the difference between these two measures but the signs remain the same.

[31] To check the importance of entry, Figure 8 in the Appendix shows the number of firms that entered and exited from 1996-2006. Except for 1996 and 2006, they tend to move together so that net entry is not large. Both measures are about 60 on average (to put that into context, the census lists approximately 5000 firms per year). The value added of new entrants is on average only 5.8% of the economy.

$$U \equiv \prod_{j=1}^{J} U_j^{\beta_j} \tag{2}$$

$$U_j(M_j^e, q_j) \equiv M_j^e \int_0^{c_d^j} u_j(q_j(c)) dG_j(c) \tag{3}$$

$$\text{s.t} \sum_{j=1}^{J} M_j^e \int_0^{c_d^j} p_j(c) q_j(c) dG_j(c) = w \tag{4}$$

Notice the first line assumes that there is Cobb-Douglas aggregation across sectors. This will allow me to focus on the intrasector misallocation and assign each sector a constant weight β_j.[32] I assume that utility within sectors takes a variable elasticity form and is *additively separable across products*. Although this allows for any range of demand elasticities, I restrict myself to preferences where the inverse demand elasticity is increasing with quantity.[33] For this reason more productive firms (producing a differentiated good with a lower marginal cost) will produce higher quantity, but have more market power and charge higher markups than their less productive counterparts. For the rest of this section I mostly drop the subscript j to concentrate on the within-sector analysis and return to the definition of industries in the empirical framework (below I will selectively use j subscripts in some equations to highlight that they hold at sector level). Notice that the Cobb-Douglas aggregation assumes there will be no interesting sectoral interaction, although misallocation could also be present across sectors.[34]

For each variety there is inverse demand of, $p(q(c)) = \frac{u'(q(c))}{\delta}$, where the shadow price of income is $\delta = M^e \int_0^{c_d} u'(q(c))q(c)dG$. There is a competitive labor market with labor mobile across sectors, so that firms take as given a common wage, w. This common wage can thus be normalized to one. Firms pay a fixed entry cost, f_e, to choose a cost from the distribution, and then only active firms pay a fixed cost of production, f. These firms maximize profits, $\pi(c) = [p(q(c)) - c]\, q(c)L - f$. With monopolistic competition firms set their marginal revenue equal to marginal costs and the the markup rate is equal to the

[32]To tie this to the empirical application, a sector will be comprised of products within a 2-digit ISIC code. Aggregate measures will be constructed and vary at this level.

[33]This is the case most often chosen in the literature, which Mrazova and Neary (2018) call "Marshall's Second Law of Demand". It is also the pro-competitive case in Krugman (1979). I am partial to Paul Krugman's words that to get reasonable results, "I make this assumption without apology".

[34]See Epifani and Gancia (2011) for a nice application of inter-sector misallocation. In Chile, the labor share of 2-digit industries are remarkably constant during the 13 years I study. Behrens et al. (2016) derive separate intra-sector and inter-sector allocative efficiency measures for a similar model. With the Cobb-Douglas upper-tier, the two inefficiencies can be studied independently.

inverse demand elasticity: $\mu(q) = |\frac{qu''(q)}{u'(q)}| = |dlnp(q)/dlnq| = (p(c) - c)/p(c)$. The Lerner index, or the degree of market power, was my measure of markups in Table 1. Free entry implies the following sector-specific conditions: $\pi(c_d)) = 0$ and $\int \pi(c)dG = f_e$. Therefore, in the language of Dixit and Stiglitz (1977), the social optimum is a "constrained optimum" since firms need to be compensated for the chance of losing the entry cost.

In the market equilibrium, firms charge variable markups. The firms' first order conditions imply that for all firms: $u'(q(c)) + u''(q(c))q = \delta c$, or $u'(q(c)) = \frac{\delta c}{1-\mu(q(c))}$. Given that $p = u'(q(c))/\delta$:

$$p(q(c)) = \frac{1}{1 - \mu(q(c))}c \tag{5}$$

With non-homothetic preferences, the price is not a constant over marginal cost because $\mu(q(c))$ is a function of firm-varying productivity (or marginal cost). In other words, market power is heterogeneous across firms within a sector, and firms do not equate marginal rates of transformation.

3.2 Allocative Inefficiency

Dhingra and Morrow (2019) show that the model above leads to distortions not present in the standard CES model because the market equilibrium is socially optimal only when preferences are CES. Building on their work, this paper identifies the difference in the growth rate of revenue due to reallocation in the variable elasticity model relative to the commonly used CES framework. I motivate the importance of revenue growth in a welfare decomposition and use the definition of revenue (separating prices and quantity) to measure the bias inherent in the CES assumption relative to the generalized demand.

3.2.1 Utility with CES

I start by decomposing welfare when utility is homothetic, the knife-edge case where welfare is proportional to revenue, and compare that case to a generalization where utility is non-homothetic. If the sub-utility is assumed CES, aggregate real revenue is proportional to welfare because $u(q) \propto qu'(q)$, which means we can relate utility to aggregate real revenue ($qu'(q) \propto p(q)q$). From the definition of preferences and the consumer budget

constraint, the following describes utility in each sector j:

$$U_j = M_j^e \int u_j(q_j) dG_j \propto M_j^e \int_0^{c_d^j} u_j'(q_j(c)) q_j(c) dG_j(c)$$
$$\propto \lambda_j L_j \big(\int_0^{c_d^j} \frac{1}{1-\mu_j(q_j(c))} dG_j(c) \big) \big(L_j - N_j f - M_j^e f_e \big) \tag{6}$$

where N_j is once again the mass of varieties supplied and λ_j is the Lagrange multiplier in the social problem of utility maximization.[35] The last line uses the budget constraint (the total resources in the economy), and that $Cov(\frac{1}{1-\mu(q(c))}, cq(c)) = 0$. In this case, welfare is proportional to the average markup times the total labor used for production.

3.2.2 Utility with VES

I now generalize to the non-homothetic case where the subutility is not CES, which imples $Cov(\frac{1}{1-\mu(q(c))}, cq(c)) \neq 0$. In this case utility and aggregate revenue will diverge, and I show how to decompose this divergence. Since cost advantages are not fully passed through to prices, some firms under-produce and others over-produce, which distorts total revenue relative to the CES benchmark. [36] Furthermore, it has been known since Dixit and Stiglitz (1977) that an inefficiency exists even with homogeneous firms due to a distortion in the number of available varieties.[37] I decompose the full welfare expression in my model to express clearly how the misallocation term in my model captures a distortion from the CES case and builds on the variety distortion described in this earlier work.

Aggregate revenue is defined as, $R = M^e L \int_0^{c_d} p(q(c)) q(c) dG(c)$. I will work with the conditional distribution of $g(c)$ on $(0, c_d]$, defined as follows:

$$h_d(c)dc = \begin{cases} \frac{g(c)}{G(c_d)} dc & \text{if } c \leq c_d, \\ 0 & \text{if } c > c_d \end{cases} \tag{7}$$

It will be useful to define the average price level, $P \equiv \int_0^{c_d} p(q(c)) h_d(c) dc$ and aggregate physical production sold, $Q \equiv NL \int_0^{c_d} q(c) h_d(c) dc$. Let the "elasticity of utility" be: $\epsilon(q) = \frac{\partial u(q)}{\partial q} \frac{q}{u(q)}$, the proportional increase in utility given an increase in the quantity of a variety. Then, as in Dhingra and Morrow (2019), the (utility-weighted) average elas-

[35] A sector aggregate that represents the shadow value of resources.

[36] Notice that this framework is consistent with the results of Edmond et al. (2015) and Arkolakis et al. (2019), who both find that it is the *joint distribution* of markups and production that matters. Alternatively, the intuition is that the whole distribution of markups matters, not the unweighted mean.

[37] Chamberlin (1933) also argued for "excess capacity" which resulted in excess entry.

ticity of utility is $\bar{\epsilon} = \frac{\int \epsilon(q)u(q)}{\int u(q)}$. Using this definition, the indirect utlity function is defined as $V = \frac{1}{\bar{\epsilon}} M^e L \int u'(q(c))q(c)dG(c)dc$. I then plug in the inverse demand function, $u'(q(c)) = \delta p(q)$ (with δ as the marginal utility of income), and conduct algebraic manipulations on revenue, to decompose revenue within the indirect utility function:

$$V = NL\frac{\delta}{\bar{\epsilon}} \int_0^{c_d} p(q)q(c)h_d(c)dc$$

$$\Delta ln(V) = \Delta ln(1-\bar{\epsilon}) + \Delta ln(\delta) + \underbrace{\Delta ln\,(Q) + \Delta ln\,(P) + \Delta ln\left(\frac{\tilde{R}}{Q}\right)}_{\Delta ln(R)}, \tag{8}$$

where $\tilde{R} = \frac{R}{P}$. Vives (2001) on page 170 refers to $(1-\epsilon(q))$ as "the proportion of social benefits not captured by revenues when introducing a new variety." Since the elasticity of utility is constant under CES preferences, in that case $-\Delta ln(\bar{\epsilon})$ is zero.[38] Along with the change in the marginal utility, it will be a part of the change in indirect utility that I do *not* capture by focusing only on $\Delta ln\,(R)$, the part that is captured in the data. Note also that with a Pareto distribution of firm costs, entry is optimal (Arkolakis et al., 2019; Feenstra and Weinstein, 2017). Equation 8 motivates why we care about revenue: the change in revenue is a component of welfare growth.

The central motive for this decomposition is to establish *the bias in revenue growth over time using CES demand relative to the generalized variable elasticity demand*, within the monopolistic competition framework described above. I argue that the last term is the bias in aggregate revenue that is not captured by the allocatively efficient case, which requires CES subutility. To get an intuition about the last term in 8, it is helpful to derive it from the aggregate revenue equation. I decompose aggregate revenue in terms of mean and variances using the covariance: $\text{Cov}(p,q) = \int_0^{c_d}(p(q(c)) - P)(q(c) - \frac{Q}{NL})h_d(c)dc$. Then,

$$\Delta ln\left(\frac{\tilde{R}}{Q}\right) \approx \Delta\left(\frac{\text{Cov}(\frac{1}{1-\mu}, l)}{\int_0^{c_d} p(q(c))h_d(c)dc \int_0^{c_d} q(c)h_d(c)dc}\right). \tag{9}$$

$\Delta \ln\left(\frac{\tilde{R}}{Q}\right)$ is therefore proportional to the covariance of the markup and labor allocation in each firm.[39]

In order for $\Delta \ln\left(\frac{\tilde{R}}{Q}\right)$ to represent the change in allocative efficiency captured by aggre-

[38] In which case: $\frac{1}{\bar{\epsilon}} = \frac{1}{1-\bar{\mu}} = \frac{\sigma}{\sigma-1}$, with σ the constant elasticity of substitution.

[39] Full details are in Appendix B. I write the covariance in terms of markups and labor by using $pq = \frac{p}{c}cq$, where cq is the labor required to produce q units.

gate real revenue I will show that it is zero only in the case where there is no inefficiency, which is true when demand is CES. Furthermore, I eliminate the effects due to changes in the cost cutoff by assuming that $G(c)$ is a Pareto distribution.[40] In other words, the measure provides a sufficient statistic for the correction in real revenue due to reallocation that is not captured in the Melitz-Chaney framework. For this to be true, the following proposition is necessary:

Proposition 1. *In the VES framework described above, and if $G_j(c)$ is a Pareto distribution, then $\Delta \ln(\tilde{R}_j) = \Delta \ln(Q_j)$ if and only if demand in sector j is CES.*

Proof. In Appendix B I take the case of CES preferences and Pareto distribution of costs and show that the right hand side of Equation 9 is zero.

The second part of the proof is to show that if $\Delta \ln \left(\frac{\tilde{R}}{Q}\right) = 0$ then preferences must be CES. Assume preferences are not the knife-edge CES case, then the within-sector preferences described in 3 are non-homothetic. Then, the price is a function of quantity and this contradicts that the left hand since of Equation 9 is equal to 0. ■

Given that the market power distortion exists only in the non-efficient market equilibrium, I label the change in the covariance term as ΔAE:

$$\Delta(AE_j) = \Delta ln \left(\frac{\tilde{R}_j}{Q_j}\right) \tag{10}$$

Notice that this term is the log change of the *aggregate* markup, which depends on the joint distribution of markups and production. In the market allocation, high (low) markup firms under- (over-) produce. *Reallocating production to the high-markup firms raises the aggregate markup and welfare*. The rise in welfare would be due solely to a compositional effect.[41] With constant markups this term is constant – in fact $\Delta \ln(V) = \Delta \ln(Q)$ in the CES case. Lastly, notice that this holds for each industry j separately.

3.3 Discussion

Equation 10 measures the real revenue change due to a change in the distribution of the price-cost ratios, holding productive efficiency constant. This measure is therefore similar to the Holmes et al. (2014) (HHL) allocative efficiency index that is separable from

[40] For this reason, the analysis should be viewed as reallocation across existing firms. Data availability would make it very difficult to capture the effect of entry on welfare without a more stylized model.

[41] This is a result emphasized in recent papers: (Macedoni and Weinberger, 2019; Peters, 2018; Edmond et al., 2018; Baqaee and Farhi, 2017). However, the sufficient statistic that captures changes in this weighted markup is new. I will connect foreign shocks, firm reallocation, and misallocation transparently in 4.2.

production efficiency to measure gains from trade. However that index only holds for homothetic tastes. The features of that model make it very useful for analyzing the effects of a symmetric trade cost, but not for taking into consideration how firms might pass-through costs shocks to prices which has been shown to be a big part of short-term adjustments to trade liberalization (DeLoecker et al. (2016)) (DGKP). The comprehensive studies of misallocation in Edmond et al. (2015) (EMX) and Hsieh and Klenow (2009) (HK) also focus on an aggregate TFP index due to homothetic demand. My measure is not tied to an aggregate TFP measure but instead non-homothetic preferences result in a welfare distortion in market allocations – differentiated from the supply-side interpretation of those papers. Therefore, the measure in Equation 10 can be viewed as a complement to those papers.

ACDR propose a gains from trade decomposition with variable markups that is closest to the one in this paper, as it includes a misallocation distortion.[42] In that paper, a reduction in output tariffs has two effects: the distortion among domestic firms is *reduced* due to an increase in competition, but the distortion *increases* among foreign firms as they face lower marginal costs. In my paper, aside from providing a more direct measure of changes in allocative efficiency which is captured with aggregate data, I extend their work by separating foreign shocks and isolating their effect on misallocation – in the next section I focus on an exchange rate shock, but a similar analysis can be done for a rise in competition due to foreign entry. In fact, in Appendix E I provide comparative statics of the effect of both supply and demand shocks on the *markup differences* between firms. In a very stylized model, I show that the reallocation response of domestic firms differs across these shocks.

4 Allocative Efficiency in Response to Exchange Rate Shocks

In Sections 2.3 and 2.4, I provided suggestive evidence of how real exchange rate shocks might affect allocative efficiency by plotting the time series of markup dispersion, real revenue growth, and physical production growth. After appreciation in the price of copper – Chile's main export commodity – markup dispersion increased and the growth rate of physical production was larger than the growth rate of real revenue. A framework with variable markups provides a rationale for these observations: a terms of trade appreciation reduces the cost of importers, which allows them to produce more, but some of those costs are passed on to markups.

[42] They show that this distortion is proportional to the covariance of markups and firm-level employment shares (of domestic firms).

With heterogeneous firms, where larger firms are characterized by both higher markups and a larger exposure to imported inputs, a REER shock will induce the compositional effect that determines (10). For example, if more productive firms can pass-through a larger proportion of cost-savings to their markups, production is reallocated to less productive firms so that the growth rate in revenue is smaller than implied by the case where pass-through is equal across firms. In this section I show that in fact the real exchange rate variation in Chile led to relatively different changes in allocative efficiency in industries depending on their exposure to imported inputs relative to exports. Then, I clarify the mechanisms at work with a structural model where firms face VES demand and imported inputs generate productivity gains. As the exchange rate varies, the model generates changes in misallocation based on the reallocation responses emphasized in the previous section.

4.1 Empirical Evidence

To identify the aggregate allocative efficiency responses to Chile's exchange rate volatility, I estimate a difference-in-difference specification in which industries are treated differentially based on their exposure to the real exchange rate shock. This is analogous to the specification in Section 2.3, now measuring the aggregate response of the same shock, allowing for reallocation across firms with heterogeneous markups.

The main Industry outcome is ΔAE, the misallocation measure introduced in Section 3.2. I also check ΔQ, to confirm that production behaves as expected given the exchange rate shock, along with $\Delta cov(markup, inputs)$, which should move in conjunction with allocative efficiency given in (9).[43] The specification will follow very closely the firm-level regressions (1), but at the industry level:

$$\Delta AE_{jt} = \alpha_j + \alpha_t + \psi \Delta REER_t * Expos_j + \zeta \Delta Z_{jt} + u_{jt}, \tag{11}$$

where the key outcome variable will be the growth rate of allocative efficiency as constructed in the structural model. α_t and α_j represent time (t) and industry (j) fixed effects respectively. I use $REER$ to represent the aggregate shock, but in the appendix I also report results with the TOT.[44] ψ measures the differential effect on industries "exposed"

[43] I calculate the covariance of the markup with both material inputs and labor costs. In the theory, labor is the only input (costing cq), so my reported results reflect the covariance of the markup with wages paid.

[44] One might worry about an omitted variable bias as exchange rates respond endogenously to other macro shocks. In Chile, the appreciation in 2003-2006 is plausibly exogenous as it is due to an unexpected boom in copper prices. I eliminate all copper-related industries from the regressions and interpret the

to the shock relative to other industries. The question is: "do industries that are more exposed to the large terms of trade appreciation have relatively lower growth in allocative efficiency?" In this case, 'Net Exposure" is constructed so that the industry level exposure to inputs relative to exports is the average of all firms in the 2-digit ISIC sector (still fixed over time). As the net exposure becomes more negative this identifies a sector that imports a larger share of its inputs than its export share of sales. I also use an average of importer (but not exporter) and exporter (but not importer) dummies as a separate measure of exposure, with results in the Appendix. Given the results in the previous section, I expect that a REER appreciation will have a larger impact on the allocative efficiency of industries with a larger share of importers. I control for growth rates in important sector-level time-varying characteristics: the 2-digit Herfindahl index (HHI) of sales, and the sector import tariff. The former controls for changes in misallocation due to a different competition environment.

The regressions use value added weights for sectors, which gives an empirical counterpart to β_j in Section 3.1. In using variation across sectors and years, I of course measure only *within-industry* misallocation. However, the similarity in my weighted average of industry allocative efficiency measure with an aggregate measure that pools together all industries is evidence that inter-sector reallocation was not an important part of changes in manufacturing allocative efficiency. As in 2.4, I eliminate firms that do not produce for 6 consecutive years. In results that do not eliminate these entering firms, the regression results are very similar both qualitatively and quantitatively.

Regression Results Tables 2 displays the results for industry level outcomes in response to variability in the REER. The first two columns illustrate that industries with a higher exposure to importing intermediates than exporting their final product become more misallocated in response to an increase in the terms of trade. In comparison to the first column, the second specification includes changes in the Herfindahl index. One way to interpret this coefficient is to compare industries with different extreme values of net exposure. For example, an industry with firms that import all of their inputs but do not export will have a net exposure of -1. Net exposure of 0 means the ratio of exports to sales is equal to the ratio of imports to total inputs (or it could signify no import or exports). Therefore, the coefficients in the first row of Columns (1) and (2) can be interpreted as: an industry with net exposure of 0 has allocative efficiency growth that is 1.96-2.75 percentage points larger than the industry with net exposure of -1 in response to a 1% increase in the growth of the real effective exchange rate. As expected, the importing industries become more mis-

copper price shock as exogenous to other industries.

allocated with terms of trade gains.[45] Column (3) decomposes exposure into the export and import shares. It is the larger import share that drives reductions in misallocation in response to TOT shocks. This is important because the elasticity of exports and imports to the exchange rate could differ, which is not picked up by only using a *net* exposure. In Tables 9 and 10 in the Appendix, I check that these results hold for varying exposure shares, and with TOT variation instead of the REER.

Column (4) investigates the response of physical production, and as expected quantity production increases for negatively exposed firms in response to an appreciation. In columns (5) and (6), the signs are consistent when replacing the allocation efficiency measure with the covariance of markups and labor expenditure. The interpretation of these columns is consistent with higher misallocation as a result of allocating inputs away from the more productive firms. The lower input costs allow for productivity gains, but more productive firms pass-through relatively more of their cost reductions to markups, thus raising misallocation. [46]

In summary, in response to appreciations in the terms of trade, industries that have a higher share of importers relative to exporters become more misallocated. Figure 4 reiterates that only the industries with exposure to importing are responding to the terms of trade shocks with the expected sign. I separate industries into a binary "negative" exposure if industries are below the median of exposure across all industries, with the rest being "positive". I then take the average ΔAE for all industries in each category per year (weighting by value added), and plot this against the change in the terms of trade.[47] Each point represents one year (for each type of industry). ΔAE has a clear negative relationship with changes in the terms of trade *only for* negatively exposed industries, while this relationship is horizontal for the other industries.

4.2 Quantitative Model of Importing with VES Preferences

The previous subsection provides reduced form evidence of changes in allocative efficiency in response to a shock in the exchange rate. In this subsection I highlight the mech-

[45] Another way to interpret the magnitude of these results is to create a binary variable for "exposure." Given the sector averages, I define an industry as negatively exposed ($NegativeExposure = 1$) if the average net exposure is less than -0.1 (this was the median exposure across industries). The results are available upon request but not shown in the Table since they tell the same story.

[46] In Table 11 (Appendix), the main result is that when the REER increases, industries with a larger fraction of importers (that are not exporters) suffer in terms of allocative efficiency (first row of Column (1)). The last two columns yield the same interpretation as the results in the previous table.

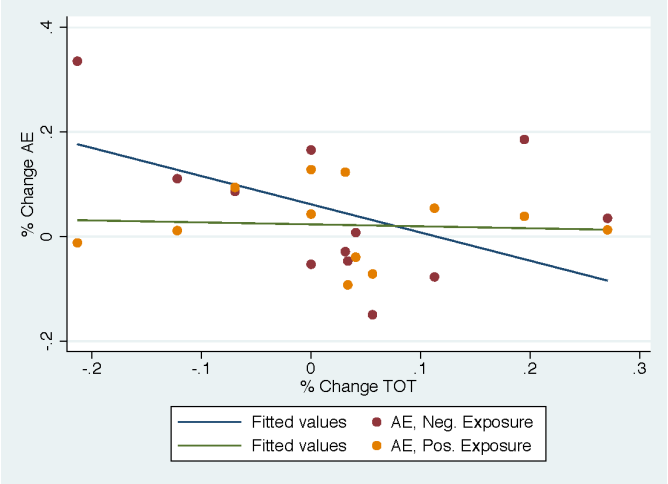
[47] This highlights the effect of the appreciation in the latter part of the sample, which is the focus of the next section.

Table 2: Industry Responses to Exchange Rate Shock: Fixed Exposure Shares

	Δ AE			Δ Q	Δ Cov(markup,inputs)	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ REER*Net Exposure	2.747***	1.964**		-1.163	3.607***	
	(0.806)	(0.755)		(0.845)	(0.559)	
Δ REER*Imported Share			-3.449***			-2.891***
			(1.145)			(0.525)
Δ REER*Exported Share			-2.050			8.487**
			(6.385)			(3.236)
Δ Tariff	0.034	0.034	0.033	0.020	-0.010	-0.009
	(0.032)	(0.030)	(0.032)	(0.048)	(0.013)	(0.014)
Δ HHI Index		1.944***		0.344***		
		(0.522)		(0.095)		
Fixed Effects	Year,Sector	Year,Sector	Year,Sector	Year,Sector	Year,Sector	Year,Sector
R^2	0.243	0.328	0.244	0.323	0.200	0.203
N	192	192	192	192	204	204

This Table reports the response in aggregate industry measures in response to REER shocks, where industry exposure is based on the Ekholm et al. (2012) Net Exposure measure. Dependent variables are ΔAE (columns (1)-(3)), ΔQ (column (4)), and Δ Cov(markup,inputs) (columns (5)-(6)). Between columns (1-2) and (3), and between (5) and (6), the difference is in the measure of exposure: "Net" versus splitting up export share and import share. These are all at the 2-digit ISIC level. For the covariance I use first differences, while the first 4 columns are growth rates. $\Delta REER$ is a one year growth rate. Net Exposure is defined as (Export Sales/Total Sales)-(Imported Inputs/Total material input costs). The prior two components are "Exported Share" and "Imported Share." These shares are fixed over time. Industry controls are changes in: the Herfindahl Index of concentration and industry output tariffs. All regressions include sector and year fixed effects and standard errors clustered at the 2-digit industry level (in parenthesis). I drop the basic metal industry (ISIC 27). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Figure 4: Δ AE vs Δ TOT for Positively Exposed vs Negatively Exposed Industries



Industry categorized as "Negatively Exposed" if the average net exposure variable is less than the median across all industries. Within each year, I take the average Δ AE for the two categories by aggregating using value added shares. Each point represents a combination of the change in the terms of trade and change in allocative efficiency in a year, for each type of industry.

anisms that underlie how reallocation improves or worsens misallocation. I introduce an aggregate shock and show how the differential adjustment across firms sets off a reallocation that either increases or decreases misallocation. Generally, the shock is felt by firms in two separate ways: either as a demand shock that affects the level of competition, or a supply shock through the cost of its inputs. To connect the theoretical framework to the

case of Chile, which experienced large real exchange rate variation during the sample period highlighted above, I introduce an importing margin to a general equilibrium model that features consumers with an explicit form of non-homothetic preferences nested in the general environment of the Section 3. As the aggregate shock alters the relative price of domestic versus foreign goods, domestic firms respond differentially to the change in the cost of importing, and the model generates an endogenous change in the allocative efficiency highlighted above. Although I choose to model the supply shock as an explicit example that isolates one potential effect of globalization, the allocative efficiency effects of a competition shock can be modeled as well.[48]

4.2.1 Closed Economy

Consumer Problem I drop the multiple-sector assumption in order to focus exclusively on the distributional effects within sectors. Preferences will now take the specific functional form:

$$U = \left(\int_{\omega \in \Omega} (q(\omega) + \bar{q})^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}, \tag{12}$$

where $q(\omega)$ is individual consumption, Ω is the set of potentially produced goods, $\bar{q} > 0$ is a constant, and $\sigma \geq 1$ governs curvature. These preferences are a type of "generalized" CES that correspond to a Dixit-Stiglitz utility function with a displaced origin, and have been applied recently by Jung et al. (2019) and Arkolakis et al. (2019). For $\bar{q} = 0$, the utility function is of the CES form, but a positive $\bar{q}$ implies that firms face less elastic demand as their sales increase. I make the latter assumption to guarantee that more productive firms have larger markups and higher sales, as is clear in the data. For more details on the advantages of this preference specification, see Jung et al. (2019).

In the estimation of the model, I use the specification above which allows for flexibility in the curvature of firm demand. For exposition purposes, in this section I will take the case where $\sigma \to 1$ which allows for closed form solutions.[49] The utility function becomes:

$$U = \int_{\omega \in \Omega} \ln(q(\omega) + \bar{q}) d\omega, \tag{13}$$

which essentially restricts the substitution across varieties.

[48] Appendix E illustrates a case where supply and demand shocks affect markup differences between firms in two separate ways. That is a stylized model with predictions about firm reallocation, whereas the model in this section introduces importing, and therefore supply shocks, in a way where I can then show changes in ΔAE.

[49] The appendix includes details for the general model.

Firm Problem Aside from the ability to import, the supply side is mostly identical to the setup in Section 3.1. However, to simplify the estimation, I drop the production fixed cost. The preferences outlined above feature a bounded marginal utility, which means that consumers do not have positive demand for all potentially produced goods. Firms pay an entry cost to draw their intrinsic productivity from a Pareto distribution with κ the shape parameter of that distribution. Given the attributes of demand, there exists a choke price such that demand is zero for goods whose price exceed it. Since details of this model can be found in Simonovska (2015), I relegate large parts to Appendix C.2 and include only the most important components in the main text.

Given the preferences above and the setup described, the consumer first order conditions imply an inverse demand function that sets up the following firm profit equation (where firms are identified from their cost draw, c):

$$\pi(c) = \frac{L}{\lambda}\left(\frac{q(c)}{q(c)+\bar{q}}\right) - Lq(c)wc, \tag{14}$$

where the aggregate environment is taken as given for each firm and summarized by $\lambda = \int_0^{c^d} \frac{q(c)}{q(c)+\bar{q}} dG(c)$[50], and w represents the real wage. Firms choose $q(c)$ to maximize profits, and setting quantity equal to zero yields the cutoff cost draw with positive demand:

$$c^d = (\lambda w \bar{q})^{-1} \tag{15}$$

As is common in the literature (e.g. (Arkolakis et al., 2019; Jung et al., 2019)), the competitive environment is summarized by a single aggregator, the market determined cost cutoff. Quantity demanded, price, and revenue can be written as a function of the the firms' cost relative to the market cutoff:

$$p(c) = wc\underbrace{\left(\frac{c^d}{c}\right)^{\frac{1}{2}}}_{\text{markup}} \tag{16}$$

$$q(c) = L\bar{q}\left[\left(\frac{c^d}{c}\right)^{\frac{1}{2}} - 1\right] \tag{17}$$

$$r(c) = L\bar{q}w\left[c^d - \left(c^d c\right)^{\frac{1}{2}}\right] \tag{18}$$

[50] A useful feature of this framework is that it will not be necessary to solve for the general equilibrium term, λ, as is clear below.

4.2.2 Importing

The ability to import inputs will generate interesting implications from a shock to the relative price of foreign and domestic goods. I introduce the ability to import in the simplest way possible, by making the assumption that the share of foreign inputs in the total input cost is proportional to firm productivity. This restricts the set of firms to have positive imports, so in mapping this to the data one interpretation is to take a balanced panel of importers with heterogeneous import shares, so that an aggregate shock shifts the import share for all firms.[51] Although this is a stark assumption, it generates the differential response to a shock in the cost of imports, which is clearly enough to highlight the mechanism through which this shock generates changes in allocative efficiency.

I incorporate imports by differentiating between the intrinsic cost c, and a unit cost, u, which includes the cost reductions that result from importing. To model the cost reductions, I rely on the theoretical results of Blaum et al. (2018) which allows me to skip the sourcing decision. They show that in a variety of structural models[52], the unit cost is proportional to the domestic share of inputs. Directly from their results, I take the following specification for unit costs:

$$u = wc(\tau\xi)^{-\beta}, \qquad \xi = \frac{1}{s_D} = c^{-1/\theta} \tag{19}$$

where $(\tau\xi)$ captures the cost advantage of importing, governed by a parameter β, s_D is the share of inputs that are sourced domestically, and w captures the domestic input costs which will play no role. c continues to be the firm cost draw, but the unit cost is now given by u, which can be computed by substituting in the expression for ξ. Notice that ξ is simply the inverse of the domestic input share, so that firms receive a cost reduction from raising their imported share. Further, the inverse of the domestic share is proportional to the cost draw of the firm, which captures the fact that there is a sorting in the data where more productive and larger firms have larger import shares.[53] τ is a parameter

[51] Although I don't model this explicitly, it is consistent with conditioning on the firms that have paid a fixed entry cost for importing. A full structural model with non-importers and importers would include the entry cost f_e above, plus a second fixed cost to enter the import market (akin to the export fixed cost in Melitz (2003)). However, with non-homothetic preferences, the decision to pay the second entry cost becomes burdensome, and would not add any insight in terms of generating changes in allocative inefficiency when all firms respond to the aggregate shock. The general model would produce *quantitatively* different results as only a subset of firms would receive a cost shock (the original importers plus any firm that enters/exits the import market).

[52] Halpern et al. (2015), Gopinath and Neiman (2014), Antras et al. (2017), Kasahara and Rodrigue (2008), Lu et al. (2016), Goldberg et al. (2010), Amiti et al. (2014).

[53] This assumption allows for a parsimonious framework to incorporate imports. One could combine a

which alters the relative price between imported and domestic inputs, which I set equal to 1 to start and then vary it to mimic the exchange rate volatility. In comparison to the structure of Blaum et al. (2018), β governs the productivity gains from importing, which in their model is identified by the ratio of the material output elasticity and the elasticity of substitution between inputs. Firms' imported input shares will depend on θ, which I discipline using the average import share across all importers in the Chilean data.

The aggregate environment can be expressed as a function of τ. First, notice that given the new unit cost of the firm, the competitive environment is governed by a new cost cutoff:

$$c^* = (w\lambda\bar{q})^{-\frac{\theta}{\beta+\theta}}\,\tau^{\frac{\theta\beta}{\beta+\theta}}, \tag{20}$$

which increases with τ. To see the firm-level effect, for example take the case when the relative price of imports is lower. Every firm sources more products from abroad, and one can show that markups increase for all firms (Appendix C.2), reflecting the incomplete pass-through of a productivity shock to prices.

To see the aggregate consequences on allocative efficiency I derive the measure of misallocation in Section 3:

$$\frac{\tilde{R}}{Q} = \left(\frac{c^*(\tau)}{c^*(\tau=1)}\right)^{-\frac{1}{2}(\frac{\beta}{\theta}+1)} \left[\frac{1-\frac{\kappa}{\frac{1}{2}(\frac{\beta}{\theta}+1)+\kappa}}{\frac{\kappa}{\frac{1}{2}(\frac{\beta}{\theta}+1)+\kappa}\left(\frac{\kappa}{\kappa-\frac{1}{2}(\frac{\beta}{\theta}+1)}-1\right)}\right] \tag{21}$$

Notice a convenient feature of this formulation is that the cutoffs can be normalized by the case when $\tau = 1$. In this framework, the general equilibrium object is a function of parameters and I normalize it to one in the benchmark case.[54] Then, with importing shocks, $\frac{c^*(\tau)}{c^*(\tau=1)} \neq 1$ with varying τ, as the competitive environment changes in the domestic economy. An interesting case is $\kappa \to \infty$, as would be the case without firm heterogeneity: in that case, $\frac{\tilde{R}}{Q} = 1$. The reason is that the misallocation I capture in this paper is due solely to firm heterogeneity, as the Pareto assumption guarantees that entry is optimal.

Why does this matter? The ratio of aggregate real revenue relative to aggregate production, or the aggregate markup, decreases with τ, which I show quantitatively below. The intuition is simple: responses to the shock are heterogeneous as they depend of the

fixed importing cost with a separate draw from a distribution of domestic input shares. This structure could be used to create a closer relationship with the correlation of sales and imports in the data (by breaking the sorting that happens due to my assumption), but would only alter the results quantitatively.

[54] See Jung et al. (2019) for an expression of cutoff costs in general equilibrium with *multiple* countries, which are proportional across preferences that are directly additive and feature a Pareto distribution of costs. For example, the $\sigma \to 1$ case has a cutoff proportional to the general case. With multiple countries, cutoffs are a function of wages and trade costs.

firm-specific ratio of their cost to the cutoff. More productive firms import more and are larger, hence have larger markups, so the differential response to changes in the aggregate environment result in a reallocation of labor. The reallocation is inefficient in the case where τ leads to a lower price of foreign inputs relative to domestic ones. In fact, by using the model to simulate firms in this economy, I can show that a rise (drop) in τ generates relatively greater (smaller) increases in labor in firms with initially lower markups. This is the compositional change that drives Equation 10, and confirms the reduced form empirical evidence that heterogeneous markup responses to this shock lead to changes in the level of misallocation (in the correct direction). As production is reallocated to lower markup firms, aggregate real revenue increases more slowly than production.

4.2.3 Estimation

Constructing changes in allocative efficiency requires estimating only 3 parameters: $(\kappa, \frac{\beta}{\theta}, \sigma)$.[55] I use a simulated method of moments (SMM) approach and target 3 moments: the average firm import share, the correlation between markups and sales, and a moment of relative firm sales across the 99th and 90 percentiles. The first moment is necessary to discipline the relationship between productivity and importing. With VES preferences, non-constant markups generate a non-trivial relationship between sales and markups that varies with the demand curvature, which is the purpose of the second moment. Finally, κ and σ will affect the distribution of sales, so I use relative sales at varying percentiles to capture the sales distribution in the data.

The procedure consists of simulating a large enough number of draws so as to best approximate the entire continuum of firms that exist in the model. I follow the application in Jung et al. (2019), and relabel firm-level indicators that can be simulated from a parameter-free uniform distribution. Recall that the pdf of the cost distribution is given by $h(c) = \frac{\kappa}{c^{-\kappa-1}}$[56]. I draw 500,000 realizations of the uniform distribution on the [0; 1] domain, $U \sim [0; 1]$, and order them in decreasing order, and find the maximum realization, denoted by u_{max}. Then, the firm cost is $c = (u/u_{max})^{1/\kappa} c^*$. I normalize ($c^*$) to one in the case where $\tau = 1$, so that $(w\lambda\bar{q})^{-\frac{\theta}{\beta+\theta}} = 1$ (as explained above). This is without loss of generality, as all moments can be written without these parameters. Then, by construction

[55] With a general σ, there is no closed form solution for prices and therefore I cannot derive an equivalent to 21. However, it can be solved numerically (Jung et al., 2019). Qualitatively it shares the same features of the $\sigma \to 1$ case in terms of the relationship between productivity and markups, but quantitatively it affects the correlation between sales and productivity as the parameter governs the level of market power for each firm. Appendix C.2 describes the two types of preferences.

[56] Generally, the distribution includes a support parameter that determines the lower bound of the distribution. In this case, I set the support equal to 1 as it would just cancel out in the estimation if included.

$c \in [0, 1]$. As τ changes, $c^* = \tau^{\frac{\beta\theta}{\beta+\theta}}$.[57]

The simulated firms are used to construct the model-implied moments described above. The theoretical moments are matched to their counterpart in the data in order to identify the model parameters. Let $F_i^m(\kappa, \frac{\beta}{\theta}, \sigma)$ be the vector of model generated moments, where i represents each moment, and let F_i^d denote the corresponding value of the empirical moments. The data targets are constructed using all importing firms from 1995-2001.[58] Identification consists of choosing the parameter set that minimizes the sum of the squared errors between empirical and theoretical moments:

$$\min_{\kappa, \frac{\beta}{\theta}, \sigma} \sum_{i=1}^{3} \left(F_i^d - F_i^m(\kappa, \frac{\beta}{\theta}, \sigma) \right)^2 . \tag{22}$$

I compute bootstrap standard errors by running the estimation above 100 times, each time taking a bootstrap sample of the data. I take the average parameter estimates $(\hat{\kappa}, \frac{\hat{\beta}}{\theta}, \hat{\sigma})$, and use the standard deviation of estimates to compute standard errors. I estimate the parameters under $\tau = 1$, and then study the implications for allocative efficiency when τ varies.

Table 3, Panel A, displays the parameter estimates in an exactly-identified estimation. Panel B reports the targeted data moments, and the simulated ones using the parameter estimates. Each of the model-simulated moments are almost matched exactly.[59] The parameter estimates are reasonable when compared to the existing literature. The shape parameter of the productivity distribution is 7.63, higher than that found in Simonovska and Waugh (2014), but within the range in the literature (which bounds the parameter at 8). The demand curvature parameter is 2.93, which confirms that $\sigma > 1$.[60] The main difference with the rest of the literature is that in this model firms raise their productivity by importing. I estimate $\frac{\beta}{\theta}$ equal to 1.80, which magnifies the sales heterogeneity across firms relative to a model without these productivity gains (where $\frac{\beta}{\theta} \to 0$). This extra heterogeneity implies that a larger κ is necessary to match the sales distribution in the data. A reliable measure of β can be obtained from Blaum et al. (2018). They report that $\beta = \frac{.61}{2.38-1}$, using measures of the elasticity of demand between imports and the output

[57] For changes in τ, I re-compute the set of firms.

[58] I have also ran the estimation with one year of data, say 1997, to not capture the depreciation period. The effect on the parameter estimates is minimal. Appendix D provides details on this estimation procedure.

[59] Given that I cannot match these exactly, I compute the bootstrap standard errors.

[60] Jung et al. (2019) estimate a slightly lower κ. There are two important differences in the estimation relative to the rest of the literature. First, in this paper I match moments using *importing* firms only. Second, I allow more productive firms to have an extra cost advantage through imports.

elasticity of materials. Adopting their parameter implies $\hat{\theta} = 0.24$.[61]

Table 3: Calibration Results

Panel A: Parameter Estimates

Data Targets	$(\hat{\beta/\theta})$	$\hat{\kappa}$	$\hat{\sigma}$
Mean Import Share	1.80	7.63	2.93
Markup-Sales Correlation	(0.18)	(0.84)	(0.13)
Sales Percentiles (99 vs 90)			

Panel B: Moments: Data versus Model

Moment	Data	Model
Mean Import Share	0.35	0.35
Markup-Sales Correlation	0.11	0.11
90-90 Sales	1.71	1.70

Panel A reports the parameter estimates, with standard errors in parenthesis below the estimate. I calibrate the model for all importing firms. I target Data Moments by averaging across the 1995-2001 period. In Panel B, I compute moments in the data and using the simulated firms.

4.2.4 Counterfactual

With the necessary parameters in hand, I next investigate how changes in τ affect the allocative efficiency of the economy. Consider a counterfactual where τ ranges between 0.8 and 1.2, where $\tau < 1$ reflects a depreciation in the real exchange rate and $\tau > 1$ an appreciation. In Figure 5 Panel A, I plot the aggregate real income to quantity ratio, $\left(\frac{\tilde{R}}{Q}\right)$, at varying τ's. As an example, a 10% appreciation, which reduces the relative price of foreign goods, generates a reduction in allocative efficiency of 2.86% in a sector where the net exposure starts at 0.35 (average import share at $\tau = 1$) relative to an industry where net exposure is 0 (no importers).[62] Although aggregate revenues *increase* due to an appreciation, the rise in welfare is dampened by the welfare-reducing reallocation.

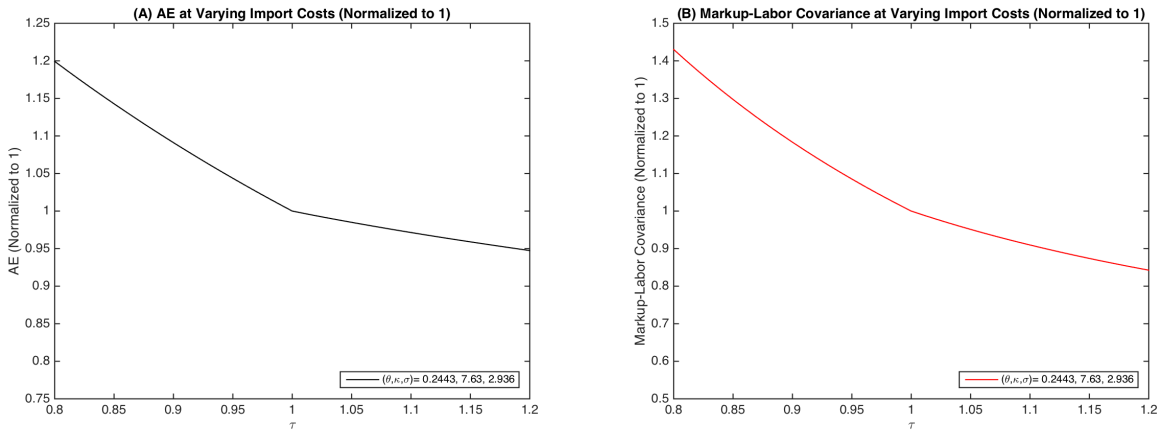
Panel B of Figure 5 plots the markup-labor covariance for $\tau \in (0.8, 1.2)$. A real exchange rate appreciation, which lowers the cost of imported inputs and raises markups,

[61] They use French customs data to estimate the elasticity of demand between imported inputs (ϵ in that paper), and firm data to estimate the material output elasticity (γ in that paper). Both are required for the productivity advantage of imported inputs, equal to $\frac{\gamma}{\epsilon-1}$. Since I do not have this data I do not attempt to split β and θ, but my estimation only requires the ratio $\frac{\beta}{\theta}$.

[62] By comparing across these sectors, I ignore the effect on output competition.

also reduces the markup-labor covariance.[63] This reflects the fact that there is a reallocation of labor from high-markup to low-markup firms relative to the case where all firms choose a constant markup. It is the mechanism through which exchange rate shocks affect the allocative efficiency of the economy. The opposite effect is true for an exchange rate depreciation, where Figure 5 reports that the markup-labor covariance and allocative efficiency both increase.

Figure 5: Effects of a Cost Shock on Allocative Efficiency



Panel A reports the allocative efficiency statistic (normalized to 1 when $\tau = 1$) as τ varies between 0.8 and 1.2. Panel B reports the markup-labor covariance, once again normalized to one. These are calculated using the parameter estimates from Table 3: (θ, κ, σ) = (0.24,7.63,2.93).

The counterfactual can be compared to the reduced form regressions that measure the changes in allocative efficiency in response to changes in the REER. In this simplified model, on average each 1% appreciation generates on average a reduction in allocative efficiency of 0.6 percentage points in industries where the import share is around 0.35. This can be interpreted relative to an industry where no firms import, so that a change in τ has no effect. The *upper bound* of the reduced form results in Table 2 (column (1)) suggest that each 1% increase in the growth of the real exchange rate reduce allocative efficiency by 0.96% in industries where net exposure is -0.35 (2.75*0.35). The model therefore captures around 61% of the change in allocative efficiency implied by the reduced form analysis (and 87% using the lower estimate in column (2) of Table 2). There are a few possible reasons for the discrepancy. First, the reduced form results could be capturing further effects on allocative efficiency not included in the model. To show the mechanisms transparently, I simplify the model such that all firms import and one parameter maps productivity to

[63] Average markups are 12.9% in the benchmark estimation, and rise to 14.2% when $\tau = 1.1$. I do not target the level of markups in the data. These are slightly lower than in the data, but within the range typically found in the literature. A lower σ or κ would generate higher markups.

import shares. Second, translating the empirical result to an industry with a net exposure of -0.35 might hide non-linearities in terms of changes in allocative efficiency as the import exposure increases. Overall, it is reassuring that the model-simulated effect on misallocation is in the quantitative vicinity of the reduced form estimate.

5 Conclusion

This study examines how changes in allocative efficiency can be a result of shocks that alter the relative price of foreign and domestic goods. The distortion that keeps the market economy away from allocative efficiency is the heterogeneity in market power. In this framework, exchange rate shocks are passed through to prices differentially across firms and I show this effect can have important welfare effects using the case of Chile. Chile experiences large exchange rate variations, partly due to demand shocks for its commodities. In response, exposed firms respond by varying their markup and this generates a reduction in markup dispersion during the depreciation period and an increase in dispersion during the appreciation period. In this context, the mechanism I find most compelling is incomplete pass-through of revenue productivity gains that are heterogeneous across the firm distribution within an industry. Changes in misallocation suggest that the assumption of preferences that yield constant markup would result in a mismeasurement of how reallocation impacts real revenue growth.

Chile can be characterized as an exporter of natural resources, especially copper, and importer of intermediate goods. It is therefore not surprising that there is a significant benefit for Chilean firms in terms of cheaper imported inputs. On the other hand, it is not clear how much its domestic producers are affected by an increase in global competition. Other countries could have a very different composition of exports and imports. They might import mostly final goods and export goods higher up in the vertical specialization ladder. This would mean that trade liberalization can have a more dramatic effect in terms of increasing competition in the manufacturing sector, as is convincingly shown in Feenstra and Weinstein (2017). Future research should consider the importance in the composition of imports and exports to how domestic firms respond to global shocks.

References

Ackerberg, Daniel, Kevin Caves, and Garth Frazer, "Identification Properties of Recent Production Function Estimators," *Econometrica*, 2015, *83* (6).

Amiti, Mary and Jozef Konings, "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia," *American Economic Review*, 2007, *97* (5), 1611–1638.

_ , _ , and Oleg Itskhoki, "Imports, Exports, and Exchange Rate Disconnect," *American Economic Review*, 2014, *104* (7), 1942–78.

Antras, Pol, Teresa C. Fort, and Felix Tintelnot, "The Margins of Global Sourcing: Theory and Evidence from U.S. Firms," *American Economic Review*, 2017, *107* (9), 2514–2564.

Arkolakis, Costas, Arnaud Costinot, Dave Donaldson, and Andres Rodrigues-Clare, "The Elusive Pro-Competitive Effects of Trade," *Review of Economic Studies*, 2019, *forthcoming.*

Baqaee, David Rezza and Emmanuel Farhi, "Productivity and Misallocation in General Equilibrium.," *National Bureau of Economic Research Working Paper*, 2017.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta, "Cross-Country Differences in Productivity: the Role of Allocation and Selection," *American Economic Review*, 2013, *103* (1), 305–334.

Basu, Susanto and John Fernald, "Aggregate Productivity and Aggregate Technology," *European Economic Review*, 2002, *46*, 963–991.

Behrens, Kristian, Giordano Mion, Yasusada Murata, and Jens Suedekum, "Trade, Wages, and Productivity," *International Economic Review*, 2014, *55* (4), 13051348.

_ , _ , _ , and _ , "Distorted Monopolistic Competition," *DICE DISCUSSION PAPER No. 237*, 2016.

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

Berthelon, Matias, "Chilean Export Performance: The Role of Intensive and Extensive Margins," Working Paper Series 615, Banco Central de Chile 2011.

Bertoletti, Paolo, Eilen Fumagalli, and Chiara Poletti, "Price-increasing Monopolistic Competition? The case of IES preferences," working paper no. 15, IEFE 2008.

_ , Federico Etro, and Ina Simonovska, "International Trade with Indirect Additivity," *American Economic Journal - Micro*, 2018, *10*, 1–57.

Blaum, Joaquin, Claire Lelarge, and Michael Peters, "The Gains from Input Trade With Heterogeneous Importers," *AEJ: Macroeconomics*, 2018, *10*, 77–127.

Cavallari, Lilia and Federico Etro, "Demand, Markups, and the Business Cycle," Working Paper No. 09/WP/2017, Ca'Foscari University of Venice 2017.

Chamberlin, E.H., *The Theory of Monopolistic Competition*, Cambridge: Harvard University Press, 1933.

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

DeLoecker, Jan and Frederic Warzynski, "Markups and Firm-Level Export Status," *American Economic Review*, 2012, *102* (6), 2437–2471.

_ and Pinelopi Koujianou Goldberg, "Firm Performance in a Global Market," 2014, *6*, 201–227.

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

Demidova, Svetlana, "Trade Policies, Firm Heterogeneity, and Variable Markups," *Journal of International Economics*, 2017, *108*, 260–273.

Desormeaux, Jorge, Pablo Garca, and Claudio Soto, "Terms of trade, commodity prices and inflation dynamics in Chile," in Bank for International Settlements, ed., *Monetary policy and the measurement of inflation: prices, wages and expectations*, Vol. 49 of *BIS Papers chapters*, Bank for International Settlements, Autumn 2010, pp. 115–128.

Dhingra, Swati and John Morrow, "Monopolistic Competition and Optimum Product Diversity Under Firm Heterogeneity," *Journal of Political Economy*, 2019, *forthcoming.*

Dixit, Avinash K and Joseph E Stiglitz, "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, June 1977, *67* (3), 297–308.

Edmond, Chris, Virgiliu Midrigan, and Daniel Xu, "Competition, Markups and the Gains from International Trade," *American Economic Review*, 2015, *105* (10), 3183–3221.

_ , _ , and Daniel Yi Xu, "How Costly Are Markups?," *National Bureau of Economic Research Working Paper*, 2018.

Ekholm, Karolina, Andreas Moxnes, and Karen Ulltveit-Moe, "Manufacturing Restructuring and the Role of Real Exchange Rate Shocks," *Journal of International Economics*, 2012, *86*, 101–117.

Epifani, Paolo and Gino Gancia, "Trade, markup heterogeneity and misallocations," *Journal of International Economics*, 2011, *83*, 1–13.

Feenstra, Robert and Hiau Looi Kee, "Export variety and country productivity: Estimating the monopolistic competition model with endogenous productivity," *Journal of International Economics*, 2008, *74* (2), 500–518.

Feenstra, Robert C. and David E. Weinstein, "Globalization, Markups, and the U.S. Price Level," *Journal of Political Economy*, 2017, *125* (4), 1041–1074.

_ , Robert E. Lipsey, Haiyan Deng, Alyson C. Ma, and Hengyong Mo, "World Trade Flows: 1962-2000," Working Paper 11040, National Bureau of Economic Research January 2005.

Goldberg, Pinelopi Koujianou, Amit Kumar Khandelwal, Nina Pavcnik, and Petia Topalova, "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India," *The Quarterly Journal of Economics*, November 2010, *125* (4), 1727–1767.

Gopinath, Gina and Brent Neiman, "Trade Adjustment and Productivity in Large Crises," *American Economic Review*, 2014, *104* (3), 793–831.

Halpern, Laszlo, Miklos Koren, and Adam Szeidl, "Imported Inputs and Productivity," *American Economic Review*, 2015, *105* (12), 3660–3703.

Head, Keith, Thierry Mayer, and Mathias Thoenig, "Welfare and Trade without Pareto," *American Economic Review*, May 2014, *104* (5), 310–16.

Holmes, Thomas J., Wen-Tai Hsu, and Sanghoon Lee, "Allocative Efficiency, Markups, and the Welfare Gains from Trade," *Journal of International Economics*, 2014, *94* (2), 195–206.

Hsieh, Chang-Tai and Pete Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

Jung, Jae Wook, Ina Simonovska, and Ariel Weinberger, "Exporter Heterogeneity and Price Discrimination: A Quantitative View," *Journal of International Economics*, January 2019, *116*, 103–124.

Kasahara, Hiroyuki and Beverly Lapham, "Productivity and the decision to import and export: Theory and evidence," *Journal of International Economics*, 2013, *89* (2), 297–316.

_ **and Joel Rodrigue**, "Does the use of imported intermediates increase productivity? Plant-level evidence," *Journal of Development Economics*, 2008, *87* (1), 106–118.

Krugman, Paul R., "Increasing Returns, Monopolistic Competition, and International Trade," *Journal of International Economics*, 1979, *9* (4), 469–479.

Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economics Studies*, 2003, *70* (2), 317–342.

Lu, Dan, Asier Mariscal, and Luis-Fernando Mejia, "How Firms Accumulate Inputs: Evidence from Import Switching," *Working Paper*, 2016.

Macedoni, Luca and Ariel Weinberger, "Quality Heterogeneity and Misallocation: The Welfare Benefits of Raising Your Standards," *Working Paper*, 2019.

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

_ **and Giancarlo Ottaviano**, "Market Size, Trade and Productivity," *Review of Economic Studies*, 2008, *75*, 295–316.

Mrazova, Monika and Peter J. Neary, "Selection Effects with Heterogeneous Firms," *Journal of the European Economic Association*, 2018.

_ **,** _ **, and Mathieu Parenti**, "Sales and Markup Dispersion: Theory and Empirics," Technical Report 2017.

Olley, G. and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1298.

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

Peters, Michael, "Heterogeneous Markups, Growth, and Endogenous Misallocation," Technical Report 2018.

Petrin, Amil and James Levinsohn, "Measure Aggregate Productivity Growth using Plant-Level Evidence," *The RAND Journal of Economics*, 2012, *43* (4), 705–725.

Restuccia, Diego and Richard Rogerson, "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, 2008, *11* (4), 707–720.

Simonovska, Ina, "Income Differences and the Prices of Tradables," *Review of Economic Studies*, 2015, *82* (4), 1612–1656.

_ **and Michael Waugh**, "Trade Models, Trade Elasticities, and the Gains from Trade," *NBER Working Paper No. 20495*, 2014, (20495).

Vives, Xavier, *Oligopoly Pricing: Old Ideas and New Tools*, Vol. 1 of *MIT Press Books*, The MIT Press, 2001.

Weinberger, Ariel, "Importing, Productivity and the Exchange Rate," mimeo 2015.

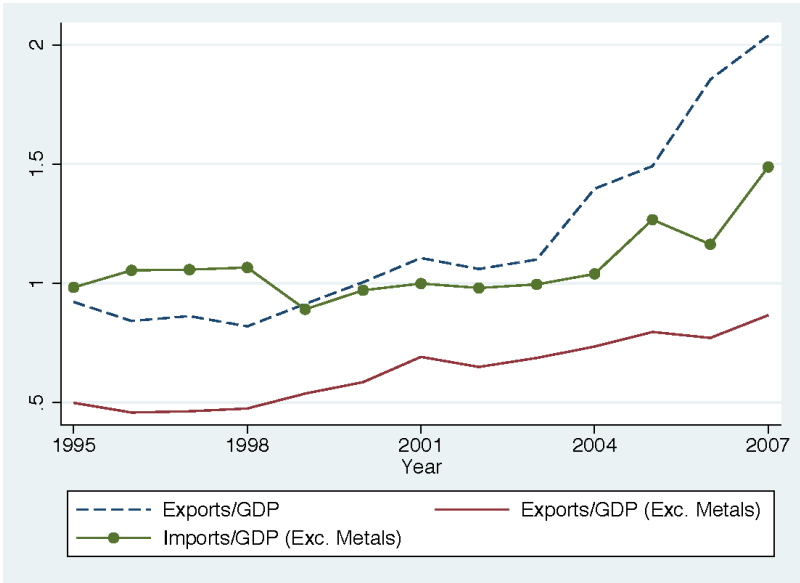
Zhelobodko, Evgeny, Sergey Kokovin, Mathieu Parenti, and JacquesFranois Thisse, "Monopolistic Competition: Beyond the Constant Elasticity of Substitution," *Econometrica*, 2012, *80* (6), 2765–2784.

Appendices

A Open Economy Data and Firm Level Responses

A.1 Open Economy Trends

Figure 6: Exports and Imports as a share of GDP, 1995-2007



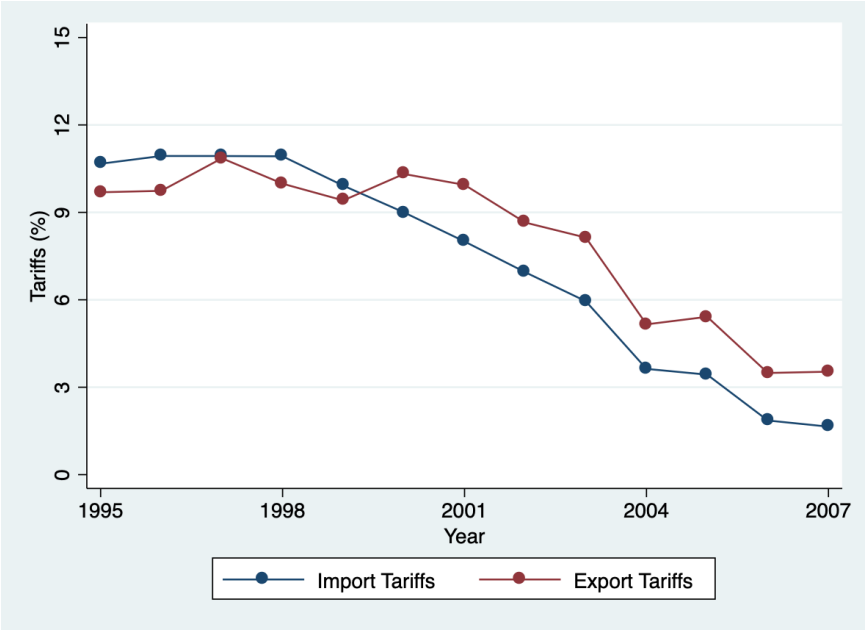
Sources: Trade data from Feenstra et al. (2005), and manufacturing GDP from Banco Central de Chile. Manufacturing GDP and manufacturing exports/imports are both in thousands of current US dollars.

A.2 Aggregate Data Definitions

Here I describe the measures constructed from the right hand side of Equation 10. $\tilde{R}$ equivalent to the Aggregate Productivity Growth (APG) that is used in Petrin and Levinsohn (2012) and Basu and Fernald (2002), which tracks welfare without taking into account variety. In words, $\tilde{R}$ is the sum of deflated value added, subtracting out the growth in inputs. $\Delta ln(\tilde{R}_t) = \Delta ln(Y_t) - \Delta ln(L_t)$, where Y_t (sum of deflated value added) is real revenue if all production income goes towards final demand. $\Delta ln(L_t)$ corrects for changes in expenditure on labor (wage growth in the data) so that the APG measure is not driven by differential wage trends across sectors or labor reallocating across sectors.

Measurement of Y_t ("Final Demand"): At the firm (i) level, $Y_i = Q_i - \sum_j X_{ji}$, where X_{ji} are inputs sourced from some firm, j. By the National Accounting Identity, aggregate final demand is equal to aggregate value added: $\sum_i P_i Y_i = \sum_i VA_i = \sum_i P_i Q_i - \sum_i \sum_j P_{ij} X_{ji}$.

Figure 7: Average Applied Import and Export Tariffs 1995-2007



Source: Comtrade Database, downloaded from World Integrated Trade Solution (WITS). Bilateral tariffs are aggregated to 4-digit level using an unweighted average of 6-digit tariff lines, and then weighted by trade shares to get an average applied tariff rate across all trade partners. Export tariffs are an average of tariffs charged by all importers of Chilean goods.

Information on the construction of aggregate price indices can be found at: `http://www.ine.cl/canales/chile_estadistico/estadisticas_economicas/industria/enia/pdf/deflactor_dos_completo_07_09.pdf` (Note: This is in Spanish). The index is calculated using a Laspeyres index and is aggregated to the 4 digit ISIC using data on 7-digit products. Deflators are constructed for both output and input prices, so that the value added is "double deflated."

Information on the construction of the quantity index can be found at: `http://www.ine.cl/canales/chile_estadistico/estadisticas_economicas/industria/series_estadisticas/archivos/base2002/manufacturera_metodologico_base_promedio_2002.pdf` The goal as described by the INE is to "measure the evolution of quantities and qualities at the product level by eliminating the influence of prices." They sample a set of firms from 1989-2002 and 2001-2007 (the overlap makes it possible to have a continuing time series of growth rates). Although the sampled firms are not the universe of firms in the census, they do make up about 80% of manufacturing value added. However, it does mean I am not picking up the smallest firms and some new enterers, which is why I only use firms that exist for more than 6 years in the construction of real income (though the results look similar without dropping these firms). As with the price indices, the INE constructs a Laspeyeres index with value of sales as weights at a disaggregated product level and then aggregate

up to the 3 digit level.

A.3 Production Function and Markup Estimation

The production function must follow the following functional form:

$$Y_{it} = F(L_{it}, X_{it}, K_{it}; \beta)exp(\omega_{it})$$

β is the vector of output coefficients, ω_{it} is a firm's (i) productivity at time t, ϵ_{it} the measurement error, and $\{L_{it}, X_{it}\}$ are the set of variable inputs (labor and materials). Given data constraints, Y_{it} is deflated total sales.[64] I take logs and use a Gross Output, Translog production function:

$$y_{it} = \beta_l l_{it} + \beta_{ll} l_{it}^2 + \beta_k k_{it} + \beta_{kk} k_{it}^2 + \beta_x x_{it} + \beta_{xx} x_{it}^2 + \beta_{lk} l_{it} k_{it} + \beta_{lx} l_{it} x_{it} + \beta_{kx} k_{it} x_{it} + \beta_{lkx} l_{it} k_{it} x_{it} + \omega_{it} + \epsilon_{it}$$

l, k, x refer to the logged value of labor, capital and intermediate inputs respectively. I estimate each 2-digit industry separately, using 4-digit industry input and output deflators provided by the Chilean Statistics Institution (INE). Notice that this Translog production specification allows for heterogeneous firm level output coefficients.[65] Importantly, I incorporate exporter and importer dummies into the ACF methodology as state variables to the firms' production decisions. This allows exporters and importers to follow a different production technology, following the strategy of Kasahara and Rodrigue (2008) (they add an importer dummy as a state variable), and DeLoecker and Warzynski (2012) (they use export status similarly). Specifically, in the first step of the ACF procedure for the production function estimation, I add imports and exports into the intermediate input demand function of the firm.[66][67] Furthermore, these dummy variables are used in the estimation of survival probabilities (using a Probit function) that control for non random exit of firms as a determinant of next-period productivity.[68]

[64] Labor is the number of total workers. I combine skilled and unskilled although they can be split up using a subjective classification of labor categories. Capital and materials are both expressed as total deflated value of the input.

[65] Given the production function above, the output elasticity of materials for example is: $\theta_{it}^x = \beta_x + 2\beta_{xx} x_{it} + \beta_{lx} l_{it} + \beta_{kx} k_{it} + \beta_{lkx} l_{it} k_{it}$. βs are constant by sector for all years, however notice that θ_{it}^x depends on firm and year specific input values. Output elasticities are therefore firm and year specific.

[66] For a full account of the 2-step procedure see Olley and Pakes (1996), Levinsohn and Petrin (2003), or Ackerberg et al. (2015).

[67] Or in the Olley and Pakes (1996) framework, the investment demand function. This gets inverted to get a non-parametric function for the unobserved productivity shock.

[68] See Olley and Pakes (1996) for a full discussion about the necessity to account for exit/survival.

Table 4: Factor Coefficients and Markups by 2-digit ISIC Sectors

				(1)		
	Obs	θ_L	θ_K	θ_M	Ret Scale	Median Markup
Food products and beverages	19475	0.218	0.073	0.757	1.048	1.192
Manufacture of textile	3462	0.336	0.083	0.666	1.085	1.206
Wearing apparel	3846	0.349	0.047	0.665	1.062	1.219
Tanning and leather	2095	0.433	0.054	0.657	1.145	1.034
Manufacture of wood	4382	0.240	0.051	0.773	1.064	1.264
Manufacture of paper	1803	0.187	0.089	0.745	1.020	1.358
Publishing, printing	3017	0.285	0.111	0.633	1.029	1.323
Manufacture of chemicals	3740	0.283	0.105	0.667	1.055	1.360
Manufacture of rubber and plastics	4085	0.221	0.072	0.734	1.027	1.352
Other non-metallic mineral products	2837	0.191	0.064	0.802	1.057	1.540
Manufacture of basic metals	1503	0.128	0.139	0.747	1.015	1.412
Fabricated metal products	4760	0.243	0.059	0.675	0.977	1.189
Machinery and equipment	2923	0.508	0.098	0.489	1.095	0.993
Electrical machinery	1199	0.246	0.074	0.682	1.002	1.260
Manufacture of instruments	365	0.178	0.046	0.778	1.002	1.774
Manufacture of motor vehicles	752	0.490	0.091	0.656	1.237	1.529
Manufacture of other transport	595	0.338	0.074	0.603	1.016	1.119
Manufacture of furniture	3229	0.180	0.033	0.812	1.025	1.544

Production function coefficients and median markups calculated using the methods of Ackerberg et al. (2015) and DeLoecker and Warzynski (2012) as described in the text. The production function is estimated with past export and import status (as well as exit probability) as state variables. Robustness analysis has also been done by excluding import and export status from the production function.

I estimate firm level markups from the gap (or "wedge") between the output elasticity of materials (θ^x_{it}) and the cost share of materials (α^x_{it}) in total costs. The only assumption necessary is that firms minimize costs, so that the output elasticity is then set equal to its cost share. Markups could also be estimated using the same gap in the labor input, though labor requires more adjustment costs than materials and is less variable. This would make it a worse measure of markups, but I do compare some results to using the labor "wedge" as well. Specifically, my markup measure, at the firm-time level, is represented by:

$$\frac{1}{1-\mu_{it}} = m_{it} = \frac{\theta^x_{it}}{\alpha^x_{it}} \tag{23}$$

A.4 Robustness Regressions: Firm Responses to Exchange Rate Shock

Table 5: Firm Level: Differential Effect on Markup by Degree of Exposure: Shows All controls

	Markup (Lerner)				Markup (Profit)
	(1)	(2)	(3)	(4)	(5)
REER*NetExposure	-0.087***		-0.108***	-0.128**	-0.061
	(0.030)		(0.026)	(0.062)	(0.050)
REER*ImportShare		0.174***			
		(0.039)			
REER*ExportShare		0.058			
		(0.044)			
Tariff*NetExposure			0.007		
			(0.004)		
TFP	0.111***	0.111***	0.111***		
	(0.004)	(0.004)	(0.004)		
REER*MNC=1	0.002	0.002	0.002	-0.004*	-0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
REER*Kintensity	0.002***	0.002***	0.002***	-0.012**	0.003***
	(0.000)	(0.000)	(0.000)	(0.005)	(0.000)
REER*Skill/Unskilled	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.004)	(0.001)
Fixed Effects	Firm, Year	Firm, Year	Firm, Year	Firm, Year	Firm, Year
R^2	0.77	0.77	0.77	0.25	0.59
N	32212	32212	32212	32212	32212
Avg Markup	0.23	0.23	0.23	0.39	0.61

This table is a duplicate of Table 1, but reports all controls. All regressions include firm and year fixed effects and standard errors clustered at the firm level (in parenthesis). I drop the basic metal industry (ISIC 27). $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table 6: Firm Level: Differential Effect on Markup by Degree of Exposure to Competition: Alternative Specifications

	Markup (Lerner)				
	(1)	(2)	(3)	(Year≥2001)	(Year≤2000)
REER*NetExposure	-0.118***	-0.096***		-0.103***	-0.105***
	(0.032)	(0.031)		(0.031)	(0.036)
REER*ImportShare			0.148***		
			(0.039)		
REER*ExportShare			-0.053		
			(0.049)		
REER*MNC=1	0.003**	0.002	0.003**	-0.001	0.002**
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
TFP	0.067***	0.110***	0.067***	0.087***	0.111***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
REER*Kintensity	0.001***	0.002***	0.001***	0.001***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
REER*Skill/Unskilled	-0.000	0.001	-0.000	-0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
NetExposure				0.407***	0.420**
				(0.138)	(0.166)
Fixed Effects	Firm, Industry-Year	Firm, Region-Year	Firm, Industry-Year	Firm, Industry, Year	Firm, Industry, Year
R^2	0.852	0.772	0.852	0.874	0.818
N	32117	32212	32117	23082	25730

This table conducts alternative specifications of the baseline results. The first three columns are as in Table 1, but with industry-year (columns 1 and 3) and region-year (column 2) fixed effects. The last two columns repeat the benchmark firm and year fixed effects specification but only after 2000 and before 2001. Note that only in these cases I allow the exposure measure to vary, so that it is different in the two time periods. Dependent variable is log markup measured using the procedure outlined in DeLoecker and Warzynski (2012) (DLW). All regressions include standard errors clustered at the firm level (in parenthesis). I drop the basic metal industry (ISIC 27). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 7: Firm Level: Differential Effect on Markup by Varying Exposure to Competition

	Markup (Lerner)			Markup (Profit)	Markup (Lshare)
	(1)	(2)	(3)	(4)	(5)
REER*NetExposure	-0.054**		-0.068***	-0.104**	-0.067*
	(0.023)		(0.021)	(0.051)	(0.040)
REER*ImportShare		0.108***			
		(0.030)			
REER*ExportShare		0.036			
		(0.033)			
Tariff*NetExposure			0.006**		
			(0.003)		
TFP	0.107***	0.110***	0.107***		
	(0.003)	(0.003)	(0.003)		
Fixed Effects	Firm, Year	Firm, Year	Firm, Year	Firm, Year	Firm, Year
R^2	0.79	0.79	0.79	0.30	0.61
N	49711	49711	49701	49711	49711
Avg Markup	0.23	0.23	0.23	0.40	0.61

This table examines the differential markup responses to foreign shocks depending on firm exposure. In comparison to Table 1, net exposure is allowed to vary over time (so the term is also used as a control). Dependent variable for the first 3 columns is the Lerner index, which the price-cost ratio measured using the procedure outlined in DeLoecker and Warzynski (2012) (DLW). TFP measurement also follows DLW. REER, TOT, and output tariffs are in logs. Column (4) uses a profit share measure of the markup: "Markup (Profit)" = $\frac{Sales_{it}-wages_{it}-capitalcosts_{it}-inputscosts_{it}}{Sales_{it}}$. The last column uses a Lerner index of the inverse labor share: "Markup (Lshare)" = μ^{Lshare}. The last two columns *do not* include TFP as a control as they are not constructed using the productivity estimation procedures. All columns include firm and year fixed effects (for industry-year interacted FEs see Appendix). I interact the following firm characteristics with the foreign shock to use as controls: capital intensity, a dummy if the firm is a multinational, the ratio of skilled to unskilled labor. The table only displays the results for the REER interaction. Standard errors are clustered at the firm level. I drop the basic metal industry (ISIC 27). $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table 8: Firm Markups and Characteristics

	Markup (Lerner)		Markup (Profit)		Markup (Lshare)	
	(1)	(2)	(3)	(4)	(5)	(6)
TFP	0.073***	0.075***	0.275***	0.276***	0.341***	0.343***
	(0.003)	(0.003)	(0.005)	(0.005)	(0.007)	(0.007)
Employment	0.021***	0.020***	0.021***	0.021***	0.023***	0.023***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Capital Intensity	0.006***	0.006***	-0.006***	-0.006***	0.024***	0.024***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Multinational	0.008*	0.007*	-0.003	-0.004	0.009	0.008
	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)
Skill/Unskill	-0.005**	-0.004**	-0.002	-0.002	0.010***	0.010***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Import Share		0.086***		0.042***		0.057***
		(0.004)		(0.005)		(0.006)
Fixed Effects	Firm, Industry-Year	Firm, Industry-Year	Firm, Industry-Year	Firm, Industry-Year	Firm, Industry-Year	Firm, Industry-Year
R^2	0.86	0.86	0.73	0.73	0.73	0.73
N	49674	49674	51836	51836	51836	51836
Avg Markup	0.23	0.23	0.41	0.41	0.62	0.62

This table measures the relationship of two markup measurements with other firm characteristics. The benchmark Lerner markups is the outcome in the first two columns. The alternative markups are the outcomes in the last four columns. In all cases I include industry-year and firm fixed effects, so that coefficients can be interpreted as comparing firms with industries and years. Standard errors (in parenthesis) are clustered at the firm level. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

A.5 Net Exposure Variable

In this Appendix I describe the identification assumption used by Ekholm et al. (2012) to relate the firm level "net currency exposure" to firm level outcomes.

Taking into consideration both domestic and export sales, the optimal revenue of a firm i is $r_i = p_i q_i + E p_i^* q_i^*$, where p_i and p_i^* are prices in local currency set at home and abroad, q_i and q_i^* are sold quantities at home and abroad, and E is the nominal exchange rate (domestic currency per unit of foreign currency). Then the real exchange rate is $REER_i = p_i/(Ep_i^*)$. Ekholm et al. (2012) consider a small change in the $REER_i$ holding output constant:

$$\frac{\partial r_i}{\partial REER_i} \frac{REER_i}{r_i} = -\frac{ep_i^* q_i^*}{r_i}. \tag{24}$$

Notice that this elasticity is equal to the firm export share.

Then, they define a firms' costs as $C_i = c_i v_i + E c_i^* v_i^*$, where c_i and c_i^* are prices of domestic and imported inputs, and v_i and v_i^* are quantities of domestic and imported inputs. Then again consider a small change in the real exchange rate holding inputs constant[69]:

$$\frac{\partial C_i}{\partial REER_i} \frac{REER_i}{C_i} = -\frac{E c_i^* v_i^*}{C_i}. \tag{25}$$

This elasticity is equal to the share of inputs in total costs.

Finally, this allows for a relationship between the profits and the net effect of the export share and import share in inputs. The elasticity of profits with respect to the REER is shown to be:

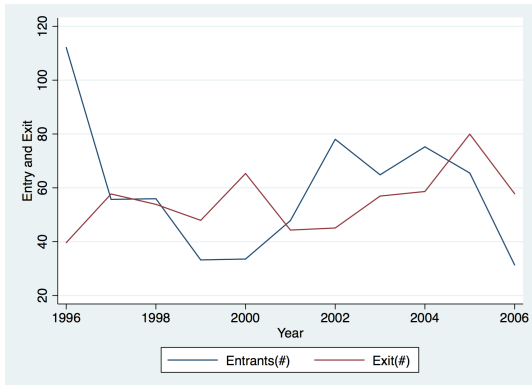
$$\frac{\partial \pi_i}{\partial REER_i} \frac{REER_i}{\pi_i} = -\frac{E p_i^* q_i^*}{r_i} - \frac{\frac{E p_i^* q_i^*}{r_i} - \frac{E c_i^* v_i^*}{C_i}}{\pi_i / r_i}. \tag{26}$$

In my empirical analysis I am interested in how the currency shock affects firm level markups and industry level allocative efficiency. Since markups are directly relative to profits, I make the same identification assumption as Ekholm et al. (2012) that a positive net currency exposure increases the competitive pressure on firms when there is an appreciation shock, while a negative net exposure reduces the competitive pressure.

[69] I ignore the differences between the REER measured by output prices and the REER measure by input prices since I don't have these separately in the data anyways.

A.6 Entry and Exit of Firms into Sample

Figure 8: Entry and Exit (number of firms)



Entry is defined as a firm not in the census in the previous year. Exit is the number of firms not in the census, that were in the census the previous years. Both measures are total number of firms in manufacturing. For context, there are about 5,000 firms in each year of the census.

B Price-Quantity Covariance

This appendix shows the derivation of Equation 9 and establishes the result that it is zero in the case when the sub utility function is CES and the added assumption of Pareto distribution of marginal costs.

First, I start by decomposing revenue using the definition of the covariance. The aggregate revenue equation can be manipulated in the following way:

$$R = NL \int_0^{c_d} q(c)h_d(c)dc \int_0^{c_d} p(c)h_d(c)dc + NL\left[\text{Cov}(p,q)\right], \tag{27}$$

where I use the definition of the covariance: $\text{Cov}(p,q) = \int_0^{c_d}(p(q(c)) - P)(q(c) - Q)h_d(c)dc$. The last term is a residual that represents the deviation of aggregate revenue from physical production. Equation 27 can be further expanded substituting for $\tilde{R}$ and Q, and then taking logs to get growth rates:

$$\begin{aligned} \frac{\tilde{R}}{Q} &= 1 + \frac{\text{Cov}(p,q)}{\int_0^{c_d} p(q(c))h_d(c)dc \int_0^{c_d} q(c)h_d(c)dc} \\ \Delta ln\left(\frac{\tilde{R}}{Q}\right) &\approx \Delta\left(\frac{\text{Cov}(p,q)}{\int_0^{c_d} p(q(c))h_d(c)dc \int_0^{c_d} q(c)h_d(c)dc}\right) \end{aligned} \tag{28}$$

The last line uses the approximation that $ln(1+x) \approx x$. In the main text, I rewrite the covariance using markups and labor, which is equivalent to the equation above.

Next, I show that (28) is zero in the case where preferences are CES and costs are drawn from a Pareto distribution. This is the first part of the proof necessary for Proposition 1. Using the definition of the covariance above:

$$\Delta\left(\frac{\text{Cov}(p,q)}{\int_0^{c_d} p(q(c))h_d(c)dc \int_0^{c_d} q(c)h_d(c)dc}\right) = \Delta\left(\frac{\int_0^{c_d} p(q(c))q(c)h_d(c)dc}{\int_0^{c_d} p(q(c))h_d(c)dc \int_0^{c_d} q(c)h_d(c)dc} - 1\right) \quad (29)$$

When preferences are CES, $p(c) = \frac{1}{1-\mu}c$ with μ constant, and $q(c) = c^{-\sigma}\left(\frac{1}{1-\mu}\right)^{-\sigma}\left(\frac{R}{\tilde{P}}\right)$ with $\tilde{P}$ the aggregate "ideal" price index and R the aggregate revenue. Additionally, $h_d(c)dc = \frac{g(c)}{G(c_d)} = \theta c^{\theta-1}c_d^{-\theta}$. Thus I can input all this information into Equation 29 and reduce the numerator and denominator separately:

$$\begin{aligned}
\int_0^{c_d} p(q(c))q(c)h_d(c)dc &= \left(\frac{R}{\tilde{P}}\right)\frac{1}{1-\mu}\left(\frac{1}{1-\mu}\right)^{-\sigma}\int_0^{c_d} cc^{-\sigma}\theta c^{\theta-1}c_d^{-\theta}dc \\
&= \left(\frac{R}{\tilde{P}}\right)\left(\frac{1}{1-\mu}\right)^{1-\sigma}\left(\frac{\theta}{\theta-\sigma+1}\right)c_d^{1-\sigma} \qquad (30)\\
\int_0^{c_d} p(q(c))h_d(c)dc &= \frac{1}{1-\mu}\int_0^{c_d} c\theta c^{\theta-1}c_d^{-\theta}dc \\
&= \frac{1}{1-\mu}\frac{\theta}{\theta+1}c_d \qquad (31)\\
\int_0^{c_d} q(c)h_d(c)dc &= \left(\frac{R}{\tilde{P}}\right)\left(\frac{1}{1-\mu}\right)^{-\sigma}\int_0^{c_d} c^{-\sigma}\theta c^{\theta-1}c_d^{-\theta} \\
&= \left(\frac{R}{\tilde{P}}\right)\left(\frac{1}{1-\mu}\right)^{-\sigma}\left(\frac{\theta}{\theta-\sigma}\right)c_d^{-\sigma} \qquad (32)
\end{aligned}$$

Next, combining the three above terms into Equation 29:

$$\Delta\left(\frac{\int_0^{c_d} p(q(c))q(c)h_d(c)dc}{\int_0^{c_d} p(q(c))h_d(c)dc \int_0^{c_d} q(c)h_d(c)dc} - 1\right) = \Delta\left(\frac{(\theta+1)(\theta-\sigma)}{\theta(\theta-\sigma+1)}\right) \quad (33)$$

where the term inside the parenthesis on the RHS is constant. Therefore, under the case of CES sub utility and Pareto $G(c)$, the terms in Equation 9 are zero.

C Changes in Allocative Efficiency as a Response to an Exchange Rate Shock: Empirics and Structural Estimation

C.1 Robustness: Empirics

Table 9: Industry Responses to Exchange Rate Shock: Varying Exposure Shares)

	Δ AE			Δ Q	Δ Cov(markup,inputs)	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ REER*Net Exposure	5.045***	4.021***		-2.468*	4.624***	
	(1.497)	(1.200)		(1.269)	(0.587)	
Δ REER*Imported Share			-3.426			-4.497***
			(2.454)			(1.075)
Δ REER*Exported Share			5.334			4.318
			(6.645)			(2.928)
Net Exposure	-1.291*	-1.165**		0.826	-0.082	
	(0.629)	(0.527)		(0.635)	(0.607)	
Imported Share			0.939			-0.171
			(0.726)			(0.751)
Exported Share			-2.020*			-0.687
			(1.090)			(0.453)
Δ Tariff	0.040	0.039	0.038	0.016	-0.008	-0.009
	(0.030)	(0.029)	(0.028)	(0.046)	(0.013)	(0.014)
Δ HHI Index		1.911***	1.920***	0.365***	0.157	0.163
		(0.548)	(0.559)	(0.092)	(0.126)	(0.134)
Fixed Effects	Year,Sector	Year,Sector	Year,Sector	Year,Sector	Year,Sector	Year,Sector
R^2	0.255	0.337	0.339	0.350	0.210	0.212
N	192	192	192	192	204	204

This Table reports the response in aggregate industry measures in response to REER shocks as in Table 2 but with varying exposure shares. Dependent variables are ΔAE (columns (1)-(3)), ΔQ (column (4)), and Δ Cov(markup,inputs) (columns (5)-(6)). Between columns (1-2) and (3), and between (5) and (6), the difference is in the measure of exposure: "Net" versus splitting up export share and import share. These are all at the 2-digit ISIC level. For the covariance I use first differences, while the first 4 columns are growth rates. $\Delta REER$ is a one year growth rate. Net Exposure is defined as (Export Sales/Total Sales)-(Imported Inputs/Total material input costs). The prior two components are "Exported Share" and "Imported Share." These shares are fixed. Industry controls are changes in: the Herfindahl Index of concentration and industry output tariffs. All regressions include sector and year fixed effects and standard errors clustered at the 2-digit industry level (in parenthesis). I drop the basic metal industry (ISIC 27). $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table 10: Industry Responses to Terms of Trade Shock with Fixed Exposure Shares

	Δ AE			Δ Q	Δ Cov(markup,inputs)	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ TOT*Net Exposure	1.475	0.926		0.269	2.143***	
	(1.012)	(0.800)		(0.192)	(0.400)	
Δ TOT*Imported Share			-2.448**			-1.895***
			(0.859)			(0.447)
Δ TOT*Exported Share			-5.064			3.808
			(3.692)			(2.415)
Δ Tariff	0.026	0.027	0.021	0.023	-0.020	-0.019
	(0.035)	(0.031)	(0.033)	(0.044)	(0.017)	(0.018)
Δ HHI Index		1.918***				
		(0.560)				
Fixed Effects	Year,Sector	Year,Sector	Year,Sector	Year,Sector	Year,Sector	Year,Sector
R^2	0.247	0.328	0.259	0.302	0.225	0.227
N	192	192	192	192	204	204

This Table reports the response in aggregate industry measures in response to TOT shocks instead of REER shocks. Dependent variables are ΔAE (columns (1)-(3)), ΔQ (column (4)), and Δ Cov(markup,inputs) (columns (5)-(6)). Between columns (1-2) and (3), and between (5) and (6), the difference is in the measure of exposure: "Net" versus splitting up export share and import share. These are all at the 2-digit ISIC level. For the covariance I use first differences, while the first 4 columns are growth rates. $\Delta REER$ is a one year growth rate. Net Exposure is defined as (Export Sales/Total Sales)-(Imported Inputs/Total material input costs). The prior two components are "Exported Share" and "Imported Share." These shares are fixed. Industry controls are changes in: the Herfindahl Index of concentration and industry output tariffs. All regressions include sector and year fixed effects and standard errors clustered at the 2-digit industry level (in parenthesis). I drop the basic metal industry (ISIC 27). $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table 11: Industry Responses to Exchange Rate Shock: Fixed Importer Dummy (Average across Sector)

	Δ AE	Δ Q	Δ Cov(markup,inputs)
	(1)	(2)	(3)
Δ REER*Importer	-3.818	-0.377	-3.546***
	(2.270)	(1.388)	(0.934)
Δ REER*Exporter	-0.011	-6.098*	6.021***
	(5.415)	(3.079)	(1.162)
Δ Tariff	0.033	0.020	-0.010
	(0.032)	(0.047)	(0.013)
Δ HHI Index			0.201
			(0.127)
Fixed Effects	Year,Sector	Year,Sector	Year,Sector
R^2	0.242	0.316	0.206
N	192	192	204

This Table repeats the specifications from Table 2, but with fixed import and export dummies. The dummies are at the firm level, so I take an average for the sector measure. All regressions include sector and year fixed effects and standard errors clustered at the 2-digit industry level (in parenthesis). I drop the basic metal industry (ISIC 27). $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

C.2 Importing with VES Preferences

In this Appendix I provide details of the specific VES model with stone-geary preferences. Here, I show the case with importing (adding τ to the closed economy case), since the closed economy case is simply $\tau = 1$. First, using the assumptions on unit costs and the mapping from the cost draw to import shares, I can write unit costs:

$$u(c) = wc^{\frac{\beta}{\theta}+1}\tau^{-\beta} \tag{34}$$

I set $w = 1$ as domestic inputs remain unchanged in the model.

From the first order conditions, we can use the inverse demand function to get:

$$\pi(c) = \frac{L}{\lambda}\left(\frac{q(c)}{q(c)+\bar{q}}\right) - Lq(c)w\tau^{-\beta}c^{\frac{\beta}{\theta}+1}, \tag{35}$$

where the aggregate environment is taken as given for each firm and summarized by $\lambda = \int_0^{c^d} \frac{q(c)}{q(c)+\bar{q}} dG(c)$. A useful feature of this framework is that it will not be necessary to solve for the general equilibrium term, λ, which will become clearer below.

Firms choose $q(c)$ to maximize profits, which yields the following solution:

$$q(c) = L\bar{q}\left[\left(\frac{1}{\lambda w\bar{q}}\tau^{\beta}c^{-(\frac{\beta}{\theta}+1)}\right)^{\frac{1}{2}} - 1\right] \tag{36}$$

The key then is to set quantity equal to zero, which yields the cutoff cost draw with positive demand. Then, price, quantity, profits, and revenues for a firm can all be written strictly as a function of their cost relative to the cutoff:

$$c^* = (\lambda w\bar{q})^{\frac{-\theta}{\beta+\theta}}\,\tau^{\frac{\beta\theta}{\beta+\theta}} \tag{37}$$

$$p(c) = wc^{\frac{\beta}{\theta}+1}\tau^{-\beta}\left[\underbrace{(c^*c)^{\frac{1}{2}(\frac{\beta}{\theta}+1)}}_{\text{markup}}\right] \tag{38}$$

$$q(c) = L\bar{q}\left[\left(\frac{c^*}{c}\right)^{\frac{1}{2}(\frac{\beta}{\theta}+1)} - 1\right] \tag{39}$$

$$\pi(c) = L\bar{q}w\tau^{-\beta}\left[(c^*)^{(\frac{\beta}{\theta}+1)} - 2\,(c^*c)^{\frac{1}{2}(\frac{\beta}{\theta}+1)} + c^{(\frac{\beta}{\theta}+1)}\right] \tag{40}$$

$$r(c) = L\bar{q}w\tau^{-\beta}\left[(c^*)^{(\frac{\beta}{\theta}+1)} - (c^*c)^{\frac{1}{2}(\frac{\beta}{\theta}+1)}\right] \tag{41}$$

Notice that plugging in the solution for c^*, one can get, for example, the markup: $m(c) = (\lambda w\bar{q})^{-1/2}\tau^{\frac{1}{2}\beta}$. Thus, $\frac{\partial m(c)}{\partial \tau} = \frac{1}{2}\beta\tau^{\frac{-1}{2}}(\lambda w\bar{q})^{-1/2} > 0$.

Then, the aggregates are computed by assuming a Pareto distribution of costs, where κ the shape parameter of the Pareto cost distribution For now, let $\tau = 1$. The set of producing firms is in the range $(0, c^*)$. I therefore integrate across producing firms, and write aggregate production as a function of this new general equilibrium cutoff.

$$
\begin{aligned}
R &= N \int_0^{c^*} r(c) c^{\kappa-1} (c^*)^{-\kappa} \kappa dc \\
&= N L \bar{q} w \tau^{-\beta} \left[(c^*)^{(\frac{\beta}{\theta}+1)} - \frac{\kappa}{\frac{1}{2}(\frac{\beta}{\theta}+1)+\kappa} (c^*)^{(\frac{\beta}{\theta}+1)} \right] \qquad (42) \\
Q &= N L \bar{q} \left[\frac{\kappa}{\kappa - \frac{1}{2}(\frac{\beta}{\theta}+1)} (c^*)^{\frac{1}{2}(\frac{\beta}{\theta}+1)} - (c^*)^{\frac{1}{2}(\frac{\beta}{\theta}+1)} \right] \qquad (43) \\
P &= w \tau^{-\beta} \left[\frac{\kappa}{\frac{1}{2}(\frac{\beta}{\theta}+1)+\kappa} (c^*)^{(\frac{\beta}{\theta}+1)} \right]. \qquad (44)
\end{aligned}
$$

N is the total number of producing firms, which depends on entry and the Pareto shape and scale parameters, but notice it cancels out when we take the ratio of aggregate revenue and quantity. Then, one can write the ratio for aggregate markups, which determines allocative efficiency:

$$
\frac{\tilde{R}}{Q} = (c^*)^{-\frac{1}{2}(\frac{\beta}{\theta}+1)} \left[\frac{1 - \frac{\kappa}{\frac{1}{2}(\frac{\beta}{\theta}+1)+\kappa}}{\frac{\kappa}{\frac{1}{2}(\frac{\beta}{\theta}+1)+\kappa} \left(\frac{\kappa}{\kappa - \frac{1}{2}(\frac{\beta}{\theta}+1)} - 1 \right)} \right] \qquad (45)
$$

An interesting case is $\kappa \to \infty$, as would be the case without firm heterogeneity: in that case, $\frac{\tilde{R}}{Q} = 1$. The reason is that the misallocation I capture in this paper is due solely to firm heterogeneity, as the Pareto assumption guarantees that entry is optimal.

Finally, notice that since $c^* = (\lambda w \bar{q})^{\frac{-\theta}{\beta+\theta}}$ (when $\tau = 1$), it is a function of a general equilibrium object, λ. However, I use the free entry condition: $\bar{\pi} = f_e$, where f_e is the entry cost parameter. This allows one to solve for λ and plug into c^*, which becomes a function of only parameters. I normalize this $c^*(\tau = 1)$ to 1, and solve for changes in the economy due to changes in $\frac{c^*(\tau)}{c^*(\tau=1)}$, where the only change is in the τ.

C.3 Generalized CES Preferences

The estimation in the main text is conducted assuming Generalized CES (GCES) preferences, to allow for flexibility in the curvature of demand, which will affect average markups in the economy. Calibrating σ allows the model to capture the sales-markup

correlation along with moments from the sales distribution.

First, notice that the behavior of the cutoff cost is identical in GCES preferences. Equation (20) in Jung et al. (2019) shows the cost cutoff with GCES preferences in a model with multiple countries. Since in this paper I examine a one country case, wages are normalized to one and the cutoff is fixed (for the case without τ). Jung et al. (2019) also show that the cutoff cost is proportional across all preferences with directly separable demand, only differentiated by the demand parameters. Hence, the cutoff in the GCES will follow exactly from the case where $\sigma = 1$ above.

For $\sigma \neq 1$, no closed form solution exist for prices and therefore sales. However, one can solve for these numerically, as laid out in Jung et al. (2019). As in the previous subsection, there is a cutoff cost at which demand is zero and given this cutoff cost there is the following implicit function for prices:

$$(1-\sigma)p(c) + \sigma u(c) = p(c)^{\sigma+1}(c^*)^{-\sigma}, \tag{46}$$

where u is the unit cost of the firm (which in this case might differentiate from the cost draw c). Given prices, one case solve for quantity and sales:

$$q(c) = L\bar{q}\left[\left(\frac{c^*}{p(c)}\right)^{\sigma} - 1\right] \tag{47}$$

$$r(c) = L\bar{q}w\tau^{-\beta}\left[(c^*)^{\sigma}\, p(c)^{1-\sigma} - p(c)\right] \tag{48}$$

I then use numerical methods to aggregate these and compute $\tilde{R}$ and Q. Although there is no closed form analogue to (45), Figure 5 shows that allocative efficiency decreases with τ as in the $\sigma = 1$ case. One can also check numerically that $\frac{\tilde{R}}{Q} = 1$ when $\kappa \to \infty$.

D Simulated Method of Moments Estimation

In this section, I describe how to estimate the parameters of the structural model described in Section 4.2. The goal is to estimate the parameter set $(\kappa, \frac{\beta}{\theta}, \sigma)$. I solve the model via simulation because the moments in the model that pin down these parameters are created using simulated firms. In other words, for a guess of the parameters, I simulate firm-level outcomes and attempt to reproduce moments from the data, which includes firm sales, firm imports relative to total intermediate costs, and firm markups. The moments in the data are constructed using the Chilean data described in the main text.

First, I simulate a large enough number of draws so as to best approximate the entire continuum of firms that exist in the model. I follow the application in Jung et al. (2019),

and relabel firm-level indicators that can be simulated from a parameter-free uniform distribution. Recall that the pdf of the cost distribution is given by $h(c) = \frac{\kappa}{c^{-\kappa-1}}$[70]. I draw 500,000 realizations of the uniform distribution on the [0; 1] domain, $U \sim [0; 1]$, and order them in decreasing order, and find the maximum realization, denoted by u_{max}. Then, the firm cost is $c = (u/u_{max})^{1/\kappa} c^*$. I normalize ($c^*$) to one in the case where $\tau = 1$, so that $(w\lambda\bar{q})^{-\frac{\theta}{\beta+\theta}} = 1$ (as explained above). This is without loss of generality, as all moments can be written without these parameters. Then, by construction $c \in [0, 1]$. As τ changes, $c^* = \tau^{\frac{\beta\theta}{\beta+\theta}}$. For changes in τ, I therefore re-compute the set of firms.

I adopt an exact-identification strategy that targets 3 moments: 1) The log sales of the firm in the 99th percentile minus the log sales of the firm in the 90th percentile. Given simulated costs, firm sales are computed from (48), where all parameters except $(\kappa, \frac{\beta}{\theta}, \sigma)$ cancel out by using relative sales. 2) The correlation of markups and sales across the distribution of firms. I use the Lerner index for markups, or $\frac{p-u}{p}$.[71] 3) The average import share of firms. From the structure of the model, the domestic share of each firm is computed as: $s_D = 1/\xi$, where $\xi = (1/u)^{1/\theta+\beta}$.

The theoretical moments are matched to their counterpart in the data in order to identify the model parameters. The moments above can be constructed using the same census of Chilean firms used in the reduced form estimation. I restrict the sample of firms to importers only to match the fact that all simulated firms import by construction. I use data from 1995-2001, as this is arguably when firm imports are most stable, although firm moments are generally stable over time. For firm sales, I normalize by average sales in order to remove the nominal units. Markups are computed to reflect the model: value added relative to wages. I construct firm-level import shares as imported inputs relative to firm material costs, and then calculate the average across all firms.

Let $F_i^m(\kappa, \frac{\beta}{\theta}, \sigma)$ be the vector of model generated moments described above, where i represents each moment, and let F_i^d denote the corresponding value of the empirical moments. Our identification consists of choosing the parameter set that minimizes the sum of the squared errors between empirical and theoretical moments:

$$\min_{\kappa, \frac{\beta}{\theta}, \sigma} \sum_{i=1}^{3} \left(F_i^d - F_i^m(\kappa, \frac{\beta}{\theta}, \sigma) \right)^2 . \tag{49}$$

Finally, I compute bootstrap standard errors by running the estimation above 100 times, each time taking a bootstrap sample of the data. I take the average parameter estimates

[70] Generally, the distribution includes a support parameter that determines the lower bound of the distribution. In this case, I set the support equal to 1 as it would just cancel out in the estimation if included.

[71] Recall that $u = c^{\frac{\beta+\theta}{\theta}} \tau^{-\beta}$.

$(\hat{\kappa}, \frac{\hat{\beta}}{\theta}, \hat{\sigma})$, and use the standard deviation of estimates to compute a 95% confidence interval.

The strategy to estimate the parameter set $(\hat{\kappa}, \frac{\hat{\beta}}{\theta}, \hat{\sigma})$ is based on the separate ways that each parameter is identified within the model. κ governs the shape of the cost distribution, which is proportional to the shape in the sales distribution only in the CES case (Mrazova et al., 2017), which do not apply to this non-homothetic specification. The divergence in the sales and cost distribution is due to the distribution of markups. Since firm markup levels are a function of σ, this parameter affects the mapping from the cost to the sales distribution and is not collinear with κ.[72] Finally, as is clear from the definition of unit costs in (19), $\frac{\beta}{\theta}$ determines the cost advantage and the ability of firms to import. This will translate into the import share chosen by each firm.

E Global Shocks and Reallocation

In this Appendix section, I investigate how separate foreign shocks affect the reallocation of production and the implications for allocative efficiency. My strategy is to fit into a reduced-form approach two separate aggregate shocks that are generally confounded when trade costs are reduced. Globalization can affect firms through either i) their residual demand curve or ii) their marginal cost.[73] An example of the former effect is competition through foreign entry, which I show in Section E.1 reallocates production from less to more productive firms because more productive firms lower their markup relatively more. The latter case can occur through a higher terms of trade or lower input tariffs, and in this case more productive firms are able to increase their markup relatively more which reallocates production to the less efficient firms.

E.1 Global Shocks and Markups

The average productivity/selection responses from the two shocks outlined above have been studied extensively in the canonical trade model. I differentiate how the shocks can increase allocative efficiency ($\Delta AE_j > 0$), or dampen welfare gains by reducing allocative efficiency ($\Delta AE_j < 0$), purely through reallocation. My goal is to have a clear

[72] As is not the case, for example, if preferences were CES and the distribution of costs is Pareto.

[73] Focusing exclusively on output tariffs can confound the two channels since they have opposite effects. Below I outline how each channel affects the markup distribution. Though the two shocks can happen simultaneously, in the empirical section I identify the shock using the firm or industry's expected exposure. For a separate perspective on how output tariffs can be tied to welfare gains from trade in a similar model, see Demidova (2017).

and intuitive demonstration of how a shock manifests itself through the change in relative markups, which can be interpreted as a reallocation that either increases or reduces allocative efficiency.

Consider an import competition shock that occurs with a one-time increase in entry (M_e). More entry implies an increase in the marginal utility of income — taken as given by the firm – and since $p(c) = \frac{u'(q(c))}{\delta}$, prices decrease for all firms. Let $p_i(\delta', c_i)$ represent the price decision of firm i after an entry shock.

Separately, consider a shock that lowers costs of importing inputs.[74] To examine this case, I introduce imported inputs as a source of production with a constant labor requirement. To give the firm marginal cost more structure, for each firm i, let the production of one unit of output require one unit of a domestically produced task at cost: $c_i(\varphi_i) = \frac{a}{\varphi_i}$. a is a constant, and φ_i the firm's draw from a productivity distribution. With trade, firms can also import inputs with the cost of one unit of an imported task equal to $\frac{a(\tau^{\kappa_i}-1)}{\varphi_i}$ with $\tau^{\kappa_i} > 1$. The total marginal cost of production is then $c_i(\tau, \varphi_i, \kappa_i) = \frac{a\tau^{\kappa_i}}{\varphi_i}$, where τ is a scalar in the marginal cost curve that represents the cost of importing inputs and allows for a productivity shock that lowers production cost and raises markups due to incomplete pass-through (as is found in DeLoecker et al. (2016)). $\kappa_i > 1$ and is firm specific to allow the magnitude of the import shock to be heterogeneous across firms.[75] A shock that lowers the cost of imported inputs scales down $a\tau^{\kappa_i}$. The impetus for this mechanism can be a terms of trade gain or lower input tariffs.[76]

Taking both effects into consideration, price is represented by $p_i(\delta, a\tau^{\kappa_i}/\varphi_i)$ and Equation 5 is rewritten to express the markup as:

$$\frac{p_i(\delta, a\tau^{\kappa_i}/\varphi_i)}{a\tau^{\kappa_i}/\varphi_i} = \frac{1}{1-\mu(\delta, \tau^{\kappa_i}, \varphi_i)} = m_i(\delta, \tau^{\kappa_i}, \varphi_i) \tag{50}$$

Markups are a function of one firm primitive and two aggregate variables that identify the domestic environment. κ_i is allowed to vary across firms, but I will compute comparative statics using changes in τ only, while controlling for firm specific effects in the empirical analysis. For the rest of this section, I set $\kappa_i = 1\ \forall i$.[77]

To relate changes in allocative efficiency to reallocation, I concentrate on the second

[74] DeLoecker and Goldberg (2014) differentiate between shocks to the residual demand curve and shocks to the marginal cost curve as responses to output and input tariffs changes respectively.

[75] For example bigger firms might be more sensitive to changes in import prices that small firms.

[76] Since $-1 < \frac{\partial q}{\partial c}\frac{c}{q} < 0$, a reduction in marginal costs will increase the equilibrium individual consumption of each variety and increase its markup.

[77] For example, in the empirical analysis I control for firm size, and interact it with the terms of trade (the τ shock).

case from above which is also the focus of the empirical analysis of Chile. The firm-level responses to an input shock are given by $\frac{\partial m_i(\delta,\tau,\varphi_i)}{\partial \tau}$, and the reallocation effects can be interpreted as $\frac{\partial m_i^2(\delta,\tau,\varphi_i)}{\partial \tau \partial c_i}$. The first comparative static is trivial: the direction of the markup for each firm after the shock. The interpretation for the latter is the firm-specific sensitivity of the markup in response to the shock holding δ constant. The thought experiment is as follows: at a new equilibrium with a new τ, has the markup difference between (the same) two firms increased or decreased? Going back to Equation 50, $\frac{\partial m_i(\delta,\tau,\varphi_i)}{\partial \tau} < 0$, or markups decrease with τ. With the assumption of decreasing demand elasticity made in Section 3.1, it can be shown that $\frac{\partial m_i^2(\delta,\tau,\varphi_i)}{\partial \tau \partial c_i} > 0$.[78] Therefore at lower τ's, there is a *bigger* markup difference between a low cost and a high cost firm meaning that inputs are reallocated relatively to initially low markup firms. Intuitively, more productive firms pass-through relatively more of the cost reductions to markups.[79] At the industry level it shows up as a reduction in allocative efficiency as real revenues grow more slowly when there is a reallocation to low-markup firms.

I also explore a Krugman (1979) type globalization episode with tougher competition but relegate it mostly to the appendix since I cannot explore plausibly exogenous variation in competition in Chile using my data. In this case, at higher levels of competition the markup differences between two firms (which differ in their productivity) get *smaller*. Higher markup firms increase productions relatively more as they move down their demand curve and this shows up as an increase in aggregate allocative efficiency.

The details of these super/sub modularity arguments, and the derivations, are below. First, I summarize some testable predictions of the findings above.

E.2 Testable Predictions

Incorporating the global shocks allows for testable predictions. The main question of interest is how the distinct aggregate shocks, either through a cost shifter or competition, affect aggregate misallocation at the industry level. With the assumptions on demand, reallocation of production can be inferred from the observed markup response and this allows for the channel that links the shocks to aggregate misallocation. There are therefore 2 connected predictions:

Hypothesis 1. *(Firm-level) A "favorable" cost shock in intermediate inputs increases markups due to incomplete pass-through, and also reallocates production to initially low markup firms*

[78] In Mrazova and Neary's terminology, this is equivalent to markups being super-modular with respect to trade costs when demand is "less convex" than CES.

[79] I have also checked that, as expected, CES demand implies both of these derivatives are equal to 0.

because they increase markups by less. Increased import competition reduces markups, and this effect is larger for firms with initially higher markups because production is reallocated to high markup firms.

Hypothesis 2. *(Industry-level) A terms of trade appreciation or decrease in input tariffs (reduction in the cost of imported inputs) reduces allocative efficiency. Increased import competition (through a reduction in output tariffs) increases allocative efficiency.*

This hypothesis is tested by the consistency of the observed firm-level responses with the growth in aggregate allocative efficiency as defined in subsection 3.2. Notice that this differs from the ACDR exercise that explores lower output tariffs and the response from both domestic and foreign firms. Although the predictions should not be surprising given the recent literature on non-homothetic demand, this provides a clear direction for my empirical analysis that builds from the firm level up to aggregate industry results. I use Chilean data to measure growth in allocative efficiency at the 2-digit industry level as well as firm level markups using production function estimation. In the next section, I argue that the real exchange rate shock – interpreted as a cost shock – is consistent at the micro level with the predicted changes in markups and at the macro level with the implied changes in allocative efficiency.

E.3 Markup Heterogeneity: Super/Sub Modularity

E.3.1 Markup Differences in response to more entry

A shock to the residual demand curve can be produced by assuming a higher level of entry, M_e, which will increase the marginal utility of income, δ. $\frac{\partial q_i}{\partial M_e}\frac{M_e}{q_i} < -1$ (equilibrium consumption of each variety decreases), which raises the marginal utility of income. Using the same super/sub-modularity argument as in the main text (and shown in detail below), and once again assuming that demand elasticities decreases with sales, tougher competition not only lowers the average markup but also leads the lower cost firms to decrease their markup more than high cost firms.[80]. In this case we start with $\frac{\partial m_i(\delta, c_i(\tau,\varphi_i))}{\partial \delta} < 0$, so that a competition shock, or an increase in M_e, lowers the markup of each firm. To see the reallocation effects let $c_i = \frac{a\tau}{\varphi_i}$ be constant for each firm as there is no shock to τ. Then, $\frac{\partial m_i^2(\delta, c_i(\tau,\varphi_i))}{\partial \delta \partial c_i} > 0$, which means that at higher levels of competition the markup differences between two firms get *smaller*. Higher markup firms increase productions relatively more as they move down their demand curve.

[80] A result similar to Melitz and Ottaviano (2008)

E.3.2 The sub/super modularity argument

In general, the function $m_i(\delta, \tau, \varphi_i)$ is supermodular in τ and φ_i (for a given δ) if:

$$\Delta_{\varphi_i} m_i(\delta, \tau_1, \varphi_i) \leq \Delta_{\varphi_i} m_i(\delta, \tau_2, \varphi_i) \text{ when } \tau_1 \geq \tau_2 \tag{51}$$
$$\text{where } \Delta_{\varphi_i} m_i(\delta, \tau, \varphi_i) = m_1(\delta, \tau, \varphi_1) - m_2(\delta, \tau, \varphi_2) \text{ for } \varphi_1 \geq \varphi_2$$

In the main text I define $m_i(\delta, \tau, \varphi_i) = \frac{p_i(a\tau/\varphi_i)}{a\tau/\varphi_i}$. Let $a = 1$ in this section. Super-modularity holds when $\frac{\partial m_i^2(\delta,\tau,\varphi_i)}{\partial \tau \partial \varphi_i} > 0$. Therefore the markup difference between two firms differentiated by their productivity/marginal cost gets smaller or larger depending on the change in τ. Below, as in the main text, I use *cost differences* instead of productivity differences to calculate supermodularity but the same intuition holds.

To show the results in the previous subsection, I will examine two unique example of the VES utility system of Dhingra and Morrow (2019). They show in their paper that distortions are determined by two elasticities: the demand elasticity and the elasticity of utility (which determines the social markup). Therefore there are 4 different possibilities: demand elasticity and elasticity of utility both increase in production, both decrease in production, and one increases and the other decreases (2 separate cases). My paper actually eliminates two of these cases by considering only decreasing demand elasticities as that is consistent with the Chilean data (and every other firm data I am aware of).

Therefore, I take one functional form example of each possible case. This is sufficient to show that the results are general for the whole VES class I explore in this paper. Although alternative functional forms for preferences within each case have different implications for equilibrium price, quantity, markup, etc., the changes in markups in response to the shocks must move in the same direction for a unique assumption on the sign of i) change in the elasticity of utility with quantity; and ii) change in demand elasticities with quantity.

Case 1: Stone-Geary preferences (social markup decreases with quantity)

$$u(q) = (q + \alpha)^{\rho}, \tag{52}$$

where $\alpha > 0$. In order to get an analytical solution I follow Simonovska (2015) and take the special case for ρ such that $u(q) = \log(q + \alpha)$. Furthermore, for this example assume that $\alpha = 1$. This example is consistent with a decreasing social markup and decreasing demand elasticity, plus the necessary conditions that $u'(q) > 0$, $u''(q) < 0$, and $\mu(q) < 1$.

First order conditions for the firm satisfy $u'(q) + qu''(q) = \delta\tau c$. After calculating the first

and second derivatives of the utility function allows me to solve for q and the markup:

$$q = (\delta\tau c)^{-1/2} - 1, \text{ where } c \in \left(0, \frac{1}{\delta\tau}\right) \tag{53}$$

$$m = \frac{p}{\tau c} = \frac{u'(q)}{\delta\tau c} = (\delta\tau c)^{-1/2}, \tag{54}$$

Finally, the above expressions allow me to verify that: $\frac{\partial m(\delta,\tau,\varphi)}{\partial\tau} < 0$, $\frac{\partial m(\delta,\tau,\varphi)}{\partial\delta} < 0$, $\frac{\partial m(\delta,\tau,\varphi)^2}{\partial\tau\partial c} > 0$, and $\frac{\partial m(\delta,\tau,\varphi)^2}{\partial\delta\partial c} > 0$.

Case 2: HARA or "bipower" preferences (social markup increases with quantity) The HARA system is the specific utility system explored by Dhingra and Morrow (2019) and is also used in Cavallari and Etro (2017):

$$u(q) = aq^{\rho} + bq^{\alpha}, \tag{55}$$

where $\rho \neq \alpha$, $a < 0$, and $b > 0$ to satisfy the conditions that the social markup increases with quantity and the demand elasticity decreases with quantity. An example that satisfies the necessary restrictions and is easy to work with is: $\rho = 3$ and $\alpha = 1$.

Again, the first order conditions for the firm satisfy $u'(q) + qu''(q) = \delta\tau c$ which allows me to solve for q and the markup:

$$q = \sqrt{\frac{\delta\tau c - b}{9a}}, \text{ where } c \in \left(0, \frac{b}{\delta\tau}\right) \tag{56}$$

$$m = \frac{p}{\tau c} = \frac{1}{3} + \frac{2}{3}\left(\frac{b}{\delta\tau c}\right) \tag{57}$$

Once again this allows me to verify that $\frac{\partial m(\delta,\tau,\varphi)}{\partial\tau} < 0$, $\frac{\partial m(\delta,\tau,\varphi)}{\partial\delta} < 0$, $\frac{\partial m(\delta,\tau,\varphi)^2}{\partial\tau\partial c} > 0$, and $\frac{\partial m(\delta,\tau,\varphi)^2}{\partial\delta\partial c} > 0$.