

Trade Liberalization, the Price of Quality, and Inequality: Evidence from Mexican Store Prices*

Benjamin Faber[†]

15 November 2012

Job Market Paper

Abstract

This paper considers a novel distributional channel of developing country trade liberalization that operates through differences in cost of living inflation between rich and poor households. Motivated by the observed pattern of vertical differentiation across Mexican households in consumption and plants in production, I consider quality choice as a channel that links differences in consumption baskets between the rich and the poor to differences in plant technologies and, thus, relative price changes. Guided by this theoretical framework, I draw on a new collection of microdata covering Mexican households, plants, and stores to estimate this channel empirically in the context of Mexico's trade liberalization under NAFTA. In particular, I present evidence that cheaper access to US inputs reduces the relative price of higher quality products in Mexican cities. In turn, because richer households consume higher product quality, I find that this relative price effect has led to a significant increase in Mexican real income inequality due to NAFTA over the period 1994-2000.

Keywords: Market integration; inequality; product quality

JEL Classification: D63; F13; F15; O24

*I am grateful to Robin Burgess, Dave Donaldson, Henry Overman, Steve Redding and Daniel Sturm for their advice and support on this project. I would like to thank Etienne Gagnon at the Federal Reserve Board in Washington D.C. for allowing me access to the more recent years of the Mexican store price data, and Gabriel Romero Velasco at INEGI in Aguascalientes, Mexico for his tireless support on the Mexican plant data. I am grateful to Pol Antras, David Atkin, Michael Best, Francisco Costa, Arnaud Costinot, Swati Dhingra, Esther Duflo, Jason Garred, Horacio Larreguy, Guy Michaels, John Morrow, Thomas Sampson and Anthony Venables for very helpful conversations. I would also like to thank seminar participants at Harvard, LSE, MIT and UCL for comments and suggestions. This project was supported by research funding from the UK Economic and Social Research Council.

[†]London School of Economics and Centre for Economic Performance; Email: b.s.faber@lse.ac.uk.

1 Introduction

How does globalization affect inequality? The canonical approach to analyze this question is through the lens of Stolper-Samuelson, whereby trade across homogeneous final goods affects the relative returns to domestic factors of production. This paper departs from this tradition in three ways. First, I consider household price indexes in the denominator of real income, rather than nominal incomes in the numerator, as a channel through which trade liberalization can affect inequality. Second, I emphasize access to imported inputs, rather than final consumer goods. Third, I analyze relative price changes across vertically differentiated products within disaggregate product groups, rather than across sectors.

In particular, I consider product quality choice by households in consumption and by plants in production as a channel that links differences in the consumption baskets between the rich and the poor to differences in imported input shares in production. Trade liberalization can thus affect cost of living inflation asymmetrically across the income distribution because product quality differentiation gives rise to a correlation between differences in household expenditure shares in consumption and differences in plant technologies in production. Drawing on a new collection of microdata covering Mexican households, plants, and stores, I test this channel empirically in the context of Mexican trade liberalization under NAFTA.

The paper's main contributions can be summarized as follows. Motivated by the observed pattern of vertical differentiation in Mexican microdata on household consumption and plant production at the beginning of NAFTA in 1994, I propose a model of quality choice in a setting with heterogeneous consumers and producers. Guided by this theoretical framework, I present evidence that cheaper access to US intermediate inputs reduces the relative price of higher quality products in Mexican cities. In turn, because richer households consume higher product quality, I find that this relative price effect has led to a significant increase in Mexican real income inequality due to NAFTA over the period 1994-2000. Finally, I find that the observed pattern of quality sorting also has one more general implication for real income inequality. The estimations suggest that differences in household quality choices translate into differences in weighted average plant productivities, so that the consumption baskets of the poor embody significantly higher quality adjusted prices compared to the rich. This increases real income inequality in the cross-section of households compared to a world under conventional assumptions without quality sorting by households and plants.

The paper's analysis is motivated by three more general observations about inequality and trade. The first of these concerns what we already know about real income inequality. There is pervasive evidence that relative consumer price changes can have significant implications for real income inequality due to differences in cost of living inflation between the rich and the poor (e.g. Muellbauer, 1974; Deaton, 2003; Moretti, 2011; Handbury, 2012).¹ However, when moving from the measurement of inequality to analyzing the effects

¹Rather than focusing on prices over time, Handbury (2012) evaluates relative prices across locations.

of policy or market shocks on inequality, economists have mainly focused on nominal wages while treating differences in cost of living inflation as exogenous.² The second and third observations concern developing country imports. Developing country imports are strongly dominated by intermediate inputs, rather than final consumer goods. The case of Mexico is no exception to this, with roughly 90% of total imports from the US between 1994-2000 accounted for by intermediates. Finally, the majority of the variation in use of imported inputs across the product space appears to be between plants within disaggregate product groups, rather than across sectoral averages.³ Taken together, these three insights suggest i) that relative consumer prices matter for inequality, and ii) that access to imported inputs and price changes across varieties within product groups play a significant role in capturing the consequences of developing country trade liberalization.

The analysis proceeds in several steps. I begin by documenting a set of stylized facts about vertical differentiation in Mexican consumption and production at the beginning of NAFTA in 1994. A meaningful analysis of relative prices, production technologies, and household consumption within consumer product groups requires data on unit values (prices per physical unit), plant characteristics, and household expenditures at a very fine level of product aggregation. I draw on Mexican plant surveys, including rich product line level information, in combination with household consumption surveys, including individual purchase prices and quantities, to document a set of relationships between unit values and plant characteristics in production, and between unit values and household characteristics in consumption. In particular, the Mexican microdata suggest that: 1) Plant product line unit values are increasing in imported input shares in production; 2) plant product line unit values are increasing in product sales; and 3) household purchase unit values are increasing in household income in consumption.

To capture these observed moments in the microdata, I propose a model of quality choice by households in consumption and plants in production. The model serves two main objectives. First, it formalizes a product quality interpretation of the stylized facts. Second, it yields testable predictions on NAFTA's effect on Mexican consumer prices, and guides the estimation of the cost of living implications of these relative price changes across the Mexican income distribution. While several existing theoretical frameworks have been proposed to capture separately either the consumption side or the production side of the stylized facts, the two have so far not been considered in a unified framework of quality choice.⁴ I show that the introduction of heterogeneous household quality evaluations into

²For discussions of the literature on the causes of inequality see for example Leamer (1996), Aghion and Williamson (1998), and Goldin and Katz (2008). Notable exceptions to the focus on nominal incomes are Deaton (1989), Porto (2006), Broda and Romalis (2008), and Cortes (2008). See discussion of related literature at the end of this section.

³Appendix Figures A.1.1 and A.1.2 provide a graphical illustration of these insights.

⁴Existing models have focused either on quality choice across households while abstracting from plant heterogeneity (e.g. Choi et al., 2009; Fajgelbaum et al., 2011; Handbury, 2012), or on quality choice across plants while abstracting from household heterogeneity (e.g. Johnson, 2011; Kugler and Verhoogen, 2011; Feenstra and Romalis, 2012). See the discussion of related literature at the end of this section for further

a model of quality choice by heterogeneous plants poses one key challenge, which is that *physical* product quality is distinct from *perceived* product quality in the market place. The definition of this distinction is of interest more generally because it reveals the concept of product quality that has been implicitly estimated from product prices and market shares in a prominent strand of empirical work in industrial organization (e.g. Berry, 1994; Berry, Levinsohn and Pakes, 1995) and international trade (e.g. Khandelwal, 2010; Hallak and Schott, 2011; Feenstra and Romalis, 2012), when acknowledging that we live in a world with non-homothetic tastes.

The theoretical framework is then used to guide the empirical estimation in the three remaining sections. Section 5 draws on the barcode level store price microdata of the Mexican Consumer Price Index to empirically test the model's predictions on the effect of NAFTA's tariff cuts on Mexican consumer prices. Section 6 estimates the cost of living implications of these relative price effects across the Mexican income distribution by drawing on observable moments in the household consumption microdata. Finally, Section 7 imposes additional parameter assumptions in order to estimate differences in product quality and quality adjusted prices across the income distribution.

In support of the predictions, the store price regressions suggest that products with initially higher unit values experience a stronger reduction in their relative price within product groups that are subject to larger tariff cuts on their intermediate inputs over the period 1994-2000. That is, the relative price of initially more expensive items decreases in product groups that gain cheaper access to US inputs. These results are based on a novel identification strategy to relate import access to domestic outcomes. It is a common concern that tariff changes may be correlated with omitted factors that also affect mean sectoral outcomes.⁵ Focusing instead on relative price changes within disaggregated product groups (e.g. antibiotic pills, electric irons) allows me to rely on the much weaker identifying assumption that tariff cuts are plausibly exogenous at the level of individual barcode product lines, especially in the case of intermediate inputs which are shared throughout the domestic economy. To address potential concerns that tariff cuts may have been targeted at particular segments of the plant distribution within product groups, I also propose an instrumental variable strategy that can be applied more generally in the context of input tariff changes. The instrument is based on the insight that tariff targeting should be of much less concern for a subset of intermediate inputs, such as basic chemicals, which are used widely across domestic output sectors. Finally, I exploit the richness of the collected store price microdata to report three different placebo falsification tests.

The model also makes predictions about the heterogeneity of the relative price effect of input tariff cuts. In particular, the observed effect should be driven by differentiated product groups in which initial price differences provide stronger signals about differences

references.

⁵See for example discussions in Goldberg and Pavcnik (2005; 2007). See Slaughter (1998) for a survey of the empirical literature on trade induced relative price changes across industries.

in quality and plant technologies. To test this prediction empirically, I follow a two stage procedure. In the first stage, I use the model's estimation equation for sectoral scopes of quality differentiation in terms of observable moments in the plant microdata. In the second stage, regression results then confirm that the observed average effect of input tariff cuts on relative store prices is driven by product groups that have been estimated to be differentiated in the Mexican plant microdata. A final prediction concerns the effect of access to imported inputs on the reallocation of market shares towards higher quality product lines. To test this prediction, I draw on detailed monthly listings of product entry and exit in the store price surveys, and present evidence in support of this effect.

To evaluate the consequences of NAFTA's observed store price effects for differences in household cost of living inflation, the model yields a convenient estimation equation in terms of observable moments in the household consumption microdata. I discuss the two key assumptions underlying this expression and outline the empirical strategy to estimate it from the data. I find that the average tariff cut under NAFTA between 1993-2000 (12 percentage points) has led to at least 1.7 percentage points higher cost of living inflation in tradable consumption of the poorest urban income quintile compared to the richest over the six year period 1994-2000. This estimate increases to 2.6 percentage points in what I refer to as the baseline specification, and to 3.9 percentage points in what I refer to as the upper bound estimate. In terms of real income inequality, these effects are equivalent to approximately 25-55% of the total observed increase in nominal income inequality among the same groups of households over the period 1994-2000.⁶

In the final section, I impose a parameter assumption on the elasticity of substitution in demand in order to estimate differences in weighted average quality as well as quality adjusted prices across households. In particular, the model captures the observed pattern in the Mexican microdata which suggest that more productive plants sort into higher quality products. The implication is that differences in household quality choices translate into differences in quality adjusted prices. I estimate that the poorest quintile of urban Mexicans consume 20-50% lower weighted average product quality among differentiated product groups. In turn, I find that these observed consumption differences translate into at least 2-7% lower weighted average plant productivities, and thus higher quality adjusted prices, in tradable consumption of the poorest quintile compared to the richest.

The paper relates and contributes to several strands of literature. It is related to empirical work on trade and inequality in developing countries. A comprehensive review of this literature is given in Goldberg and Pavcnik (2007), and more recent contributions include Verhoogen (2008) and Topalova (2010). The focus of this literature has been on trade induced differences in nominal income growth across skill or income groups. This paper, on the other hand, analyzes a distributional channel of developing country trade liberalization that links changes in the relative price of quality to differences in household cost of living

⁶The comparison is adjusting for the fact that tradable consumption accounts for 54% of total consumption in 1994.

inflation.⁷

There are a number of notable exceptions to the focus on nominal income inequality. Porto (2006) combines scheduled Argentinian tariff changes under Mercosur with household expenditure shares across seven consumption sectors to predict household inflation differences. Broda and Romalis (2008) analyze the link between consumer good imports from China and household inflation differences using homescanner data in the US. Outside the focus on trade, Cortes (2008) analyzes the price index implications of low-skilled immigration in US metropolitan areas, and earlier work by Deaton (1989) predicts the cost of living implications of agricultural price changes using household consumption surveys.⁸ To the best of my knowledge, this paper is the first to i) look at the cost of living implications of relative price changes within product groups in a developing country context, ii) empirically estimate these relative price changes in the context of a major trade liberalization episode, and iii) guide this analysis within a theoretical framework of quality choice by heterogeneous households and plants.

The paper also relates to recent contributions on quality choice in a setting with ex ante heterogeneous firms (Mandel, 2010; Johnson, 2011; Kugler and Verhoogen, 2011; Feenstra and Romalis, 2012; Gervais, 2012). This paper introduces quality choice by heterogeneous households into this setting and draws attention to, and estimates empirically the distributional implications that arise when differences in consumption baskets across the income distribution are linked to differences in plant technologies through quality choice.

Finally, the paper is related to existing literature on non-homotheticity in international trade. Non-homothetic preferences were originally introduced to explain part of the variation of cross-country trade flows left unaccounted for by neoclassical trade theory (Markusen, 1986; Bowen *et al.*, 1987; Trefler, 1995; Matsuyama, 2000; Choi *et al.*, 2009; Fieler, 2011).⁹ Rather than focusing on the consequences for trade flows, this paper analyzes the implications of non-homotheticity for the distributional effects of trade in a developing country.¹⁰

The remainder of the paper is structured as follows. Section 2 describes the background and data. Section 3 documents stylized facts about vertical differentiation in Mexican consumption and production at the beginning of NAFTA in 1994. Section 4 presents the model. Section 5 presents the empirical estimation of NAFTA's effect on Mexican consumer prices. Section 6 presents the estimation of the cost of living implications of these relative price effects. Section 7 presents the estimation of differences in quality and quality adjusted

⁷The paper also relates to recent literature on the effects of access to imported inputs in a developing country context (e.g. Amiti and Konings, 2007; Goldberg *et al.*, 2010; Halpern *et al.*, 2011).

⁸More recent contributions on the measurement of cost of living inflation in developing countries include Deaton and Dupriez (2011) and Li (2012).

⁹Atkin (2010) investigates the implications of regional taste differences for the gains from trade, rather than trade flows.

¹⁰In a recent theoretical contribution, Fajgelbaum *et al.* (2011) consider the implications of non-homotheticity in the context of both across and within country income distributions. While focusing on trade in final goods, their model yields predictions that are in line with the empirical findings presented in this paper, that trade increases inequality in the poorer country through inflation differences across the domestic income distribution.

prices across the Mexican income distribution. Section 8 concludes.

2 Background and Data

2.1 Mexican Trade Liberalization

Mexican trade liberalization began as part of government stabilization efforts in response to the severe economic crisis at the beginning of the 1980s. When Mexico joined the General Agreement on Tariffs and Trade (GATT) in 1986, it initially agreed to bind tariffs at a ceiling of 50 percent. In December of 1987 the government then implemented another major consolidation of its tariff schedule whereby all non-agricultural import tariffs were set at either zero, five, ten, fifteen, or twenty percent (Kate, 1992). Following this first wave of liberalization, the Mexican tariff schedule remained largely unchanged between the end of the 1980s until the beginning of NAFTA in January 1994.

NAFTA represented a significant second wave of Mexican import tariff reductions. While in 1993 only 10% of manufacturing imports from the US fell into a tariff category of 15% or less, this fraction increased to 60% in January of 1994 (Lopez-Cordova, 2002). In contrast, NAFTA had a smaller effect on US tariffs on Mexican exports as these were at already low levels before NAFTA took effect.¹¹ Figure A.2 provides an illustration of average Mexican tariff changes on US imports and their sectoral variation over the period 1993-2000.

Concerning the importance of other trade partners, US imports have consistently accounted for 75-80% of total Mexican imports during the 1990s. In particular, the period under study precedes China's admission to the World Trade Organization in 2001 and the subsequent surge of Chinese imports into Mexico (Iacovone *et al.*, 2010).

Any analysis of NAFTA's consequences in Mexico must address the empirical challenge that the beginning of NAFTA coincided with a severe economic crisis that unfolded in Mexico in 1995, the adverse consequences of which are apparent in Mexican real income data until the beginning of the following decade (Attanasio and Binelli, 2010). As discussed in more detail in the following sections, the present empirical analysis addresses such concerns by focusing on parts of the variation in relative price changes that are plausibly unrelated to the consequences of economy wide macroeconomic shocks.

2.2 Data

The following subsections provide a description of the datasets used in this paper, and further details can be found in the Appendix.

¹¹The average export tariff was approximately 2% in December 1993.

2.2.1 Central Bank Store Price Surveys

The great majority of countries, including Mexico, are subscribers to the ILO/IMF Consumer Price Index Dissemination Standard. This manual imposes a clear set of rules on how to compile and process data in order to report national consumer price inflation. The backbone of national CPI reporting are central bank store price surveys that are collected throughout the country and usually at several times during each month.

In a fortunate turn of events, the Artículo 20-Bis of the Mexican Código Fiscal de la Federación requires the central bank since January of 1989 to publish the store price microdata on a monthly basis in the official government gazette, the *Diario Oficial de la Federación*.¹² These publications are phone book like listings of individual city-store-barcode product combinations and their price quotes in a given month.¹³

Starting from 1989, each month of data contains approximately 30,000 individual price quotes across 35 Mexican cities and 284 product groups covering non-durables (e.g. Salchicha sausages, sanitary paper towels, antibiotic pills), durables (e.g. electric water boilers, bicycles), as well as services (e.g. language courses, taxi rides).¹⁴ For the empirical analysis, I compute average price quotes of individual items across three months in the third quarters of 1989, 1993, 1994, 2000, and the first quarter of 1995.¹⁵

These price data have a number of notable features. First, the survey is intended to capture a representative sample of Mexican household consumption and covers street vendors, markets, convenience and specialized stores, as well as supermarkets and department stores across cities. Second, any change in the presentation, appearance, size, modality, model number or otherwise is reported in an appendix of the monthly publications in the *Diario Oficial* as a product item substitution. The objective of the Mexican central bank is to compute price inflation for identical product items over time. This is to say that what I refer to as "persistent" product series are identical barcode products in the identical store over time.¹⁶ This detailed documentation also allows me to test predictions on product

¹²The commonly cited reason is concern of political influence on the computation of consumer price inflation at that point in Mexican history.

¹³While the actual product barcode is not reported, the detailed product descriptions including brand, product name, pack size, model number, and modalities (e.g. color, packaging type) provide an equivalent level of product identification.

¹⁴For product groups in food and beverages the reported monthly price quotes are averages across 2-4 monthly price quotes for each item. The number of cities and product groups increased in a revision in March 1995. The stated figures refer to cities and product groups that were consistently covered both before and after the revision in 1995.

¹⁵The price quotes of the latter three periods were provided by courtesy of Etienne Gagnon at the Federal Reserve Board in Washington D.C. A detailed description of this dataset can be found in Gagnon (2009). For the third quarter price quotes in 1989 as well as 1993, I obtain copies of archival records contained in the Archivo General de la Nación in Mexico City, and digitize these monthly price quotes by double blind data entry. The percentage of non-identical entries was approximately 1%. These cells were then double checked and corrected by hand.

¹⁶To this end, I digitize the complete series of product substitution appendices from January 1989 to the end of 1993 and I obtain the more recent substitution listings between 1994-2000 by courtesy of Gagnon (2009).

replacements and additions in addition to relative price changes of persistent barcode items.

Third, price quotes are reported in prices per common physical unit for the majority of product groups. For product groups where this is not consistently the case (e.g. measured per pack of toilet paper, or measured per bottle of body lotion), I either clean the data by hand to convert it to common physical units (e.g. per roll of toilet paper reported in the product description, or per 100ml of body lotion), or I exclude the product group from the estimations where such a correction is not feasible (e.g. tortillas or clothing where reported prices are based on store sample averages within a city so that changes in product items are unobserved).

Finally, one important limitation of the store price data is that while each city-store-barcode item has a unique identifier code, the individual store identifiers cannot be recovered from these item codes.¹⁷ Potential estimation concerns arising from this limitation will be addressed in detail in the empirical analysis.

The main estimation sample is a panel of individual city-store-barcode items within 144 processed tradable product groups that i) report individual barcode unit values rather than store sample average prices, and ii) could be matched to product groups in the plant production microdata that is described in the following subsection.¹⁸ Table A.1 presents descriptive statistics of the store price microdata together with a breakdown of the estimation sample's coverage of total household expenditure among urban Mexicans in 1994 and 2000.

2.2.2 Manufacturing Establishment Surveys

There are two general empirical challenges when empirically investigating vertical product differentiation in plant microdata. The first is that the majority of manufacturing establishment surveys do not report physical output quantities in combination with sales revenues to compute output unit values. The second is that product quality differentiation is empirically meaningful only at a very fine level of product aggregation. Most plant surveys report two digit (e.g. food processing), four digit (e.g. meat processing), or sometimes six digit (e.g. meat products except poultry) industrial classifications, which would be insufficient to match the detailed product groups that are present in the store price surveys described above.

Fortunately, the microdata reported in the Mexican Encuesta Industrial Mensual (EIM) make it possible to address both of these challenges. Starting in 1994, the survey reports monthly physical output in combination with sales at the level of several thousands of product groups within 203 six digit CMAP manufacturing sectors. Plants on average report output and sales across several products, so that the level of aggregation present in the

¹⁷The original reason for encoding the store identifier in the Diario publications was confidentiality concerns. Unfortunately, the correspondence table between the published and the actual store identifiers appears to have been lost for the period before 2002.

¹⁸See Appendix for product group descriptions and the correspondence between product groups across the different datasets.

data can be thought of as individual product lines within an establishment. These product descriptions are used to match the finest available level of product groups in the store price surveys to the product classifications in the plant microdata, resulting in 144 processed tradable product groups (see Appendix for details).

The second plant dataset that I draw on in the analysis is the Encuesta Industrial Anual (EIA) which covers the identical plants at annual intervals. In particular, I use the EIA to complement the EIM data with annual plant level information on the use of imported inputs as well as employment. Both the EIM and the EIA microdata are administered by the Instituto Nacional de Estadística, Geografía, e Informática (INEGI). I obtained access to the confidential microdata for 12 months of the EIM data in 1994 and the matching annual records contained in the EIA. These establishment surveys cover all manufacturing production sectors and represent roughly 85% of total Mexican manufacturing output. The data do not cover the universe of Mexican production establishments as the surveys typically omit the tail of small producers (INEGI, 2000). Table A.2 presents descriptive statistics of the 1994 plant microdata.

2.2.3 Household Consumption Surveys

I use the microdata of the Mexican national household consumption survey in 1994 (Encuesta Nacional de Ingresos y Gastos de los Hogares, ENIGH) for information on per capita household incomes, expenditure weights and unit purchase values across 255 processed tradables product groups that overlap with the plant production product groups of the EIA/EIM plant data (see Appendix for details). These surveys are administered by INEGI from where I obtain access to the data. To be consistent with the urban only coverage of the store price surveys, I only use data on households in urban classified municipalities.¹⁹

There are several notable features of the household consumer surveys. First, they cover the whole of Mexico and report nationally representative sample weights. Second, they are collected to represent total household consumption expenditure during the third quarter of 1994 which coincides with the collection of the central bank store price series. Third, they provide a rich breakdown of product groups into several hundred product codes covering all types of consumption expenditure. Fourth, they report every single transaction within a product group made by members of the household. Fifth, they report unit values (e.g. per kilogram or per liter) for 118 product groups that pertain to food and beverages and tobacco expenditures out of the 255 total product categories that could be matched to product groups that are also present in the EIA/EIM plant data. Finally, the surveys report the store type linked to every single transaction. The store types include street vendors, markets, convenience and specialized stores, and supermarkets and department stores. Table A.1 presents an overview of the shares of consumption expenditures captured by the processed tradables estimation sample, and Table A.3 presents descriptive statistics

¹⁹In the ENIGH 1994 survey, these are municipalities with more than 2500 residents.

for the urban household ENIGH sample in 1994.

2.2.4 Input Tariffs

Intermediate input tariff changes are computed at the four digit industrial classification (NAICS) of the Mexican 2003 input output table.²⁰ I use total (direct and indirect) requirement coefficients from the import matrix to compute the weighted average US tariff changes across four digit input sectors for each destination sector. In the instrumental variable regressions, I also make use of direct and indirect requirement coefficients of the total use matrix (not just imports).

Out of a total of 101 tradable four digit NAICS sectors, I focus attention on 75 manufacturing input sectors to compute weighted average intermediate import tariff changes for each of the 101 tradable destination sectors. The focus on manufacturing inputs is due to concerns that agricultural and primary imports are subject to significant non-tariff barriers (e.g. Krueger, 1999). The Appendix provides the concordance table that was used to match the 144 product groups in the store price estimation sample to four digit destination sectors in the Mexican input output table.

Tariff changes at the four digit NAICS level are based on average tariff changes across eight digit tariff lines from the Mexican Secretaria de Economia. NAFTA tariff changes in the estimations refer to the difference between average applied rates during the year 2000 and December 1993. For tariff changes on final good imports and exports over the same period, I match the 144 store product groups to the eight or six digit levels of tariff lines.²¹ Figure A.2 provides an illustration of average Mexican tariff changes on US imports and their cross-sectoral variation over the period 1993-2000.

3 Stylized Facts about Vertical Differentiation in Mexican Consumption and Production

This section draws on the Mexican plant and household microdata to document a set of stylized facts about vertical differentiation in production and consumption at the beginning of NAFTA in 1994. These insights serve to motivate the theoretical framework in Section 4 and the empirical estimations in Sections 5-7.

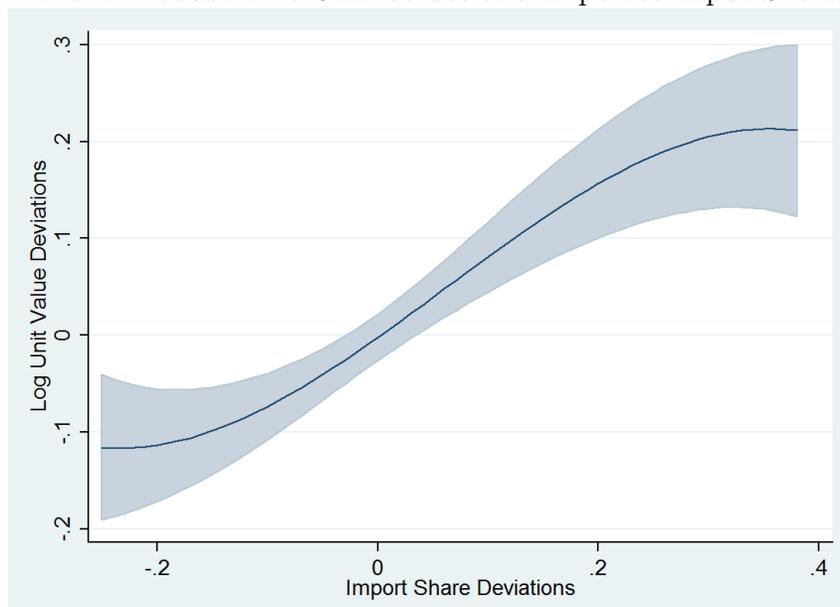
Plant Product Line Unit Values Increase in Imported Input Shares: Figure 1 depicts the first stylized fact. The graph plots the relationship between deviations of product line log unit values (prices per common physical unit) and plant level imported input shares.

²⁰This is the most recent available Mexican IO table since 1979. A data request had to be filed at INEGI in order to obtain the four digit break up of the Mexican IO table.

²¹While the available input output information dictates a four digit aggregation for intermediate tariff changes, this constraint does not apply to final product tariff changes. Six or eight digit matches depend on the levels of product aggregation in store surveys relative to tariff lines. Data on US applied tariff rates is taken from Feenstra and Romalis (2002).

Estimations are based on 2656 plants reporting across 9163 unique product lines in 1018 product groups pertaining to 79 six digit manufacturing sectors that produce consumer goods. The unit value-import share elasticity is positive and statistically significant at the 1% level. On average higher unit values within disaggregated product groups embody larger shares of imported inputs in a statistically significant way.

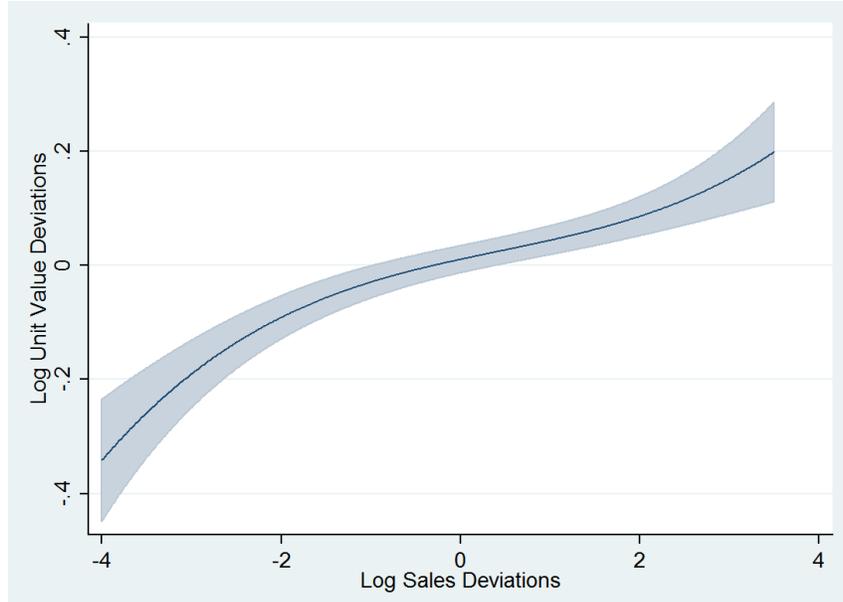
Figure 1: Plant Product Line Unit Values and Imported Input Shares in 1994



The fitted relationship corresponds to the best fitting polynomial functional form according to the Akaike Information Criterion (AIC). Standard errors are clustered at the product level and the shaded area indicates 95% confidence intervals. The y-axis depicts the residuals of a regression of log monthly product line unit values on month-by-product and month-by-state fixed effects. Estimations are based 2656 plants in 79 six digit final good manufacturing sectors and 1018 products over 12 months in 1994. The number of observations is 94741. The x-axis depicts the residuals of a regression of annual 1994 plant level imported input shares on product and state fixed effects. The number of observations in this regression is equal to the number of unique product lines (9163). The bottom and top 0.5% on the x-axis are excluded from the graph.

Plant Product Line Unit Values Increase in Market Shares: Figure 2 depicts the second stylized fact. The graph plots the relationship between deviations of product line log unit values and product line log sales from month-by-product and month-by-state means during 12 months in 1994. The estimation is based on the identical sample of plants and product lines reported for the previous figure. The unit value-sales elasticity is positive, close to log linear, and statistically significant at the 1% level. On average higher unit values within disaggregated product groups embody larger market shares in a statistically significant way.

Figure 2: Plant Product Line Unit Values and Product Market Shares in 1994

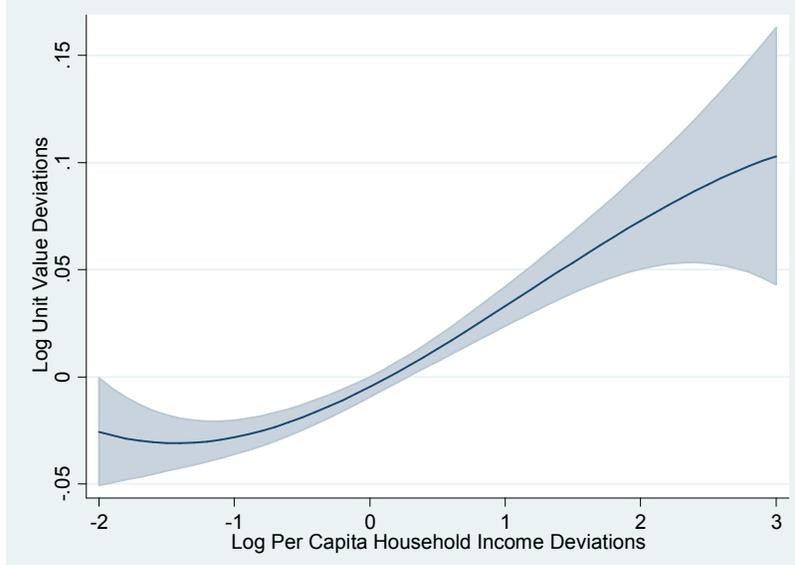


The fitted relationship corresponds to the best fitting polynomial functional form according to the Akaike Information Criterion (AIC). Standard errors are clustered at the product level and the shaded area indicates 95% confidence intervals. The y-axis and the x-axis depict the residuals of two regressions of log product-line unit values or log product line sales on month-by-product and month-by-state fixed effects. Estimations are based 2656 plants in 79 six digit final good manufacturing sectors and 1018 products over 12 months in 1994. The number of observations is 94741. The bottom and top 0.5% on the x-axis are excluded from the graph.

Household Purchase Unit Values Increase in Household Income: Figure 3 depicts the third stylized fact. The graph plots the relationship between deviations of household weighted average log unit values from city-by-product-by-store type means and deviations of log household incomes from the national mean in 1994. The weights are given by households expenditure weights attached to each reported purchase. Reported store types are markets, street vendors, convenience and specialized stores, and supermarkets and department stores.²² For purchases in the same city-by-product-by-store type cell, unit values of the average household expenditure items are statistically significantly increasing in household per capita incomes.

²²In Section 6 I report the estimated price gaps between rich and poor households both before and after the inclusion of store type fixed effects in order to learn about the potential role of heterogeneous store markups for identical barcode items in Figure 3. As discussed in more detail, the inclusion of store type fixed effects increases, rather than decreases, the estimated unit value differences.

Figure 3: Household Purchase Unit Values and Household Income in 1994



The fitted relationship corresponds to the best fitting polynomial functional form according to the Akaike Information Criterion (AIC). Standard errors are clustered at the municipality level and the shaded area indicates 95% confidence intervals. The y-axis depicts the residuals of a regression of log unit purchase values on city-by-product-by-store type fixed effects. These residuals are then averaged at the household level using reported expenditure weights. The x-axis depicts mean deviations of log household per capita incomes. Estimations are based on urban Mexican households in 1994 and subject to nationally representative sample weights. The bottom and top 0.5% on the x-axis are excluded from the graph.

To summarize, these stylized facts motivate a theoretical framework in which differences in household expenditures across the income distribution are linked to differences in plant technologies through vertical differentiation in consumption and production. While the documented moments in the Mexican microdata are neither individually novel²³, nor particularly surprising, they have so far not been considered together in a unified framework. The remainder of the paper has the two-fold objective to formalize a product quality interpretation of the documented stylized facts, and to empirically test the implications for welfare inequality that arise from this setting in the context of NAFTA in Mexico.

4 Theoretical Framework

This section proposes a model of quality choice in a setting with heterogeneous households in consumption and heterogeneous plants in production. The model serves two main objectives. First, it formalizes a product quality interpretation of the documented stylized facts. Second, it yields testable predictions on NAFTA's effect on Mexican consumer prices, and guides the estimation of the cost of living consequences of these relative price effects across the Mexican income distribution.

²³On the production side, the presented results confirm recent findings in Kugler and Verhoogen (2009; 2011) from similarly rich Colombian plant microdata. On the consumption side, for example Deaton (1997) discusses evidence that in the same village unit values within agricultural product groups are increasing in household incomes.

While several existing theoretical frameworks have been proposed to capture separately either the consumption side of quality choice in Figure 3, or the production side in Figures 1 and 2, the two have not been modeled in a unified framework. In this paper, I show that the introduction of heterogeneous household quality evaluations into a model of quality choice across heterogeneous plants poses one key challenge, which is that *physical* product quality is distinct from *perceived* product quality in the market place. The definition of this distinction is of interest more generally because it reveals the concept of product quality that has been implicitly estimated from product prices and market shares in a prominent strand of empirical work in industrial organization (e.g. Berry, 1994) and international trade (e.g. Khandelwal, 2010; Hallak and Schott, 2011; Feenstra and Romalis, 2012) when acknowledging that we live in a world with non-homothetic tastes (Figure 3).

The key features of the model are as follows. On the consumption side, I introduce non-homothetic preferences that allow household quality choice to differ across the income distribution. On the production side, I follow Kugler and Verhoogen (2011) who propose complementarity between plant efficiency and input quality in the production of output quality, and introduce the assumption that input quality is increasing in the use of imported inputs. The following provides a summary of the key features of the model while a more detailed exposition is provided in the Appendix.

4.1 Physical and Perceived Product Quality under Non-homothetic Preferences

A household h 's preferences are given by a two-tier Dixit-Stiglitz utility function in which the upper tier is Cobb Douglas across product groups denoted by subscript k , while the subutility index U_{hk} is a CES function over varieties denoted by subscript i within the product group.²⁴

$$U_h = \int_{k=0}^K U_{hk}^{\mu_{hk}} dk \quad U_{hk} = \left(\int_{i=0}^{I_k} (q_{ki}^{\varphi_h} x_{hki})^{1-1/\sigma} di \right)^{\frac{1}{1-1/\sigma}} \quad 0 < \mu_{hk} < 1 \quad \sigma > 1 \quad (1)$$

For ease of exposition, product group subscripts k are suppressed in the remainder of this subsection. Household utility is a function of physical units consumed, x_{hi} , and a variety's quality $q_i^{\varphi_h}$, where $\varphi_h > 0$ is a household specific taste parameter that determines the intensity of preferences for product quality. Product quality thus enters as a shift in utility derived from consuming a given amount of physical units, and the extent of this shift is allowed to vary across household valuations of quality. To introduce non-homotheticity across the

²⁴While the Dixit-Stiglitz structure is standard and will be convenient to solve the model, all results carry through to, for example, a nested logit demand structure as long as the source of non-homotheticity is modeled in the same way as in (1), namely through household quality evaluations rather than heterogeneity in price elasticities. In a recent contribution, Handbury (2012) draws on US homescanner data to separately estimate non-homotheticity in quality evaluations and differences in price elasticities. The evidence presented therein is in support of this assumption.

income distribution in a reduced form approach, I let φ_h be a positive function of household per capita income, as for example proposed by Hallak (2006), and empirically estimated by Handbury (2012).²⁵ Differences in household income, in turn, enter the model through differences in household endowments of effective labor units.

The first implication of (1) is that a household's expenditure shares within product groups increase in φ_h for products with above average quality, and decrease in φ_h for below average quality products.²⁶ As a consequence, the weighted average quality of the household's consumption basket increases in its quality valuation φ_h .

The second implication of (1) concerns the relationship between changes in product quality and observed market demand. For clarity, let us first consider the conventional approach to the stylized facts in Figures 1 and 2 by assuming that preferences in (1) pertain to one single representative household. In this case, the elasticity of market demand with respect to quality becomes a function of this representative agent's φ_h^* parameter, which can be set to unity without loss of generality.²⁷ Holding prices constant, total sales increase or decrease subject to the elasticity $(\sigma-1)\varphi_h^*$ when product quality changes. This insight has been used extensively in empirical industrial organization and international trade to infer differences in product quality by observing both prices per physical unit and product market shares.

Following this approach, the positive relationship between unit values and market shares in Figure 2 would imply that prices per physical unit increase close to log linearly with product quality.²⁸ It would also imply that producers offering a better quality adjusted price sort into higher quality products.²⁹ Taken together, the microdata in Figures 1-3 would thus suggest that higher quality products are **1)** produced using higher shares of imported inputs, **2)** produced by more productive plants, and **3)** consumed by richer households.

The particularly convenient assumption of the representative agent framework is that by setting $\varphi_h^*=1$, one can avoid any distinction between *physical* product quality (q_i) (e.g. degree of shininess or number of screws securing a handle) and *perceived* quality ($q_i^{\varphi_h^*}$) without any loss of generality. In the presence of non-homothetic preferences in (1), however, one cannot avoid this distinction because the observed market valuation of a product's quality characteristics reflects heterogeneous household evaluations of the identical features. To see this, I derive this reference evaluation from the elasticity of the horizontal summation of household demands with respect to physical product quality. Let y_{hi} and y_i indicate household h's expenditure on product item i within a product group and total market sales

²⁵A straight forward way to microfound this structure would be to assume complementarity between the consumption of a (normal) outside good and higher quality within differentiated product groups.

²⁶See Appendix for details.

²⁷With a representative agent, setting $\varphi_h^*=1$ is a monotonic transformation of utility in (1).

²⁸To see this, we can write the elasticity of sales ($\sum_H y_{hi}$) with respect to price as: $\frac{\partial \ln(\sum_H y_{hi})}{\partial \ln p_i} = (\sigma - 1) \left(\frac{\partial \ln q_i}{\partial \ln p_i} - 1 \right)$.

²⁹To see this, we note that $\frac{\partial \ln(\sum_H y_{hi})}{\partial \ln p_i} = (\sigma - 1) \left(\frac{\partial \ln q_i}{\partial \ln p_i} - 1 \right) > 0$ implies that $\frac{\partial \ln q_i}{\partial \ln p_i} > 1$, and so $\frac{\partial \ln(q_i/p_i)}{\partial \ln p_i} > 0$.

of variety i respectively. We get:

$$\frac{\partial \ln(\sum_H y_{hi})}{\partial \ln q_i} = (\sigma - 1) \left(\sum_H \frac{y_{hi}}{y_i} \varphi_h \right) = (\sigma - 1) \varphi_i^* \quad (2)$$

Expression (2) reveals the concept of product quality that we infer from prices and market shares when acknowledging the presence of non-homotheticity. In particular, $q_i^* = q_i^{\varphi_i^*}$ is an expenditure weighted average valuation of quality. Quality is thus measured in units of market valuation, rather than in physical units of product attributes, because the reference evaluation in (2), φ_i^* , is not constant across the spectrum of physical product qualities. In fact, (1) implies that expenditure shares $\left(\frac{y_{hi}}{y_i} \right)$ of households with lower (higher) quality evaluations are decreasing (increasing) in an item's physical quality, so that the distribution of product quality estimated from unit values and market shares will have a larger variance and be skewed to the right compared to the underlying distribution of differences in physical quality characteristics.³⁰

In the present setting, expression (2) has two main implications. First, it matters for how we model quality choice across plants on the production side in order to derive the observed unit value relationships in the microdata. That is, while all three conclusions about the stylized facts from the representative agent approach above still hold, their log linear functional forms do not, unless we relabel product quality from a physical product concept to one in terms of perceived market valuation. To see this difference more clearly, note that under non-homothetic tastes in (1) and (2), holding prices constant, a percentage increase in physical quality leads to a larger percentage increase in sales for an initially higher quality product compared to a lower quality item. In contrast, this elasticity is constant and equal to $(\sigma - 1)$ with respect to changes in perceived quality ($q_i^* = q_i^{\varphi_i^*}$).

Second, expression (2) can in principle be a source for great inconvenience when modeling quality choice on the producer side. In particular, while the finding that market shares reflect expenditure weighted household evaluations of product quality is intuitive and not particular to preferences in (1), the CES structure does impose the convenient assumption that rich and poor households respond to price changes in the same way. If this were not the case, then expenditure shares $\left(\frac{y_{hi}}{y_i} \right)$, and thus the market evaluation of a given quality characteristic, would vary across a firm's pricing decisions.³¹

³⁰To see this, we can write: $\varphi_i^* = \sum_H \frac{y_{hi}}{y_i} \varphi_h = N_H Cov \left(\frac{y_{hi}}{y_i}; \varphi_h \right)_i + \bar{\varphi}_h$. The intuition for this result is that the initial total sales of higher (lower) quality products derive to a higher (lower) extent from consumers who attach greater value to a given percentage change in quality, so that sales respond more (less) to a given change in quality.

³¹In other words, the CES structure assures that the change in sales due to a change in product quality is independent on a firm's pricing decision.

4.2 Technology

On the production side, the final goods sector consists of a continuum of monopolistically competitive plants that produce horizontally and vertically differentiated varieties within any given product group. Production of the final good can be separated into a production function of physical units, and a production function of quality that depends on plant characteristics and input quality. The production function of final goods is given by:

$$F_{Fi} = \lambda_i m_i \quad (3)$$

λ_i is a plant specific productivity parameter that in this context of quality differentiation I will refer to as technical efficiency. It defines the efficiency at which a plant converts a given number of intermediate inputs, m_i , into units of final products. Following Kugler and Verhoogen (2011), the production of final product quality is subject to complementarity between input quality, z_i and technical plant efficiency:³²

$$q_i^* = \left(\alpha \lambda_i^{\psi \vartheta} + (1 - \alpha) z_i^{\gamma \vartheta} \right)^{1/\vartheta} \quad \vartheta < 0 \quad \psi > 0 \quad \gamma > 0 \quad 0 < \alpha < 1 \quad (4)$$

Here, ϑ determines the degree of complementarity between technical plant efficiency and input quality, and the assumption $\vartheta < 0$ imposes log supermodularity of quality in plant efficiency and intermediate quality.³³ The intermediate input is produced subject to perfect competition using labor hours denoted by l . The input production function is:

$$F_{Mj} = \frac{l_j}{z_j^{\delta(\tau)}} \quad (5)$$

Intermediate unit costs thus increase in input quality. Substituting intermediate unit costs into final product unit costs, we get $c_i = \lambda_i^{-1} p_{im} = \lambda_i^{-1} w z_i^{\delta(\tau)}$, where w is the wage rate. Kugler and Verhoogen (2011) interpret the input characteristic z_i as intermediate quality which is complementary to plant efficiency in producing output quality. The simplest possible way to introduce foreign inputs into this setting is by letting input quality be increasing in shares of imported inputs. (5) captures this assumption in a simple reduced form approach by letting the elasticity of unit input costs with respect to input quality (δ) be an increasing function of foreign input costs (τ).³⁴

³²This is consistent with empirical findings in Kugler and Verhoogen (2009) and Manova and Zhang (2012).

³³See Costinot (2009). Intuitively, $\vartheta < 0$ assures that a marginal increase in input quality leads to a greater increase in final product quality for a higher λ_i plant. Note that in (4) product quality enters the production side not as physical concept (q_i), but in terms of market valuation (q_i^*). This convenient functional form assumption yields equilibrium relationships between unit values, market shares, and input usage that are consistent with the documented stylized facts from the Mexican plant microdata.

³⁴Since I deliberately abstract from relative factor income effects of import access, this is convenient but without loss of generality. To see this, consider a unit input cost function $c = c(w, \tilde{w})$ of any linearly homogeneous input production function where w and \tilde{w} are prices of a domestic and a foreign factor of production. Using Shephard's lemma, we get the unit factor requirements: $a = \frac{\partial c}{\partial w}$ and $b = \frac{\partial c}{\partial \tilde{w}}$. Differentiating

As in Melitz (2003), to enter the final-good sector, plants pay an investment cost f_e measured in domestic labor units in order to receive a technical efficiency draw λ . The distribution of this parameter is assumed to be Pareto with a c.d.f. $G(\lambda)=1-\left(\frac{\lambda_m}{\lambda}\right)^\xi$, where $\lambda_m < \lambda$, and ξ is the shape parameter. There is a fixed cost of production, f in each period, and plants exit with exogenous probability χ each period. Given zero cost of horizontal differentiation, each plant choosing the same product quality produces a distinct variety so that λ can be used to index both plants and varieties.

4.3 Predictions and Estimation Equations

In equilibrium, plants simultaneously choose output quality and prices to maximize profits, while households maximize utility in (1). In this setting, the model guides the empirical estimation in three main respects that I summarize here in the order of the three subsequent empirical sections. First, the model yields a series of predictions on NAFTA's effect on Mexican consumer prices that I am able to test empirically by drawing on the barcode level store price panels (Predictions 1.1-1.3). Second, it yields an estimation equation of the cost of living implications of these price effects in terms of observable moments in the household consumption microdata (Prediction 2). Third, it guides the estimation of differences in quality and quality adjusted prices across the Mexican income distribution in terms of observable moments in the household and plant microdata (Prediction 3).

In equilibrium, the elasticity of product unit values with respect to perceived quality is given by: $\frac{\partial \ln p_{ki}}{\partial \ln q_{ki}^*} = \eta_k = \frac{\delta(\tau)}{\gamma} - \frac{1}{\psi_k}$. Here, the parameter ψ_k from the quality production function (4) represents the equilibrium elasticity of perceived quality with respect to plant efficiency $\left(\frac{\partial \ln q_{ki}^*}{\partial \ln \lambda_{ki}}\right)$. Following Kugler and Verhoogen (2011) and earlier work by Sutton (1998), this parameter can be thought of as a product group specific scope for quality differentiation. A given distribution of ex ante plant heterogeneity leads to a wider range of product quality if the scope parameter ψ_k is greater. Intuitively, the first term in η_k represents the unit cost-quality elasticity in absence of endogenous plant sorting, while the second term captures the equilibrium link between plant efficiency and quality.

In this setting, we can derive three testable predictions about the effect of cheaper access to imported inputs on Mexican consumer prices. The first prediction concerns the average effect of input tariff cuts on the relative prices of initially more or less expensive products within product groups.

Prediction 1.1: *Input tariff cuts decrease the relative price of higher quality products.*

$$\frac{\partial^2 \ln p_{ki}}{\partial \ln q_{ki}^* \partial \tau_k} = \frac{\partial \eta_k}{\partial \tau_k} > 0 \quad (6)$$

$c=c(w, \tilde{w})$, we get: $dc=a dw+b d\tilde{w}$. Rearranging, we get: $\frac{dc}{c} = \frac{aw}{c} \frac{dw}{w} + \frac{b\tilde{w}}{c} \frac{d\tilde{w}}{\tilde{w}}$ (e.g. Bhagwati *et al.*, 2009, pp. 143-144). The elasticity of input unit costs with respect to input quality can then be written as: $\left(\frac{dc}{c} / \frac{dz_j}{z_j}\right) = \frac{d\left(\frac{b\tilde{w}}{c}\right)}{dz_j/z_j} \frac{d\tilde{w}}{\tilde{w}}$, which is increasing in foreign factor costs as long as the foreign input cost share $\left(\frac{b\tilde{w}}{c}\right)$ is increasing in input quality z_j .

Because the production of higher quality is more intensive in imported inputs, cheaper access to foreign inputs reduces the equilibrium elasticity of unit values with respect to product quality (η_k). Coupled with the observation in Figures 1 and 2 that on average $0 < \eta_k < 1$ across consumer product groups, this first prediction implies that input tariff cuts should lead to a relative reduction in the price of initially more expensive products within product groups.

The second prediction concerns the heterogeneity of this observable effect as a function of sectoral scopes for product differentiation. In particular, I can exploit a common feature of existing models of quality choice across plants in production - that unit values embody technological heterogeneity differently across product groups with different scopes for differentiation - for empirical estimation.

Prediction 1.2: *The observed average effect of input tariff cuts on relative prices is driven by differentiated product groups.*

$$\frac{\partial^2 \ln p_{ki}}{\partial \ln p_{ki}^0 \partial \tau_k} = \begin{cases} 0 & \text{for } \eta_k = 0 \text{ ("Undifferentiated")} \\ \frac{1}{\eta_k} \frac{\partial \eta_k}{\partial \tau_k} > 0 & \text{for } \eta_k > 0 \text{ ("Differentiated")} \end{cases} \quad (7)$$

To empirically estimate differences in sectoral scopes for quality differentiation, the model yields a convenient estimation equation in terms of observable moments in the plant micro-data:

$$\frac{\partial \ln p_{ki}}{\partial \ln \left(\frac{q_{ki}^*}{p_{ki}} \right)} = \frac{\eta_k}{1 - \eta_k} = (\sigma - 1) \frac{\partial \ln p_{ki}}{\partial \ln s_{ki}} \quad (8)$$

where the final term is a product group specific elasticity between unit values and market shares. In particular, notice that product group differences in this observable elasticity provide a sufficient statistic to separate product groups into sectors that show a statistically significant relationship between initial unit values and quality ($\eta_k > 0$) which I refer to as differentiated, and sectors without a significant relationship between unit values and quality (η_k close to 0) which I refer to as undifferentiated.³⁵

The third prediction concerns the effect of input tariff cuts on the distribution of product market shares across the quality distribution.

Prediction 1.3: *Input tariff cuts lead to a reallocation of market shares towards higher quality products.*

$$\frac{\partial^2 \ln s_{ki}}{\partial \ln q_{ki}^* \partial \tau_k} = \frac{\partial ((\sigma - 1)(1 - \eta_k))}{\partial \tau_k} < 0 \quad (9)$$

³⁵Intuitively, as the link between plant efficiency and product quality (captured by scope parameter ψ_k) decreases, η_k decreases. In differentiated sectors unit values increase with product quality, but less than one for one due to higher plant productivities associated with quality. In non-differentiated sectors more productivity improvements are necessary to produce units of quality so that prices are insignificantly related to higher plant capability and product quality.

It is apparent from (4) and (5) that cheaper access to higher quality inputs benefits producers of higher final good quality relatively more. By reducing η_k this increases the elasticity of plant revenue with respect to quality in the final good sector $\left(\frac{\partial \ln s_{ki}}{\partial \ln q_{ki}^*} = (\sigma - 1)(1 - \eta_k)\right)$. The testable implication is that product groups with higher tariff cuts on their imported inputs should experience a shift of market shares towards the higher end of the quality distribution. In the next section, I draw on detailed monthly listings of product exit and entry in the store price microdata to test this prediction empirically across the unit value distribution.

Because household quality evaluations (φ_h) are increasing in household income, the consumption baskets of richer households embody higher weighted average product quality than those of poorer households. Cheaper access to foreign inputs thus gives rise to differences in cost of living inflation across the income distribution. Under two assumptions that the model makes explicit, this effect can be expressed as a function of initial expenditure share differences and price growth. Following Konus (1939), a household's cost of living index is defined as the ratio of expenditures necessary to reach a reference utility level u^* subject to store price vectors at two periods p^{t0} and p^{t1} : $\frac{e(u^*, p^{t1})}{e(u^*, p^{t0})}$. Denoting a poor and a rich household by subscripts P and R respectively and taking log differences, we get:

Prediction 2: *The relative price effect of input tariff cuts increases real income inequality through differences in cost of living inflation.*

$$\ln\left(\frac{e(u^*, p^{t1})}{e(u^*, p^{t0})}\right)_P - \ln\left(\frac{e(u^*, p^{t1})}{e(u^*, p^{t0})}\right)_R = \sum_I (s_{kiP}^{t0} - s_{kiR}^{t0}) \ln\left(\frac{p_{ki}^{t1}}{p_{ki}^{t0}}\right) > 0 \quad (10)$$

Within the structure of the model, (10) presents the difference in the exact ideal price index due a change in the price of quality. Two assumptions underlie this convenient result. First, the model abstracts from general equilibrium effects on relative incomes of rich and poor households, which would affect the ideal expenditure weights due to non-homotheticity in (1) (Diewert, 1979). Second, the CES functional form in (1) abstracts from differences in the elasticity of substitution across households. That is, while households are allowed to substitute away from higher price increases (the source of the traditional CPI substitution bias), they do so at the same rate so that by taking the difference in household cost of living inflation, second period expenditure shares drop out of the expression. Notice that if either of these assumptions were violated, (10) would remain a first order approximation of the difference in cost of living inflation.³⁶

Finally, the equilibrium elasticity of product quality with respect to quality adjusted productivity is $\frac{\partial \ln q_{ki}}{\partial \ln(q_{ki}^*/c_{ki})} = \frac{1}{1-\eta_k} > 0$, so that differences in household quality choices are predicted to translate into differences in weighted average plant productivities and, thus, quality adjusted prices across the income distribution. Denoting log differences in weighted average product quality and (inverse) quality adjusted prices by $(\ln Q_P - \ln Q_R)$ and $(\ln(\frac{Q}{P})_P - \ln(\frac{Q}{P})_R)$ respectively, we get:

³⁶The Appendix provides a more detailed derivation of (10).

Prediction 3: *Differences in household quality choices translate into differences in quality adjusted prices.*

$$\ln\left(\frac{e(u^*, p^{t1})}{e(u^*, p^{t0})}\right)_P - \ln\left(\frac{e(u^*, p^{t1})}{e(u^*, p^{t0})}\right)_R = \ln\left(\frac{Q}{P}\right)_P - \ln\left(\frac{Q}{P}\right)_R = \sum_K (1 - \eta_k) (\ln Q_{Pk} - \ln Q_{Rk}) < 0 \quad (11)$$

Expression (11) represents the following comparative static result. For any given observed differences in household consumption expenditure between rich and poor people, (11) presents the difference in household cost of living inflation that would have to occur if the conventional assumption was true that quality differences do not embody differences in plant productivity. In particular, under standard assumptions $\eta_k=1$ would imply that unit values increase proportionally in product quality, so that higher quality is unrelated to quality adjusted prices. (11) is thus equivalent to a quality adjusted price index in a cross-section of households that face the identical store prices in the identical location.

To summarize, the theoretical framework guides the empirical estimation in three main respects. It yields testable predictions on the effect of NAFTA’s tariff cuts on Mexican consumer prices. It provides an estimation equation to evaluate the cost of living implications of these relative price changes. And it guides the estimation of differences in quality and quality adjusted prices in household consumption. In the remainder of the paper, I draw on the Mexican microdata to empirically estimate these in the stated order.

5 Testing NAFTA’s Effect on Relative Consumer Prices

This section draws on the central bank store price surveys to empirically test the predictions on NAFTA’s effects on Mexican consumer prices. First, I test for the average effect of input tariff cuts on the relative price of initially more expensive relative to less expensive products within product groups. Second, I test whether this observed average effect is driven by product groups that are estimated to be differentiated in terms of observable moments in the plant microdata. Finally, I draw on detailed monthly listings of product entry and exit to test for the reallocation effects of input tariff cuts on product market shares within product groups.

5.1 NAFTA’s Average Effect on Relative Prices

5.1.1 Empirical Strategy

To test for the effect of NAFTA’s input tariff cuts on the relative price of initially more expensive relative to less expensive barcode items within product groups in (6), I run the following baseline regression equation:

$$d\ln p_{ick}^{94-00} = \alpha_{ck} + \beta_1 \ln p_{ick}^{94} + \beta_2 \ln p_{ick}^{94} * d\tau_k^{93-00} + \varepsilon_{ick} \quad (12)$$

$d\ln p_{ick}^{94-00}$ is the log price change of a unique barcode-store combination i in product group k and city c from the third quarter in 1994 to the third quarter in 2000, and $d\tau_k^{93-00}$ is the weighted average intermediate import tariff change under NAFTA in percentage points across four digit input industries of product group k . α_{ck} indicates city-by-product group fixed effects. Price growth is thus regressed on initial log price levels and their interaction with a product group's intermediate input tariff change within city-by-product group cells. The coefficient β_2 captures how the relative price growth of initially higher or lower unit values within city-by-product group cells differs across product groups with higher or lower intermediate import tariff cuts. To address the concern of correlated error terms (ε_{ick}) across barcode items in the same product group, standard errors are clustered at the level of 144 product groups.

5.1.2 Baseline Results

Table 1 presents the baseline estimation results. Column 1 reports results before including the interaction of initial log prices with intermediate tariff changes. Store prices within city-by-product group cells appear to have significantly converged in Mexico over the period 1994-2000. This result could be driven by a number of economic stories including trade and the relative price of quality, as well as, for example, the very significant economic crisis that unfolded over this period. Alternatively, $\beta_1 < 0$ might just be a consequence of measurement error or temporary store price hikes (drops) in the initial period, so that initially high (low) prices within a city-product cell would have a mechanical tendency towards lower (higher) price growth.³⁷

Column 2 of Table 1 then introduces the product group's tariff interaction of interest. Product groups with higher tariff cuts on their intermediates are characterized by lower relative price growth of initially higher unit values in a statistically significant way. In Column 3 the point estimate of β_2 is unaffected by the inclusion of contemporaneous import and export tariff cuts on final consumer products. The precision of the β_2 estimate slightly increases, while no statistically significant effect of tariff changes on final goods is found. This result is consistent with the minor share of Mexican consumption expenditure on US imported final consumer goods documented in Figure A.1.1, and the fact that export tariffs to the US had already been at low levels before NAFTA.³⁸

³⁷This would be analogous to a case of Galton's fallacy as discussed by Quah (1993) in the context of the empirical literature on growth and convergence.

³⁸See also Verhoogen (2008) for a discussion of the relatively minor importance of export tariff changes compared to real exchange rate movements due to the Peso crisis in 1995.

5.1.3 An Instrumental Variable Strategy for Input Tariff Changes and Additional Robustness Results

The identification of β_2 is based on comparing the relative price growth of barcode items within product groups across product groups that have been exposed to different degrees of intermediate import tariff cuts. By focusing on the within product group dimension of relative price changes, the identifying assumption is that intermediate tariff cuts are not targeted at particular product lines within sectors. Given that imported intermediate input tariffs affect a wide range of producers because inputs are shared across sectors, this assumption appears plausible.

To address remaining concerns about potentially omitted factors that could affect relative price growth systematically across the initial price distribution, while also being correlated with the weighted average intermediate tariff cuts across product groups, I also propose an instrumental variable (IV) strategy. In particular, I notice that there is a subset of intermediate sectors, such as basic chemicals, that have significant shares of input use across a wide range of domestic destination sectors. Since endogeneity concerns revolve around the strategic targeting of tariff cuts at particular establishments within industries, such concerns are less likely with respect to input categories that are widely shared across the Mexican economy.

Guided by this logic, I adjust input-output requirements to sum to 100% for each destination product group over a subset of 20% of intermediate input sectors that have the highest median input requirement coefficients across all four digit destination sectors in the total (not just import flows) Mexican input output table. Table A.4 in the Appendix provides an overview of these sectors. I then construct an instrument for the overall weighted average intermediate input tariff cut using the weighted average across these commonly shared input categories for each product group. The IV estimation results reported in Column 4 of Table 1 confirm both the size and statistical significance of the OLS point estimate of β_2 which provides evidence against endogeneity concerns of the tariff treatments.

Table 1: Testing the Average Effect of Input Tariffs on Mexican Store Prices 1994-2000

Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Change ln(Store Price) 1994-00	OLS	OLS Baseline	OLS Controls	IV	OLS No Identical Items	Placebo 1 Only Identical Items	Placebo 2 1989-1993	Placebo 3 1994-1995
ln(Store Price 1994)	-0.183*** (0.0248)	0.271 (0.250)	0.292 (0.229)	0.336 (0.264)	0.366 (0.252)	-1.054 (0.966)	-0.232 (0.205)	0.0417 (0.0690)
ln(Store Price 1994) * Change Intermed Imp Tariff 93-00		0.0379* (0.0214)	0.0442** (0.0199)	0.0434* (0.0222)	0.0456** (0.0215)	-0.0414 (0.0810)	-0.00894 (0.0163)	0.00834 (0.00570)
ln(Store Price 1994) * Change Export Tariff 93-00			-0.00325 (0.00452)					
ln(Store Price 1994) * Change Final Imp Tariff 93-00			-0.00285 (0.00639)					
City Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Product Group Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City-By-Product Group FX	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City-By-Barcode FX	No	No	No	No	No	Yes	No	No
Obs	13,589	13,589	13,589	13,589	11,143	13,589	11,029	13,589
1 st Stage F-Stat				735.35				
N(Product Groups)	144	144	144	144	144	144	144	144
Within-R ²	0.093	0.099	0.099		0.101	0.211	0.070	0.042

All regressions include city-by-product group fixed effects. Intermediate tariff changes are weighted averages in percentage points, where weights are total (direct and indirect) requirement coefficients across four digit sectors in the import matrix of the Mexican IO table. Final product import and export tariffs are at the HS six digit level. Instrumental variable results are 2nd stage estimates after instrumenting for input tariff changes with the weighted average tariff changes of the 20% of input sectors with the highest median input requirement across destination sectors in the Mexican IO table. In Columns 5 and 6 “Identical Items” refers to multiple counts of the identical product within a city. Standard errors are clustered at the level of 144 final product groups. ***1%, **5%, and *10% significance levels.

As noted in the data section, one potentially important limitation of the Mexican store price microdata is that store identifiers cannot be recovered from the unique item identifiers. This gives rise to the concern that the estimated pattern of price dynamics within city-product group cells could be driven by relative price changes across stores rather than across vertically differentiated products. Fortunately, the detailed product descriptions can be used to estimate a robustness check on this question. Column 5 reports estimation results after excluding multiple counts of the same barcode items within a city.³⁹ In turn, Column 6 reports results after including city-by-barcode fixed effects so that the estimation is restricted to variation across multiple counts of identical barcode items within a city. The fact that the exclusion of multiple product counts slightly increases both precision and size of the point estimate of β_2 , and that no effect is found on identical items in Column 6, provide assurance that the observed tariff effects are driven by relative price changes across vertically differentiated products, rather than omitted changes in store markups.

In a final set of robustness checks, I estimate two additional placebo falsification tests. Columns 7 and 8 report regressions of store price changes during 1989-1993 and during 1994-1995 on NAFTA tariff changes 1993-2000. The first falsification test reported in Column 7 is estimated off price changes during the preceding four year period during which Mexican import tariffs remained practically unchanged (e.g. Kate, 1992). The second falsification test reported in Column 8 is estimated off price changes between the 3rd quarter of 1994 and the first quarter of 1995 which captures the spike of inflation that occurred in Mexico in the immediate aftermath of the Peso crisis in December 1994 (see Appendix Figure A.3).

These specifications address two particular concerns. The first is that NAFTA tariff changes might be associated to particular product groups that, in general, are characterized by different price distributional changes across stores and/or barcode items. The fact that the point estimate of the tariff interaction in Column 7 is close to zero and insignificant in the preceding period of price changes provides evidence against this concern. Second, tariff changes might have been correlated with product groups whose price distributions were differently affected by the Peso crisis. The fact that the β_2 point estimate in Column 8 is close to zero and insignificant provides evidence against this concern.

5.2 Testing the Heterogeneity of Tariff Effects

To test the model's prediction on the heterogeneity of the observed average effect of input tariff changes in (7), I extend the baseline specification (12) in the following way:

$$\begin{aligned} \ln p_{ick}^{94-00} = & \alpha_{ck} + \beta_1 \ln p_{ick}^{94} + \beta_2 \ln p_{ick}^{94} * d\tau_k^{93-00} \\ & + \beta_3 \ln p_{ick}^{94} * Tech_k + \beta_4 \ln p_{ick}^{94} * d\tau_k^{93-00} * Tech_k + \varepsilon_{ick} \end{aligned} \quad (13)$$

³⁹For items with multiple counts, one of the identical product items was randomly selected to remain in the regression in order not to waste information.

β_1 captures the average relative price growth between product items with initially higher versus lower unit values within city-by-product group cells. β_2 captures the effect of intermediate input tariff cuts on this relative price change among the non-differentiated product groups in the reference category ($Tech_k = 0$). β_3 captures how on average price changes of initially higher and lower unit values differ between non-differentiated and differentiated ($Tech_k = 1$) sectors. The coefficient of interest, β_4 , captures how the relative price effect of intermediate tariff cuts differs in differentiated product groups relative to the non-differentiated reference category. The prediction of the model outlined above is that $\beta_4 > 0$.

5.2.1 Estimating Technology Parameters from Plant Microdata

To empirically estimate differences in sectoral scopes for quality differentiation captured by the $Tech_k$ indicator in (8), I follow the model’s estimation equation in (13). In particular, I estimate the following specification separately across six digit industries in the 1994 monthly plant surveys in order to parameterize the final term $\left(\frac{\partial \ln p_{ki}}{\partial \ln s_{ki}}\right)$ in (8):

$$\ln p_{igmkrt} = \alpha_{rt} + \alpha_{mt} + \beta \ln s_{igmkrt} + \varepsilon_{igmkrt} \quad (14)$$

Subscript i indexes a plant-product line combination, g indexes a plant, m indexes several thousands of manufacturing product groups, k indexes six digit production sectors, r indexes 32 Mexican states, and t indexes 12 months in 1994. $\ln s_{igmkrt}$ are log monthly sales of different product lines within a plant. Log monthly output unit values are thus regressed on product group-by-month fixed effects, state-by-month fixed effects, and the product line’s log sales.⁴⁰

Following from (8), the β coefficient yields an estimate of $\left(\frac{1}{\sigma-1} \frac{\eta_k}{1-\eta_k}\right)$, either pooled across all product groups, or estimated individually for each six digit manufacturing sector. In order to empirically distinguish differentiated sectors where initial unit values in 1994 are positively related to product quality and quality adjusted productivity, and non-differentiated sectors, where no such relationship is observed, I define a binary identifier variable $Tech_k$ which takes the value 1 if β is statistically significantly positive at the 10% level.^{41,42}

To address the concern of correlated error terms within the same product category,

⁴⁰By construction of the plant surveys, sectoral fixed effects would be redundant in (14) because a plant cannot be assigned to product codes in different six digit sectors. In such cases, a novel product group is added in the principal six digit sector. See also Iacovone (2008) for a discussion of the EIM plant data.

⁴¹Kugler and Verhoogen (2011) estimate the scope for differentiation pooling across product groups within four digit industries, rather than at six digit level. I also replicate all subsequent results using the alternative four digit level of aggregation to assign product differentiation. Results are unaffected in size or statistical significance.

⁴²The model relates differences in the magnitude of the unit value-sales elasticity to the scope for differentiation captured by the parameter ψ_k . As a robustness check I follow Kugler and Verhoogen (2011) and verify that the technology estimates are related to existing "off-the-shelf" measures of vertical differentiation in statistically significant way. Appendix Figure A.4 provides graphs and Table A.6 reports regression results in which either the log sales regressor or intermediate input shares in Figure 1 are interacted with Sutton’s (1998) measure of differentiation in terms of advertising and R&D intensities across sectors.

standard errors are clustered at the level of m product groups.⁴³ Finally, following Deaton (1988), both unit values on the left hand side and sales on the right embody measurement error in prices. To address the concern of non-traditional measurement error, I follow Kugler and Verhoogen (2011) and instrument for a product line’s log monthly sales by the log of its establishment’s employment in 1994.

Table A.5 presents the pooled unit value-sales elasticities for all reporting plants, as well as for final goods sectors, and for differentiated final goods industries only. Unit value-sales elasticities are estimated in OLS for log sales and log employment, as well as by IV when instrumenting for log sales with employment in the third column. Out of the 203 reporting manufacturing six digit industries, 79 sectors can be matched to final consumption product groups present in the consumer surveys and/or store samples. Differentiated sectors refer to product groups within six digit sectors for which the unit value-sales elasticity estimate is statistically significantly greater than zero at the 10% level. This cutoff identifies about one third of processed tradable household consumption as differentiated.⁴⁴

5.2.2 Results and Robustness

The estimated technology parameters allow me to estimate regression specification (13) in order to test the prediction on the product group heterogeneity of the tariff effect. In particular, the plant data estimates in Table A.5 suggest that product groups significantly differ in the degree to which observed unit values in 1994 are related to differences in quality and plant technologies. The results reported in Table 2 confirm the prediction. In particular, the first interaction term (β_2) in (13) becomes statistically insignificant, indicating that the previously estimated average effect of intermediate import access on within product group store prices in Table 1 is indeed driven by differentiated product groups.

The point estimate of β_4 is confirmed in size and statistical significance in Column 2 when instrumenting for intermediate input tariff cuts with the weighted average input tariff cut across the 20% of input sectors with the highest median requirement across destination sectors in the Mexican input output table. Finally, Columns 3-6 report the identical battery of robustness tests as discussed for the average tariff effect in Table 1. In particular, the point estimate of β_4 slightly increases when excluding identical barcode items, and it becomes close to zero and statistically insignificant when estimated off price changes in the preceding period, or price changes in the immediate aftermath of the Peso crisis.

⁴³Reported results are unaffected by clustering at the plant level instead.

⁴⁴This proportion is close to identical when choosing to pool all product groups in the same four digit industry to estimate the technology parameters, rather than the less aggregated six digit level chosen here.

Table 2: Testing the Heterogeneity of Tariff Effects

Dependent Variable: Change ln(Store Price) 1994-00	(1) OLS 1994-2000	(2) IV 1994-2000	(3) OLS No Identical Items	(4) Placebo 1 Only Identical Items	(5) Placebo 2 1989-1993	(6) Placebo 3 1994-1995
ln(Store Price 1994)	0.135 (0.310)	0.0678 (0.268)	0.191 (0.311)	-1.115 (0.807)	-0.327 (0.388)	-0.0629 (0.104)
ln(Store Price 1994) * Change Intermed Imp Tariff 93-00	0.0246 (0.0252)	0.0191 (0.0211)	0.0292 (0.0253)	-0.0767 (0.0687)	-0.0164 (0.0296)	3.18e-05 (0.00817)
ln(Store Price 1994) * Differentiated Prod Dummy	0.676* (0.402)	0.670* (0.389)	0.696* (0.394)	1.254 (1.051)	0.200 (0.466)	0.211 (0.131)
ln(Store Price 1994) * Change Intermed Imp Tariff 93-00 * Differentiated Prod Dummy	0.0671** (0.0337)	0.0662** (0.0325)	0.0678** (0.0333)	0.112 (0.0896)	0.0159 (0.0373)	0.0168 (0.0106)
City Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Product Group Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
City-By-Product Group Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
City-By-Barcode FX	No	No	No	Yes	No	No
Obs	13,589	13,589	11,143	13,589	11,029	13,589
1 st Stage FStat		165.94				
N(Product Groups)	144	144	144	144	144	144
Within-R ²	0.108		0.110	0.219	0.070	0.044

All regressions include city-by-product group fixed effects. Intermediate tariff changes are weighted averages in percentage points, where weights are total (direct and indirect) requirement coefficients across four digit sectors in the import matrix of the Mexican IO table. Instrumental variable results are 2nd stage estimates after instrumenting for input tariff changes with the weighted average tariff changes of the 20% of input sectors with the highest median input requirement across destination sectors in the Mexican IO table. Quality differentiation is estimated from observed unit value-sales elasticities in Mexican plant microdata across 12 months in 1994. Standard errors are clustered at the level of 144 final product groups. In Columns 3 and 4 “Identical Items” refers to multiple counts of the identical product within a city. ***1%, **5%, and *10% significance levels.

5.3 Testing Reallocation Effects on Product Entry and Exit

Finally, to test for the effect of tariff cuts on the reallocation of market shares stated in (9), I draw on detailed monthly records of product additions and replacements in the central bank store price microdata. In particular, I estimate logit regressions of the form:

$$\begin{aligned} \text{Entry}_{ikc}^{94-00} \text{ or } \text{Exit}_{ikc}^{94-00} = & \alpha_{ck} + \beta_1 \ln p_{ick}^t + \beta_2 \ln p_{ick}^t * d\tau_k^{93-00} \\ & + \beta_3 \ln p_{ick}^t * \text{Tech}_k + \beta_4 \ln p_{ick}^t * d\tau_k^{93-00} * \text{Tech}_k + \varepsilon_{ick} \end{aligned} \quad (15)$$

$\text{Entry}_{ikc}^{94-00}$ and $\text{Exit}_{ikc}^{94-00}$ are a binary indicators of reported product additions or disappearances over the period 1994-2000 respectively. Superscript t indicates the third quarter in 1994 when the dependent variable is $\text{Exit}_{ikc}^{94-00}$, and the third quarter in 2000 when the dependent variable is $\text{Entry}_{ikc}^{94-00}$. Exit propensities are thus estimated as a function of initial prices, whereas entry propensities are estimated as a function of prices in 2000.

Table 3 reports logit estimation results. The estimation results provide empirical support of the model's predictions on market share reallocations towards the higher end of the quality spectrum. While for both entry and exit regressions in Columns 1 and 3, the average effect of tariff cuts across all product groups is not statistically significant, the tariff effect is significant and of expected opposite sign for entry and exit among differentiated product groups. Higher intermediate tariff cuts appear to have increased the propensity of exit at the lower end of the initial price distribution, whereas they increased the entry propensity at the higher end of the price distribution in 2000.

Table 3: Testing Predictions on Market Share Reallocation

	(1)	(2)	(3)	(4)
Dependent Variable: Additions/Replacements 1994-00	Entry	Entry	Exit	Exit
ln(Store Price 2000)	0.0720 (1.039)	1.255 (1.397)		
ln(Store Price 2000) * Change Intermed Imp Tariff 93-00	0.00535 (0.0833)	0.106 (0.110)		
ln(Store Price 2000) * Differentiated Prod Dummy		-3.415 (2.129)		
ln(Store Price 2000) * Change Intermed Imp Tariff 93-00 * Differentiated Prod Dummy		-0.303* (0.171)		
ln(Store Price 1994)			1.065 (1.369)	-6.056** (2.817)
ln(Store Price 1994) * Change Intermed Imp Tariff 93-00			0.0747 (0.109)	-0.491** (0.223)
ln(Store Price 1994) * Differentiated Prod Dummy				13.29*** (3.087)
ln(Store Price 1994) * Change Intermed Imp Tariff 93-00 * Differentiated Prod Dummy				1.031*** (0.243)
City Fixed Effects	Yes	Yes	Yes	Yes
Product Group Fixed Effects	Yes	Yes	Yes	Yes
City-By-Product Group Fixed Effects	Yes	Yes	Yes	Yes
Obs	19,277	19,277	15,591	15,591

The table presents logit regression results. All specifications include city-by-product group fixed effects. Intermediate tariff changes are weighted averages in percentage points, where weights are total (direct and indirect) requirement coefficients across four digit sectors in the import matrix of the Mexican IO table. Quality differentiation is estimated from observed unit value-sales elasticities in Mexican plant microdata across 12 months in 1994. Standard errors are clustered at the level of 144 final product groups. ***1%, **5%, and *10% significance levels.

6 NAFTA's Effect on Cost of Living Inflation across the Income Distribution

In this section, I draw on the household consumption microdata to evaluate the cost of living implications of NAFTA's observed relative price effects across the Mexican income distribution.

6.1 Empirical Strategy

The cost of living expression in (10) requires information on household expenditure share differences across product items in combination with trade induced relative price changes. The empirical strategy combines observed expenditure shares linked to purchase unit values in household consumption surveys with the causal estimate of NAFTA’s relative price effect within product groups from the store price panels. Observed expenditure shares reported for each individual purchase allow the estimation of the first term on the right hand side of (10), while the linked unit purchase prices allow to estimate predicted price growth in the second term as a function of an item’s position in the unit store price distribution. Intuitively, the estimation strategy combines the observed expenditure weighted household unit value differences depicted in Figure 3 with the estimated effect of tariff changes on relative store prices presented in Tables 1 and 2.

It is apparent from (10) that this estimation is subject to bias if errors in predicted price growth are correlated with expenditure share differences between the rich and the poor across product items.⁴⁵ The empirical challenge is that while both survey data on store prices as well as survey data on household purchase prices stem from the identical population of points of purchase in the third quarter of 1994, the collected microdata do not report a direct link between store price items and household expenditure weights across these datasets. In other words, the Mexican microdata on store price panels and household consumption are not available in homescanner format.⁴⁶

The particular concern that arises is that part of the observed unit value differences between rich and poor households depicted in Figure 3 could be driven by price differences across identical items due to rich people consuming at more expensive stores. In that case, the predicted price changes derived from store price regressions are based on initial store unit value differences in 1994 that reflect quality differentiation, whereas observed price differences between rich and poor households could simply reflect differences in store markups of identical items. The resulting bias would lead to an over-estimate of NAFTA’s true implication on differences in household cost of living inflation for the poor relative to the rich.⁴⁷

Fortunately, Mexican consumer surveys contain information that can be used to estimate a robustness test on this concern. In particular, the surveys report point of purchase types (street vendors, markets, convenience and specialized stores, supermarkets and department

⁴⁵Formally, from (10) we get:

$$\left(\sum_I (s_{kiP}^{t0} - s_{kiR}^{t0}) \ln \left(\frac{p_{ki}^{t1}}{p_{ki}^{t0}} \right) \right) = \sum_I (s_{kiP}^{t0} - s_{kiR}^{t0}) \left(\ln \left(\frac{p_{ki}^{t1}}{p_{ki}^{t0}} \right) + \epsilon_{ki} \right) = \sum_I (s_{kiP}^{t0} - s_{kiR}^{t0}) \ln \left(\frac{p_{ki}^{t1}}{p_{ki}^{t0}} \right) + N_I \text{Cov}((s_{kiP}^{t0} - s_{kiR}^{t0}), \epsilon_{ki}), \text{ where } \epsilon_{ki} = \left(\ln \left(\frac{p_{ki}^{t1}}{p_{ki}^{t0}} \right) - \ln \left(\frac{p_{ki}^{t1}}{p_{ki}^{t0}} \right) \right).$$

⁴⁶Recent contributions making use of US homescanner data from AC Nielsen include Broda et al. (2009), Broda and Romalis (2009), Handbury and Weinstein (2011), and Handbury (2012).

⁴⁷Formally, this would lead to positive correlation between expenditure share differences and prediction errors of relative price growth expressed in (10): $\text{Cov}((s_{kiP}^{t0} - s_{kiR}^{t0}), \epsilon_{ki}) > 0$.

stores) alongside household expenditures and product unit values. Table A.7 reports regressions of log purchase unit values on household income per capita quintile dummies both before and after including city-by-product-by store type fixed effects. If store markups were driving unit value differences, then one would expect the inclusion of store type fixed effects to significantly reduce the estimated price gap across income quintiles. The fact that the estimated unit value differences slightly increase due to the inclusion of store type fixed effects in Column 2 provides evidence against this concern.⁴⁸ Several explanations for this finding have been advocated, including the cost of mobility to reach cheaper stores (Caplovitz, 1963), or bulk discounting (Attanasio and Frayne, 2006).

6.2 Baseline Estimation

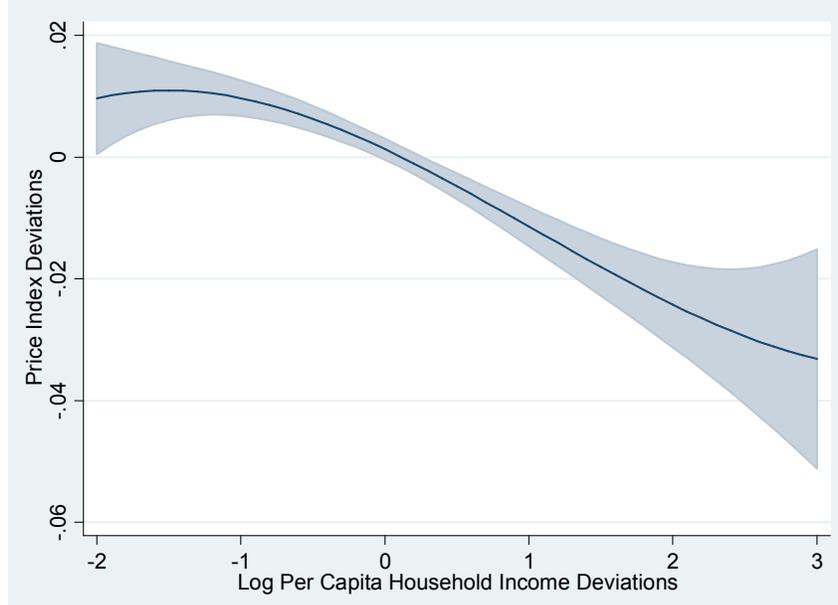
Following these insights, Figure 4 proceeds to present the baseline estimation results of the household price index effect on total tradable consumption of a 12 percentage point US import tariff cut (average of NAFTA tariff cuts 1993-2000) across the urban Mexican income distribution in 1994.⁴⁹ The baseline results are based on observed deviations of household purchase unit values from city-by-product-by-store type means in combination with the preferred estimate of NAFTA’s average relative price effect reported in Column 4 of Table 1.⁵⁰

⁴⁸This finding is consistent with results reported in Broda et al. (2009) using US barcode homescanner data. While that paper’s main conclusion is that on average poorer US households consume at slightly lower prices compared to richer households (incomes above US\$ 60,000 in 2005), Figure 2 of their paper shows that this finding is reversed over the real income range reported in Mexican consumer surveys.

⁴⁹As reported in Table 1, total tradable (i.e. non-services) household consumption accounts for on average 54% of Mexican household consumption in 1994. The store price estimation sample covers processed tradables which account for 70% of total tradable consumption. The reported estimation results are scaled to total tradable consumption, under the (conservative) assumption that no relative price effects occur among tradable products outside the estimation sample.

⁵⁰Since common physical units can be hard to define, consumer surveys report unit values only for food products, beverages, and tobacco products. Out of the 255 processed tradable product groups, 118 report unit values. I assign the weighted average household mean unit value deviation to household expenditures with missing unit information, where the weights reflect the share of household expenditures across product groups with reported unit values. This strategy is likely to be conservative as food and beverages product groups are estimated to be on average less differentiated in the plant microdata. In confirmation of this argument, the reported regressions in Table A.8 of the Appendix show that product groups that are estimated to have higher scopes for quality differentiation in the plant microdata have statistically significantly higher estimated unit purchase value gaps between rich and poor households in the consumption surveys.

Figure 4: Cost of Living Effect of Average NAFTA Import Tariff Cut 1994-2000



The fitted relationship corresponds to the best fitting polynomial functional form according to the Akaike Information Criterion (AIC). Standard errors are clustered at the municipality level and the shaded area indicates 95% confidence intervals. The y-axis depicts mean deviations of estimated household cost of living inflation of tradable consumption due to a 12% tariff cut on US imports. These estimates are based on the average effect of input tariff cuts in Column 4 of Table 1. The x-axis depicts mean deviations of log household per capita incomes. Estimations are based on urban Mexican households in 1994 and subject to nationally representative sample weights. The bottom and top 0.5% on the x-axis are excluded from the graph.

Table 4 presents the same estimates after collapsing the data to mean outcomes across five nominal income quintiles subject to nationally representative household survey weights. The reported result is that the average tariff cut under NAFTA has led to a 2.6 percentage point increase in tradable consumption inflation of the poorest quintile of urban Mexican households compared to the richest quintile.

6.3 Accounting for Product Group Heterogeneity

Because the estimation results in Figure 4 are based on the average store price effect of tariff cuts across product groups in Table 1, the implicit assumption is that all product groups are characterized by the same average scope for quality differentiation. Following the plant data technology estimates in Table A.5 and the heterogeneity of NAFTA's store price effects in Table 2, this assumption is clearly rejected in the data.

How does the observed product group heterogeneity affect the estimation results in Figure 4? In this subsection, I report two alternative estimations which are based on different assumptions about what the observed price differences between rich and poor households within municipality-by-product-by-store type cells measure in terms of differences in product quality choices.⁵¹

⁵¹The underlying empirical challenge is that product market shares are observable in combination with unit values (to estimate quality) only in the plant microdata, while consumption surveys do not report

The first approach is based on the assumption that the observed unit value differences between rich and poor households only reflect differences in product quality choices in product groups in which it is also true that prices are correlated strongly enough with product quality to estimate a statistically significant ($\eta_k > 0$, $Tech_k = 1$) in the plant microdata. That is, we assume that any effect on the relative price of quality is only present in sectors in which we are able to proxy for product quality with unit value differences in the plant data, and for which we find significant observable store price effects in Table 2. This approach thus applies the estimated store price effect in Column 2 of Table 2 only to those household consumption product groups that are estimated to be differentiated in the plant microdata. Since the plant data estimates identify only around 30% of consumer product groups as quality differentiated in this respect, I refer to this estimation approach as a lower bound estimate of NAFTA's true effect on differences in cost of living inflation across the Mexican income distribution.

The second approach is based on the opposite assumption that price differences between the rich and the poor within municipality-by-product-by-store type reflect differences in quality choices across all 255 sample product groups, despite the fact that the relationship between unit values and quality might not be strong enough to be captured in the plant production or store price microdata. The argument is that the plant and store price data include price variation across the full product space, whereas product purchase variation between the rich and poor households within the same store types is more informative to capture quality differences.⁵² This approach thus applies the estimated store price effect among differentiated sectors in Table 2 to observed household consumption price differences across all processed tradable product groups. In the estimation results reported below, I refer to this as an upper bound estimate of NAFTA's effect on differences in household cost of living inflation.

Appendix Figure A.5 presents the estimation results under these alternative assumptions in addition to the baseline estimates that are based on the average tariff effect depicted in Figure 4, and Table 4 presents these estimations after collapsing the data to mean outcomes across five nominal income quintiles subject to nationally representative household survey weights. The lower bound estimate suggests that NAFTA caused a 1.7 percentage point higher cost of living inflation for the poorest income quintile compared to the richest over the period 1994-2000. The upper bound estimate of this effect is 3.9 percentage points. As expected, these alternative estimation approaches fall on different sides of the baseline estimate of 2.6 percentage points.

product barcodes alongside expenditures and unit values.

⁵²Acknowledging noise in the price data, this conception would be fully consistent with the predictions of the model. The larger a given set of households is apart in terms of incomes (and thus quality valuations), the larger should be the signal to noise ratio of quality differences embodied in observed price differences.

Table 4: Cost of Living Effects across the Income Distribution

	Differences in Price Index Effects		
	Baseline	Lower Bound	Upper Bound
2 nd Income Quintile	-0.00388* (0.00228)	-0.00291 (0.00233)	-0.00615* (0.00354)
3 rd Income Quintile	-0.00969*** (0.00272)	-0.00789*** (0.00243)	-0.0149*** (0.00422)
4 th Income Quintile	-0.0136*** (0.00274)	-0.00873*** (0.00286)	-0.0210*** (0.00424)
5 th Income Quintile	-0.0259*** (0.00311)	-0.0170*** (0.00263)	-0.0391*** (0.00475)
Household Obs	7632	7632	7632

The table presents regression results of estimated outcomes on national urban income quintile dummies across 7632 urban Mexican households. Point estimates are based on nationally representative sample weights. Standard errors are clustered at the municipality level. ***1%, **5%, and *10% significance levels. Price index effects are based on reported household unit purchase values in combination with the estimated relative store price effect of US import tariff cuts, and based on the average NAFTA tariff reduction (12 percentage points). “Baseline” is estimated using the observed average relative price effect of input tariff cuts in Column 4 of Table 1. “Lower Bound” is estimated under the assumption that no relative price effects are present in undifferentiated sectors. “Upper Bound” is estimated under the assumption that the observed relative price effect in differentiated sectors (Column 2 of Table 2) operates in all processed tradable product groups.

In terms of the direction and magnitude of NAFTA’s consequences for inequality, the estimated effect reinforces the observed increase in nominal inequality in urban Mexico over the same period, and is equivalent to approximately 25-55% of the total observed difference in nominal income growth between the richest and the poorest income quintiles over the period 1994-2000.⁵³

7 Quality and Quality Adjusted Prices across the Income Distribution

In the final section, I shift attention from the consumer price effects of NAFTA in Mexico to a more general implication of quality sorting by households and plants for real income inequality. In particular, I impose an additional parameter assumption on the elasticity of substitution in estimation equation (8) in order to parameterize η_k across product groups. In the following, I discuss the empirical strategy to use these parameter estimates in order

⁵³This comparison adjusts for the fact tradable consumption accounts for 54% of total consumption in 1994 (i.e. I do not assume similar effects outside the estimation sample). I estimate nominal income growth differences from mean incomes by quintile in 1994 and 2000 using the identical household sample and population weights as in the cost of living estimations reported above. See also Attanasio and Binelli (2010) for a descriptive summary of changes in Mexican income inequality during the 1990s using a number of different microdata and measures of inequality.

to evaluate differences in the consumption of product quality and quality adjusted prices (plant productivity) across the income distribution in 1994.

From (11), the estimation of household differences in weighted average quality and quality adjusted prices requires information on expenditure shares and product characteristics. To estimate (11), I draw on observed unit value distributions in the household consumption surveys in combination with product group estimates of η_k from plant microdata.

The main empirical concern in this estimation is that the technology estimates described in Section 5 are based on factory gate prices in plant microdata, whereas household consumer surveys report point of purchase store prices. In particular, if distribution costs embody not just ad valorem but also per unit cost components then the factory gate price distribution would be compressed in the store price data. Because estimated differences in product quality and plant productivity are based on unit values at the factory gate in the plant microdata, the compression of the price distribution reported in consumer surveys would thus lead to an under-estimation of household consumption differences. In this light, the results reported in this section can be regarded as conservative estimates.⁵⁴

The results reported in Table 5 and Figure 5 are based on a parameter value of $\sigma=2.5$, while I report estimation results across a range of commonly estimated demand parameterizations of $\sigma=2-5$.⁵⁵ As in the cost of living estimations of the previous section, I report estimation results subject to alternative assumptions. As before, I will refer to lower bound estimates when assuming that observed household purchase price differences only reflect quality differences in “differentiated” sectors which are estimated to have a statistically significant relationship between price differences and product quality in the plant microdata in 1994. And, as before, I will refer to upper bound estimations when assuming that purchase price differences between rich and poor households capture product quality differences in all processed tradable product groups (70% of total tradable consumption).

Table 5 presents the results of these estimations after averaging across households in five income quintiles subject to nationally representative household weights. These results are based on observed deviations of household purchase unit values from city-by-product-by-store type means in combination with the estimated unit value-quality elasticities η_k from the plant microdata in Section 5. Among differentiated product groups, the poorest quintile is estimated to consume on average approximately 30% lower weighted average quality among vertically differentiated goods sectors. From (8) this estimate is decreasing in σ and varies between 20-50% across the parameter range $\sigma=2-5$.

Finally, Figure 5 depicts the resulting differences in weighted average quality adjusted

⁵⁴Formally, this leads to a positive correlation between expenditure share differences (poor minus rich) and prediction errors of quality and technology differences. Added noise in product retail prices compared to factory gate prices would have the same effect to under-estimate quality and technology differences embodied in consumption. Furthermore, to the extent not accounted for by store type fixed effects, the same holds true with respect to the estimated propensity of poorer households to consume identical barcodes at slightly higher prices compared to the rich in Table A.7.

⁵⁵This range corresponds to, for example, results reported in Broda et al. (2006).

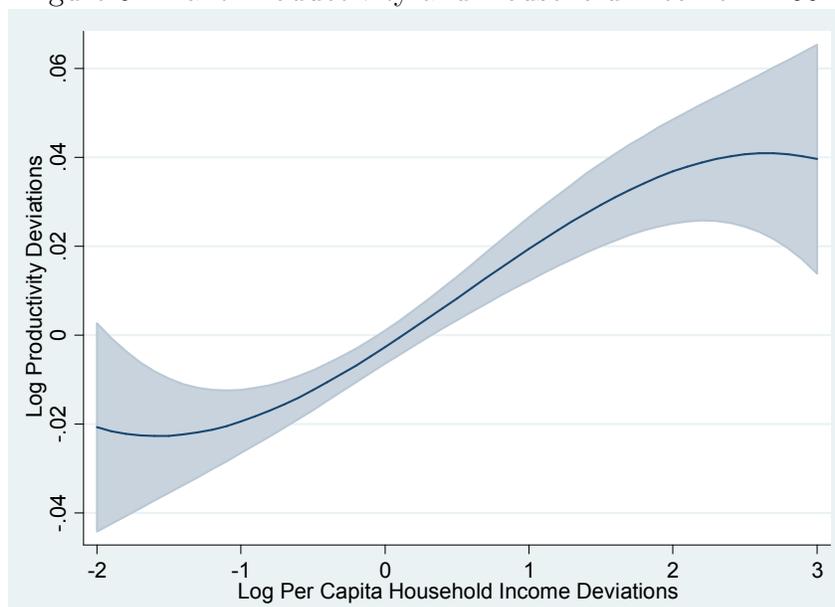
productivity that are embodied in tradable household consumption according to the lower bound estimation. When averaging the quality adjusted productivity differences across five income quintiles reported in Table 5, the poorest quintile is estimated to source tradable consumption at roughly 4.5% lower weighted average productivity compared to the richest quintile. From expression (8) this estimate is decreasing in σ and varies between 2-7% across the parameter range $\sigma=2-5$.

Table 5: Quality and Quality Adjusted Prices across the Income Distribution

	Log Quality Differences		Log Productivity Differences	
	Lower Bound	Upper Bound	Lower Bound	Upper Bound
2 nd Income Quintile	0.0142 (0.0123)	0.0389 (0.0240)	0.00719 (0.00623)	0.0198 (0.0122)
3 rd Income Quintile	0.0423*** (0.0128)	0.103*** (0.0285)	0.0214*** (0.00649)	0.0524*** (0.0145)
4 th Income Quintile	0.0467*** (0.0152)	0.144*** (0.0290)	0.0237*** (0.00769)	0.0733*** (0.0147)
5 th Income Quintile	0.0912*** (0.0142)	0.285*** (0.0339)	0.0463*** (0.00718)	0.145*** (0.0172)
Household Obs	7632	7632	7632	7632

The table presents regression results of estimated outcomes on national urban income quintile dummies across 7632 urban Mexican households. Point estimates are based on nationally representative sample weights. Standard errors are clustered at the municipality level. ***1%, **5%, and *10% significance levels. Quality differences are estimated using reported household unit purchase values in combination with the product group's estimated unit value-quality elasticity (η_k). Productivity deviations are estimated using reported household unit purchase values in combination with the product group's estimated unit value-productivity elasticity ($\frac{\eta_k}{1-\eta_k}$). Both parameter estimates are based on $\sigma=2.5$ in demand. "Lower Bound" is estimated under the assumption that no quality differences in consumption exist in product groups with $\eta_k = 0$ in the plant microdata. "Upper Bound" is estimated under the assumption that the same observed quality differences among sectors with $\eta_k > 0$ exist in all sampled processed consumer good sectors (but nowhere outside this sample).

Figure 5: Plant Productivity and Household Income in 1994



The fitted relationship corresponds to the best fitting polynomial functional form according to the Akaike Information Criterion (AIC). Standard errors are clustered at the municipality level and the shaded area indicates 95% confidence intervals. The y-axis depicts mean deviations of estimated weighted average plant productivities embodied in tradable consumption across households. Weights are reported household expenditure shares. The depicted estimates are based on a parameter value of $\sigma=2.5$ in demand, and according to the “Lower Bound” estimate in the text. The x-axis depicts mean deviations of log household per capita incomes. Estimations are based on urban Mexican households in 1994 and subject to nationally representative sample weights. The bottom and top 0.5% on the x-axis are excluded from the graph.

8 Conclusion

The question of how globalization affects real income inequality in developing countries has been a prominent policy subject in the study of international trade. This paper considers household cost of living in the denominator of real income as a channel through which trade liberalization can affect inequality in a developing country. Drawing on a new collection of microdata covering Mexican households, plants, and stores, I empirically estimate this channel in the context of Mexico’s trade liberalization under NAFTA.

The paper presents evidence in favor of the hypothesis that access to intermediates from developed countries reduces the relative price of higher quality products in a developing country. In turn, because quality choices differ across the income distribution, this relative price effect is found to have significantly increased real income inequality in urban Mexico due to NAFTA over the period 1994-2000. In terms of the direction and magnitude of this effect, NAFTA’s estimated cost of living implications reinforce the observed increase in nominal inequality in urban Mexico over the same period, and are equivalent to at least 25% of the total observed difference in nominal income growth between the richest and the poorest income quintiles.

The paper also highlights that quality choice in a setting with heterogeneous households and plants has a more general implication for real income inequality. In particular, the

observed pattern of quality sorting is such that higher quality products embody cheaper quality adjusted prices, because more productive plants sort into higher quality product lines. The empirical analysis confirms that differences in household quality choices translate into economically significant disparities in weighted average plant productivities and, thus, quality adjusted prices of low income households relative to the rich. This finding suggests that the same observed distribution of nominal incomes leads to more pronounced differences in real incomes in a world with quality sorting in consumption and production compared to a world under conventional assumptions.

For policy analysis, the presented findings serve to highlight the importance of potential price index effects in addition to the conventional focus on nominal incomes when analyzing the general equilibrium consequences of policy or market shocks for the distribution of real incomes. In this respect, the paper points to a number of interesting unanswered research questions concerning, for example, the price index effects of globalization in developed as opposed to developing countries, and the cost of living implications of other policies, such as transport infrastructure or retail sector entry (de-)regulation, in both developing and developed economies.

References

- Aghion, P., & Williamson, J. (1998). *Growth, inequality, and globalization: Theory, history, and policy*. Cambridge Univ Pr.
- Amiti, M., & Konings, J. (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from Indonesia. *The American Economic Review*, 1611–1638.
- Atkin, D. (2010). Trade, tastes and nutrition in India. *Yale University Economic Growth Center Discussion Paper No. 986, Yale Economics Department Working Paper No. 80*.
- Attanasio, O., & Binelli, C. (2010). Mexico in the 1990s: The main cross-sectional facts. *Review of Economic Dynamics*, 13(1), 238–264.
- Attanasio, O., & Frayne, C. (2006). Do the poor pay more? *University College London mimeo*.
- Berry, S. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2), 242–262.
- Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 841–890.
- Bhagwati, J., Panagariya, A., & Srinivasan, T. (1998). *Lectures on international trade*. MIT press.
- Bowen, H., Leamer, E., & Sveikauskas, L. (1987). Multicountry, multifactor tests of the factor abundance theory. *The American Economic Review*, 791–809.
- Broda, C., Greenfield, J., & Weinstein, D. (2006). *From groundnuts to globalization: A structural estimate of trade and growth* (Tech. Rep.). National Bureau of Economic Research.
- Broda, C., Leibtag, E., & Weinstein, D. (2009). The role of prices in measuring the poor's living standards. *The Journal of Economic Perspectives*, 23(2), 77–97.

- Broda, C., & Romalis, J. (2008). Inequality and prices: Does China benefit the poor in America? *University of Chicago mimeo*.
- Broda, C., & Romalis, J. (2009). The welfare implications of rising price dispersion. *University of Chicago mimeo*.
- Caplovitz, D., & of Applied Social Research, C. U. B. (1963). *The poor pay more: Consumer practices of low-income families*. Free Press of Glencoe.
- Choi, Y., Hummels, D., & Xiang, C. (2009). Explaining import quality: The role of the income distribution. *Journal of International Economics*, 78(2), 293–303.
- Cortes, P. (2008). The effect of low-skilled immigration on us prices: Evidence from CPI data. *Journal of Political Economy*, 116(3), 381–422.
- Costinot, A. (2009). An elementary theory of comparative advantage. *Econometrica*, 77(4), 1165–1192.
- Deaton, A. (1988). Quality, quantity, and spatial variation of price. *The American Economic Review*, 418–430.
- Deaton, A. (1989). Rice prices and income distribution in Thailand: A non-parametric analysis. *The Economic Journal*, 99(395), 1–37.
- Deaton, A. (1997). *The analysis of household surveys: A microeconomic approach to development policy*. Johns Hopkins University Press.
- Deaton, A. (2003). Prices and poverty in India, 1987–2000. *Economic and Political Weekly*, 362–368.
- Deaton, A., & Dupriez, O. (2011). Purchasing power parity exchange rates for the global poor. *American Economic Journal: Applied Economics*, 3(2), 137–166.
- Diewert, W. (1979). *The economic theory of index numbers: A survey*. Department of Economics, University of British Columbia.
- Fajgelbaum, P., Grossman, G., & Helpman, E. (2011). Income distribution, product quality, and international trade. *Journal of Political Economy*, 119(4), 721–765.
- Feenstra, R., & Romalis, J. (2012). International prices and endogenous quality. *unpublished paper*.
- Feenstra, R., Romalis, J., & Schott, P. (2002). *Us imports, exports, and tariff data, 1989–2001* (Tech. Rep.). National Bureau of Economic Research.
- Fieler, A. (2011). Nonhomotheticity and bilateral trade: Evidence and a quantitative explanation. *Econometrica*, 79(4), 1069–1101.
- Gagnon, E. (2009). Price setting during low and high inflation: Evidence from Mexico. *The Quarterly Journal of Economics*, 124(3), 1221.
- Gervais, A. (2012). Product quality and firm heterogeneity in international trade. *University of Notre Dame mimeo*.
- Goldberg, P., Khandelwal, A., Pavcnik, N., & Topalova, P. (2010). Imported intermediate inputs and domestic product growth: Evidence from India. *The Quarterly Journal of Economics*, 125(4), 1727–1767.
- Goldberg, P., & Pavcnik, N. (2005). Trade, wages, and the political economy of trade protection: Evidence from the Colombian trade reforms. *Journal of International Economics*, 66(1), 75–105.
- Goldberg, P., & Pavcnik, N. (2007). Distributional effects of globalization in developing countries. *Journal of Economic Literature*, 45, 39–82.
- Goldin, C., & Katz, L. (2008). *The race between education and technology*. Harvard University Press.
- Hallak, J. (2006). Product quality and the direction of trade. *Journal of International*

- Economics*, 68(1), 238–265.
- Hallak, J., & Schott, P. (2011). Estimating cross-country differences in product quality. *The Quarterly Journal of Economics*, 126(1), 417–474.
- Halpern, L., Koren, M., & Szeidl, A. (2011). Imports and productivity. *Central European University mimeo*.
- Handbury, J. (2012). Are poor cities cheap for everyone? Non-homotheticity and the cost of living across US cities. *Columbia University mimeo*.
- Handbury, J., & Weinstein, D. (2011). *Is new economic geography right? Evidence from price data* (Tech. Rep.). National Bureau of Economic Research.
- Hufbauer, G., & Schott, J. (1992). *North American free trade: Issues and recommendations*. Peterson Institute.
- Iacovone, L. (2008). Exploring Mexican plant data. *University of Sussex mimeo*.
- Iacovone, L., Rauch, F., & Winters, L. (2010). *Trade as an engine of creative destruction: Mexican experience with Chinese competition*. London school of economics and political science (LSE). Centre for economic performance (CEP).
- INEGI. (2000). Sintesis metodologica de la encuesta industrial anual. *Instituto Nacional de Estadística, Geografía, e Informática*.
- Johnson, R. (2011). Trade and prices with heterogeneous firms. *Journal of International Economics*.
- Kate, A. (1992). Trade liberalization and economic stabilization in Mexico: Lessons of experience. *World Development*, 20(5), 659–672.
- Khandelwal, A. (2010). The long and short (of) quality ladders. *Review of Economic Studies*, 77(4), 1450–1476.
- Konüs, A. (1939). *The problem of the true index of the cost of living*.
- Krueger, A. (1999). Trade creation and trade diversion under NAFTA. *National Bureau of Economic Research Working Paper*.
- Kugler, M., & Verhoogen, E. (2009). Plants and imported inputs: New facts and an interpretation. In *American economic review papers and proceedings* (Vol. 99, pp. 501–507).
- Kugler, M., & Verhoogen, E. (2011). Prices, plant size, and product quality. *Review of Economic Studies*, *Forthcoming*.
- Leamer, E. (1996). Wage inequality from international competition and technological change: Theory and country experience. *The American Economic Review*, 86(2), 309–314.
- Li, N. (2012). An Engel curve for variety: Household food consumption and welfare in India. *University of Toronto mimeo*.
- López-Córdova, J. (2002). NAFTA and Mexico’s manufacturing productivity: An empirical investigation using micro-level data. *Inter-American Development Bank Working Paper*). IDB, Washington, DC.
- Mandel, B. (2010). Heterogeneous firms and import quality: Evidence from transaction-level prices. *International Finance Discussion Papers*.
- Manova, K., & Zhang, Z. (2012). Export prices across firms and destinations. *The Quarterly Journal of Economics*, 127(1), 379–436.
- Markusen, J. (1986). Explaining the volume of trade: An eclectic approach. *The American Economic Review*, 1002–1011.
- Matsuyama, K. (2000). A ricardian model with a continuum of goods under nonhomothetic preferences: Demand complementarities, income distribution, and North-South trade.

- Journal of Political Economy*, 108(6), 1093–1120.
- Melitz, M. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6).
- Moretti, E. (2011). Real wage inequality. *UC Berkeley mimeo*.
- Muellbauer, J. (1974). Prices and inequality: The United Kingdom experience. *The Economic Journal*, 84(333), 32–55.
- Porto, G. (2006). Using survey data to assess the distributional effects of trade policy. *Journal of International Economics*, 70(1), 140–160.
- Quah, D. (1993). Galton’s fallacy and tests of the convergence hypothesis. *The Scandinavian Journal of Economics*, 427–443.
- Rodriguez, C. (1979). The quality of imports and the differential welfare effects of tariffs, quotas, and quality controls as protective devices. *The Canadian Journal of Economics*, 12(3), 439–449.
- Sato, K. (1976). The ideal log-change index number. *The Review of Economics and Statistics*, 58(2), 223–228.
- Slaughter, M. (1998). What are the results of product-price studies and what can we learn from their differences? *National Bureau of Economic Research Working Paper*, No 6591.
- Sutton, J. (1998). *Technology and market structure: Theory and history*. The MIT Press (Cambridge, Mass).
- Topalova, P. (2010). Factor immobility and regional impacts of trade liberalization: Evidence on poverty from India. *American Economic Journal: Applied Economics*, 2(4), 1–41.
- Trefler, D. (1995). The case of the missing trade and other mysteries. *The American Economic Review*, 1029–1046.
- Vartia, Y. (1976). Ideal log-change index numbers. *Scandinavian Journal of Statistics*, 121–126.
- Verhoogen, E. (2008). Trade, quality upgrading, and wage inequality in the Mexican manufacturing sector. *The Quarterly Journal of Economics*, 123(2), 489–530.

Appendix

Appendix Figures

Figure A.1: Two Notable Features about Mexican and Developing Country Imports

Figure A.1.A: What Do Developing Countries Import?

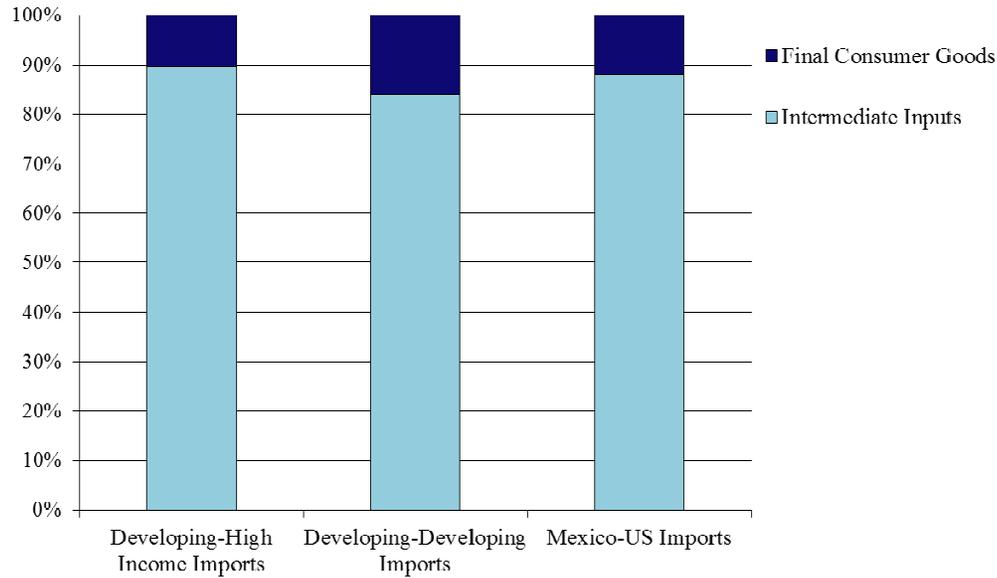


Figure A.1.B: Variation in Use of Imported Inputs

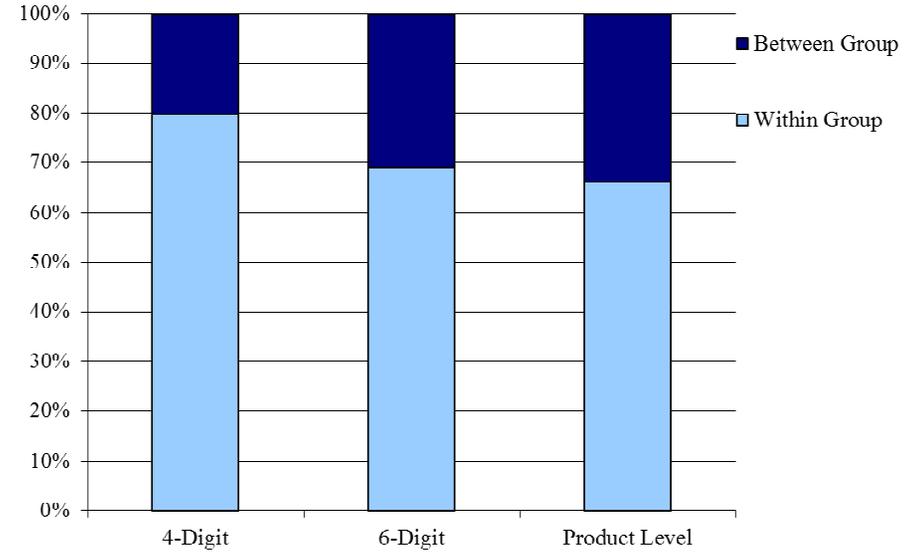


Figure A.1.A depicts end use shares in the sum of imports across different pairs of countries over the period 1994-2000. From left to right the bars depict developing country imports from high income countries, developing country imports from developing countries, and Mexican imports from the US. “Developing” and “High Income” refer to low-and-middle income and high income countries according to the World Bank’s classification in 2010 respectively. Data on import flows are from the World Bank’s WITS database. End-use classifications into final consumption and intermediate goods are based on BEC classifications. Figure A.1.B depicts the variance decomposition of plant level imported input shares into between and within product group components. The estimates are from Mexican plant data in 1994 and conditional on state fixed effects. “4-digit”, “6-digit”, and “Product Level” refer to 80, 203, and roughly 3234 manufacturing product groups respectively.

Figure A.2: Mexican Tariffs on US Imports 1993-2000

Figure A.2.1: Average Applied Tariff Rates

Figure A.2.2: Sectoral Variation

14

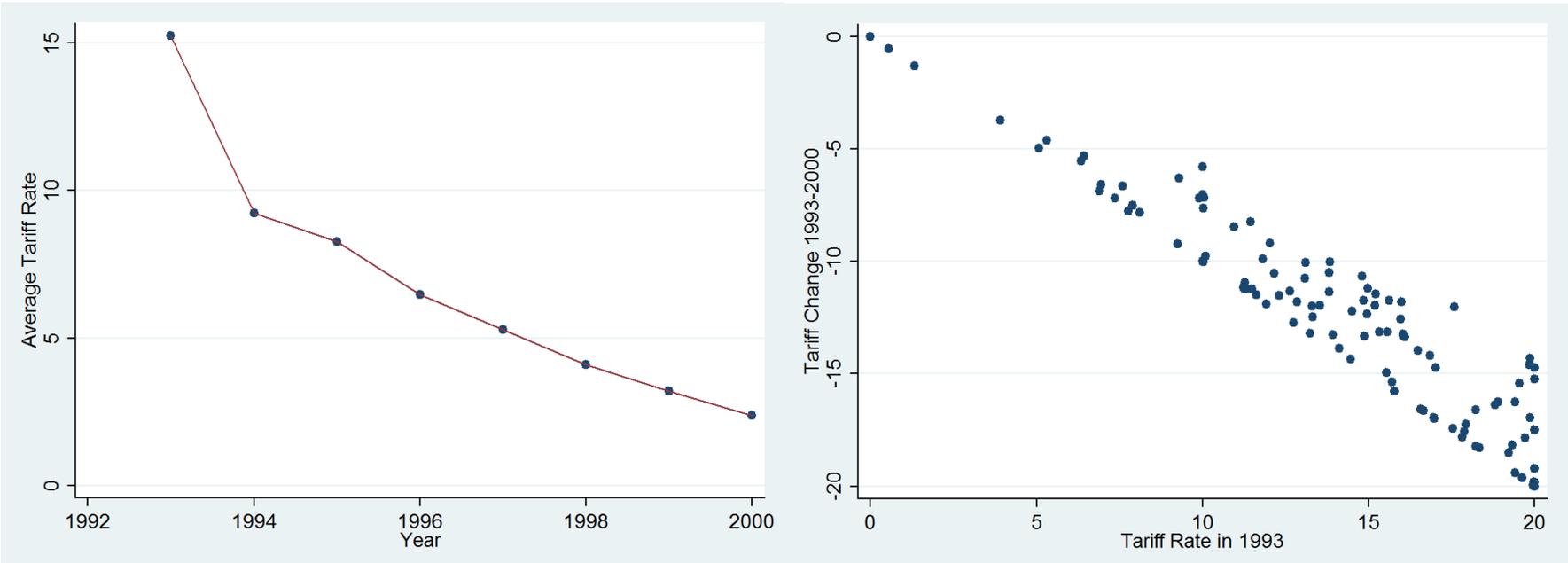
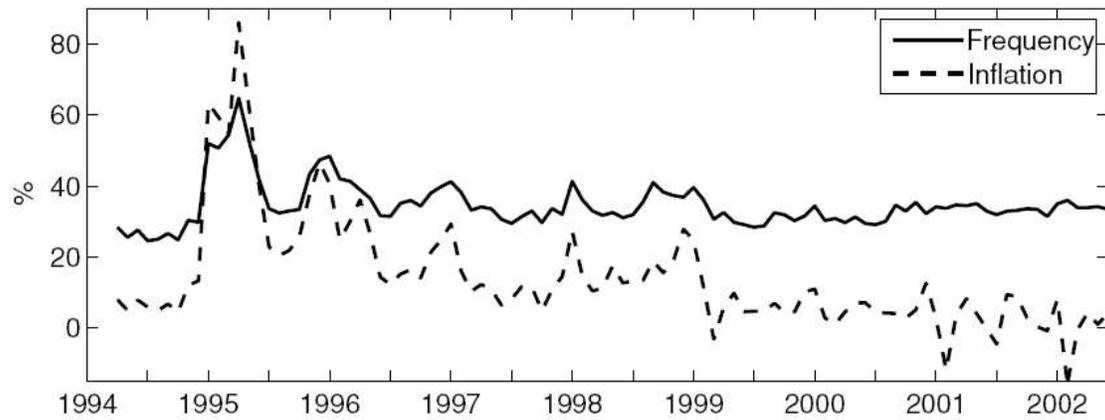


Figure A.2.A depicts average applied tariff rates on US manufacturing imports over the period 1993-2000. Figure A.2.B depicts the relationship between changes and initial levels of average applied tariff rates across four digit US manufacturing sectors between 1993-2000. Source: Secretaria de Economia.

Figure A.3: The Peso Crisis and Spike of Price Changes 1994-1995

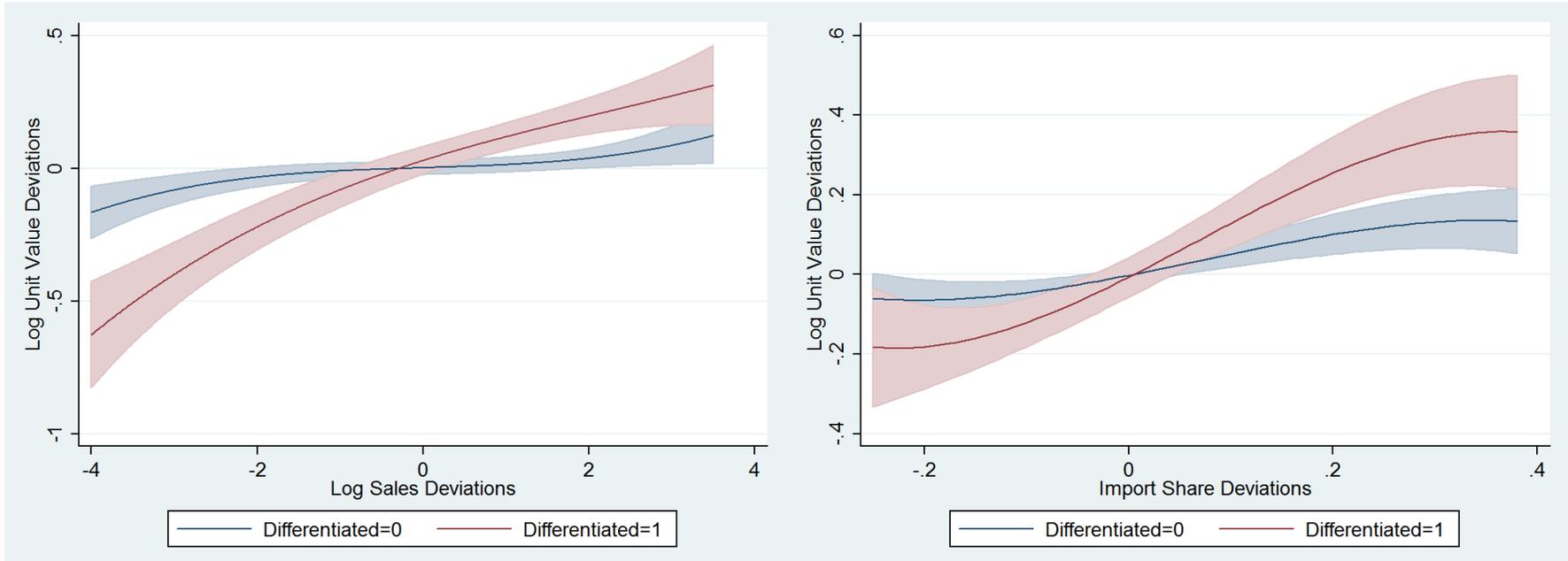


The figure is taken from Gagnon (2009, pp. 1233). It depicts monthly frequencies of price changes and CPI inflation for non-regulated goods and services.

Figure A.4: Unit Value Elasticities and Alternative Measures of Vertical Differentiation

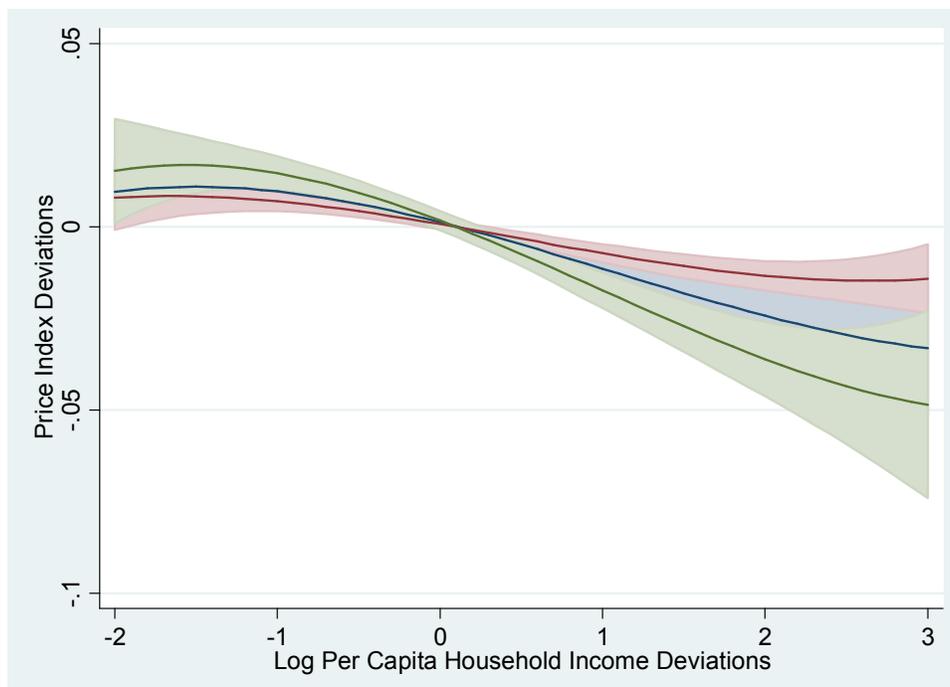
Figure A.4.1: Unit Values and Sales

Figure A.4.2: Unit Values and Imported Input Shares



The fitted relationships correspond to the best fitting polynomial functional form according to the Akaike Information Criterion (AIC). The y-axis in both graphs depict the residuals of a regression of log product-line unit values on month-by-product and month-by-state fixed effects. The x-axis in A.4.1 depicts residuals of a regression of log product line sales on the same fixed effects. The x-axis in A.4.2 depicts residuals of a regression of annual plant level imported input shares on state and product fixed effects. The sample is broken down into four-digit sectors with above or below mean shares of R&D and advertisement over sales following Sutton (1998) (Differentiated=1,0 respectively). The bottom and top 0.5% are excluded from the graph. Standard errors are clustered at the product level and the shaded area indicates 95% confidence intervals.

Figure A.5: Alternative Estimates of NAFTA's Effect on Cost of Living Inflation



The fitted relationship corresponds to the best fitting polynomial functional form according to the Akaike Information Criterion (AIC). Standard errors are clustered at the municipality level and the shaded area indicates 95% confidence intervals. The y-axis depicts mean deviations of estimated household cost of living inflation of tradable consumption due to a 12% tariff cut on US imports. The steepest, middle, and flattest functions are based on “upper bound”, “baseline”, and “lower bound” estimations respectively. Estimations are based on urban Mexican households in 1994 and subject to nationally representative sample weights. The bottom and top 0.5% on the x-axis are excluded from the graph.

Appendix Tables

Table A.1: Coverage of Store Price Sample in Total Household Consumption

	July, August, September 1994	July, August, September 2000	1994-2000
Number of Monthly Store Price Quotes in 35 Cities	28515	40280	
Tradable Sample	24089	33699	
Processed Tradeable Estimation Sample	16792	22235	
Persistent Barcode Series in Estimation Sample	13589	13589	13589
Product Replacements in Estimation Sample			3768
Basket Net Expansion in Estimation Sample			5866
Tradable Coverage in Total Urban Household Consumption (from ENIGH Consumer Surveys)	0.54	0.50	
Processed Tradable Sample Coverage in Tradable Consumption	0.70	0.66	

Table A.2: Plant Data Descriptive Statistics

	Full Sample	Final Good Sectors Only
Number of 6-Digit Sectors	203	79
Number of Establishments	6341	2762
Number of Products Reported Over 12 Months	3234	1331
Number of Month * Establishment *	257736	136440
Product Observations		
Average Number of Products Per Establishment	3.4	4.1
Median Employment Size	103	122

Table A.3: Household Consumption Survey Descriptive Statistics

Urban Household Sample	
Number of Households	7764
Number Of Municipalities	236
Total Number of Reported Transactions Across all Expenditure Categories	524782
Number of Reported Transactions In Processed Tradables Sample (255 Product Groups)	279584
Number of Transactions In Processed Tradables Sample With Unit Values	122572
Share of Processed Tradables Transactions at Markets	0.081
Share of Processed Tradables Transactions at Street Vendors	0.112
Share of Processed Tradables Transactions at Convenience and Specialized Stores	0.485
Share of Processed Tradables Transactions at Supermarkets and Department Stores	0.264

Table A.4: Input Sectors with Highest Median Total Requirement Coefficients across Mexican Destination Sectors

Rank	NAICS 4-Digit Sector	NAICS Description
1	3251	Fabricación de productos químicos básicos
2	3241	Fabricación de productos derivados del petróleo y del carbón
3	3261	Fabricación de productos de plástico
4	3222	Fabricación de productos de papel y cartón
5	3221	Fabricación de celulosa, papel y cartón
6	3363	Fabricación de partes para vehículos automotores
7	3252	Fabricación de hules, resinas y fibras químicas
8	3231	Impresión e industrias conexas
9	3259	Fabricación de otros productos químicos
10	3255	Fabricación de pinturas, recubrimientos, adhesivos y selladores
11	3311	Industria básica del hierro y del acero
12	3399	Otras industrias manufactureras
13	3132	Fabricación de telas
14	3211	Aserrado y conservación de la madera
15	3312	Fabricación de productos de hierro y acero de material comprado

Table A.5: Technology Parameter Estimates from Plant Microdata

Estimation Samples	Dependent Variable: ln(Unit Value)	(1) OLS	(2) OLS	(3) IV
All 6-Digit Manufacturing Sectors	ln(Sales)	0.0486*** (0.00594)		0.0444*** (0.0123)
	ln(Employment)		0.0363*** (0.0102)	
	Obs	170240	167449	160835
	N(Plants)	5779	5665	5665
	Within-R ²	0.017	0.009	
Final Goods 6-Digit Manufacturing Sectors	ln(Sales)	0.0557*** (0.00781)		0.0733*** (0.0159)
	ln(Employment)		0.0598*** (0.0133)	
	Obs	94741	93154	91064
	N(Plants)	2656	2602	2602
	Within-R ²	0.024	0.013	
Differentiated Final Goods 6-Digit Manufacturing Sectors	ln(Sales)	0.104*** (0.0159)		0.240*** (0.0315)
	ln(Employment)		0.209*** (0.0300)	
	Obs	24312	24123	23392
	N(Plants)	672	666	666
	Within-R ²	0.065	0.066	

All regressions include state-by-month and month-by-product fixed effects. Product groups refer to several thousand disaggregate product descriptions within 203 6-digit manufacturing sectors. Unit values and sales vary across plants, product lines within plants, and months. Annual employment varies across plants. The first stage regressions of ln(Sales) on ln(Employment) not reported here are highly statistically significant. Standard errors are clustered at the product level. ***1%, **5%, and *10% significance levels.

Table A.6: Unit Value-Sales Elasticities and Alternative Measures of Vertical Differentiation

Dependent Variable: ln(Unit Value)	(1)	(2)	(3)	(4)	(5)	(6)
ln(Sales)	0.0557*** (0.00781)	-0.00720 (0.00831)				
ln(Sales) * R&D/Advert Intensity		1.091*** (0.126)				
ln(Employment)			0.0598*** (0.0133)	-0.0410*** (0.0150)		
ln(Employment) * R&D/Advert Intensity				1.880*** (0.279)		
Import Share					0.576*** (0.103)	0.00849 (0.108)
Import Share * R&D/Advert Intensity						10.23*** (1.708)
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State*Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Product*Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Obs	94741	94741	93154	93154	93996	93996
N(Plants)	2656	2656	2602	2602	2625	2625
Within-R ²	0.024	0.041	0.013	0.026	0.026	0.026

All regressions include state-by-month and month-by-product fixed effects. “R&D/Advert Intensity” refers to US shares of R&D and advertising expenditures in firm sales, averaged to the four digit SITC industry sectors. These measures are reported in Kugler and Verhoogen (2011) and were matched to Mexican six digit final goods manufacturing industries. Standard errors are clustered at the product level. ***1%, **5%, and *10% significance levels.

Table A.7: Do the Rich and the Poor Consume Identical Items at Different Prices?

Dependent Variable: ln(Unit Value)	(1)	(2)
2 nd Per Capita Household Income Quintile	0.00297 (0.0106)	0.00962 (0.00873)
3 rd Per Capita Household Income Quintile	0.0288** (0.0125)	0.0426*** (0.0110)
4 th Per Capita Household Income Quintile	0.0454*** (0.0148)	0.0573*** (0.0122)
5 th Per Capita Household Income Quintile	0.0803*** (0.0197)	0.0983*** (0.0146)
City Fixed Effects	Yes	Yes
Product Group Fixed Effects	Yes	Yes
City-Product Group Fixed Effects	Yes	Yes
Store Type Fixed Effects	-	Yes
City-Store Type Fixed Effects	-	Yes
Product-Store Type Fixed Effects	-	Yes
City-Product-Store Type Fixed Effects	-	Yes
Obs	122,572	122,572
N(Households)	7632	7632
R ²	0.894	0.929

The dependent variable is reported purchase unit values in 118 out of a total of 255 processed tradable product groups in the 1994 household consumption survey. Income quintiles are based on per capita household incomes. Household consumption surveys include the following store types: Markets, street vendors, convenience and specialized stores, and supermarkets and department stores. Standard errors are clustered at the municipality level. ***1%, **5%, and *10% significance levels.

Table A.8: Rich-Poor Price Gaps and Product Group Differentiation

Dependent Variable: ln(Unit Value)	(1)
ln(Income per Capita)	0.0317*** (0.00559)
ln(Income per Capita)*Scope	0.0750*** (0.0278)
City Fixed Effects	Yes
Product Group Fixed Effects	Yes
City-Product Group Fixed Effects	Yes
Store Type Fixed Effects	Yes
City-Store Type Fixed Effects	Yes
Product-Store Type Fixed Effects	Yes
City-Product-Store Type Fixed Effects	Yes
Obs	122,572
N(Households)	7632
R ²	0.930

The dependent variable is reported purchase unit values in 118 out of a total of 255 processed tradable product groups in the 1994 household consumption survey. Income quintiles are based on per capita household incomes. Household consumption surveys include the following store types: Markets, street vendors, convenience and specialized stores, and supermarkets and department stores. “Scope” is the product group specific estimate of the unit value-sales elasticity observed in the plant microdata in 1994. Standard errors are clustered at the municipality level. ***1%, **5%, and *10% significance levels.

Additional Results of the Model

For ease of exposition product group subscripts k are suppressed unless indicated otherwise.

Preferences

From preferences in (1) of the paper, consumer optimization yields the following expression for household expenditure:

$$p_i x_{hi} = y_{hi} = \frac{p_i^{1-\sigma} (q_i^{\varphi_h})^{\sigma-1}}{\int_{j=0}^N p_j^{1-\sigma} (q_j^{\varphi_h})^{\sigma-1} dj} y_h = s_{hi} y_h$$

s_{hi} is the household’s expenditure share on variety i and y_h is its total expenditure on product group k. Within product group expenditure shares increase in quality valuation for products with above average quality, and decrease in quality valuation for below average quality products:

$$\frac{\partial s_{hi}}{\partial \varphi_h} = (\sigma-1) s_{hi} \left(\ln q_i - \sum_I s_{hi} \ln q_i \right)$$

Weighted average quality of the household’s consumption basket increases in quality valuation:

$$\frac{\partial \left(\sum_I s_{hi} (\ln q_i - \left(\frac{1}{I} \sum_I \ln q_i \right)) \right)}{\partial \varphi_h} = I * \text{Cov} \left(\left(\frac{\partial s_{hi}}{\partial \varphi_h} \right), (\ln q_i - \overline{\ln q_i}) \right) > 0$$

Finally, I solve for the elasticity of market sales with respect to product quality to get:

$$\frac{d(\sum_H s_{hi} y_h)}{(\sum_H s_{hi} y_h)} \bigg/ \frac{dq_i}{q_i} = (\sigma - 1) \left(\sum_H \frac{y_{hi}}{y_i} \varphi_h \right) = (\sigma - 1) \varphi_i^*$$

And following the discussion in the text, I define: $q_i^* = q_i^{\varphi_i^*}$.

Technology

The unit cost functions for production technologies in (3) and (5) of the paper are $c_i = \lambda_i^{-1} p_{im}$ and $c_j = w z_j^\delta$ respectively. Substituting intermediate unit costs into final product unit costs, we get: $c_i = \lambda_i^{-1} w z_i^\delta$. Plant profits are given by:

$$\pi_i = (p_i - c_i) x_i - f = (p_i - \lambda_i^{-1} w z_i^\delta) x_i - f$$

In quality adjusted terms, profits can be written as:

$$\pi_i = \left(\frac{p_i}{q_i^*} - \frac{c_i}{q_i^*} \right) q_i^* x_i - f = \left(\frac{p_i}{q_i^*} - \lambda_i^{-1} w z_i^\delta q_i^{*-1} \right) q_i^* x_i - f$$

Equilibrium

Firms simultaneously choose product quality and quality adjusted prices to maximize profits. From the profit equations it is clear that maximizing profits with respect to product quality, implies minimizing $\frac{c_i}{q_i^*}$ with respect to q_i^* , that is minimizing the average variable cost per unit of product quality.⁵⁶ The additional parameter restriction $\gamma > \delta$ assures a well behaved optimum. This provides an expression for equilibrium product quality as a function of technical efficiency:

$$q_i^* = \left(\frac{\alpha \gamma}{\gamma - \delta} \right)^{1/\theta} \lambda_i^\psi$$

We next solve for intermediate input quality as a function of plant efficiency:

$$z_i = \left(\frac{1}{1 - \alpha} \left(\frac{\alpha \gamma}{\gamma - \delta} \right)^{\psi/\theta} - \frac{\alpha}{1 - \alpha} \right)^{1/\theta \gamma} \lambda_i^{\psi/\gamma}$$

Equilibrium intermediate input quality is thus increasing in plant efficiency. Alternatively, we can solve for λ_i as a function of product quality and substitute back into the unit cost function to derive the equilibrium relationship between final product quality and unit costs:

$$c_i = w \left(\frac{1}{\alpha} - \frac{\delta}{\alpha \gamma} \right)^{-1/\theta \psi} \left(\frac{1}{1 - \alpha} - \frac{\delta}{\alpha \gamma} \right)^{\delta/\theta \gamma} q_i^{*\eta}$$

$\eta = \frac{\delta}{\gamma} - \frac{1}{\psi}$ is the elasticity of unit costs with respect to final product quality. The equilibrium relationship between product quality and the inverse of quality adjusted marginal costs becomes:

⁵⁶The same condition is present in Mandel (2010), Johnson (2011), and Feenstra and Romalis (2012). The latter paper attributes this finding to Rodriguez (1979).

$$\frac{\partial \ln q_i^*}{\partial \ln (q_i^*/c_i)} = \frac{1}{1 - \eta}$$

Given CES preferences, the equilibrium relationship between observed unit values and product quality can then be expressed as:

$$p_i = \frac{\sigma}{\sigma - 1} c_i = \frac{\sigma}{\sigma - 1} w \left(\frac{1}{\alpha} - \frac{\delta}{\alpha \gamma} \right)^{-1/\theta \psi} \left(\frac{1}{1 - \alpha} \frac{\delta}{\gamma} \right)^{\delta/\theta \gamma} q_i^{*\eta}$$

Finally, the equilibrium relationship between unit values and quality adjusted productivity becomes:

$$\frac{\partial \ln p_i}{\partial \ln (q_i^*/c_i)} = -1 + \frac{\partial \ln q_i^*}{\partial \ln (q_i^*/c_i)} = \frac{\eta}{1 - \eta}$$

Starting from an initial equilibrium outcome of input quality and output quality choices across plants, the observed final product unit value elasticity with respect to product quality is as derived in the above: $\frac{\partial \ln p_i}{\partial \ln q_i^*} = \frac{\partial \ln c_i}{\partial \ln q_i^*} = \eta = \frac{\delta}{\gamma} - \frac{1}{\psi}$. Import access then lowers the relative cost of inputs with higher foreign value shares which is captured by a change in δ . Denote import tariffs by τ_k , then the cross-derivative expression of interest becomes:

$$\frac{\partial^2 \ln p_i}{\partial \ln q_i^* \partial \tau} = \frac{\partial^2 \ln c_i}{\partial \ln q_i^* \partial \tau} = \frac{1}{\gamma} \frac{\partial \delta}{\partial \tau} > 0$$

The general equilibrium solution of the model closely follows Melitz (2003). To assure finite means in efficiency draws and final product plant revenues, the shape parameter of the pareto distribution needs to have a lower bound at $\xi > \max(\psi(\sigma - 1)(1 - \eta), 1)$. The cut-off values are determined by two conditions.

First, profits of the marginal plant must be zero: $\pi(\lambda^*) = \frac{r^*(\lambda)}{\sigma} - f = 0$. Second, free entry implies that ex ante expected profits are zero: $(1 - G(\lambda^*)) \sum_{t=0}^{\infty} (1 - \chi)^t \left(\frac{E(r^*(\lambda))}{\sigma} - f \right) - f_e = 0$. Using these two conditions, and that $\frac{r^*(\lambda)}{r^*(\lambda^*)} = \left(\frac{\lambda}{\lambda^*} \right)^{\psi(\sigma - 1)(1 - \eta)}$, we get: $E(r^*(\lambda)) = \frac{\xi}{\xi - \psi(\sigma - 1)(1 - \eta)} \sigma f$. It follows that:

$$\lambda^* = \lambda_m \left(\frac{f}{f_e \chi} \left(\frac{\xi}{\xi - \psi(\sigma - 1)(1 - \eta)} - 1 \right) \right)^{1/\xi}$$

Finally, the free entry condition in combination with the condition that in steady state the mass of new entrants is equal to the mass of exiting firms $M_e(1 - G(\lambda^*)) = \chi M$, and labor market clearing ($L = M E r^*(\lambda) - \Delta + M_e f_e$), where L is labor supply and Δ is the difference between final sector revenues and profits, pin down the mass of final good producers in steady state:

$$M = \frac{L(\xi - \psi(\sigma - 1)(1 - \eta))}{\xi \sigma f}$$

Cost of Living Implications

Based on the work of Sato (1976) and Vartia (1976), the ideal price index for a homothetic CES utility function is:

$$\left(\frac{e(u^*, p^{t1})}{e(u^*, p^{t0})}\right)_h = \prod_I \left(\frac{p_{ki}^{t1}}{p_{ki}^{t0}}\right)^{\omega_{hki}}, \text{ where } \omega_{hki} = \frac{s_{hki}^{t1} - s_{hki}^{t0}}{\ln(s_{hki}^{t1} - s_{hki}^{t0})} / \left(\sum_I \frac{s_{hki}^{t1} - s_{hki}^{t0}}{\ln(s_{hki}^{t1} - s_{hki}^{t0})}\right)$$

I is the number of all varieties pooled across all product groups k in the economy. Household cost of living inflation is a weighted geometric mean of price changes where the weights are ideal log changes of household budget shares. In the following, I will refer to two representative consumers that can be thought of as a poor and a rich household denoted by subscripts P and R. Taking log differences in household cost of living inflation between a poor and a rich household, we then get:

$$\ln\left(\frac{e(u^*, p^{t1})}{e(u^*, p^{t0})}\right)_P - \ln\left(\frac{e(u^*, p^{t1})}{e(u^*, p^{t0})}\right)_R = \sum_I (s_{kiP}^{t0} - s_{kiR}^{t0}) \ln\left(\frac{p_{ki}^{t1}}{p_{ki}^{t0}}\right)$$

which is (10) in the text. In the presence of non-homotheticity in (1), the Sato-Vartia ideal price index, in principle, does not hold because income changes affect expenditure shares so that the ideal weights (ω_{hki}) cease to hold (e.g. Diewert, 1979). Within the structure of the model, however, (10) represents the difference in the exact ideal price index between two representative agents due to a *ceteris paribus* change in the relative price of quality. The two underlying assumptions are that i) CES preferences in (1) hold so that elasticities of substitution are the same across households, and ii) we abstract from general equilibrium consequences of import access on relative incomes. If either of the model's assumptions is violated, then (10) remains an approximation of the true difference in cost of living to the first order, because as in a Laspeyres price index its weights are based on differences in initial expenditure weights.

The second welfare distributional implication concerns price levels rather than price changes. Using the same notation as above, we start with an expression for log differences in weighted average product quality between a poor and a rich household denoted by $\ln Q_P - \ln Q_R$:

$$\ln Q_P - \ln Q_R = \sum_I (s_{kiP}^{t0} - s_{kiR}^{t0}) \left(\ln q_{ki}^* - \overline{\ln q_{ki}^*}^k \right)$$

Substituting for product quality by the equilibrium relationship to quality adjusted costs, we get:

$$\ln Q_P - \ln Q_R = \sum_K \frac{1}{1 - \eta_k} \sum_{Ik} (s_{kiP}^{t0} - s_{kiR}^{t0}) \left(\ln \left(\frac{q_{ki}^*}{c_{ki}} \right) - \overline{\ln \left(\frac{q_{ki}^*}{c_{ki}} \right)}^k \right)$$

Denoting differences in weighted average inverse quality adjusted prices by $\ln \left(\frac{Q}{P} \right)_P - \ln \left(\frac{Q}{P} \right)_R$, we get:

$$\ln \left(\frac{e(u^*, p^{t1})}{e(u^*, p^{t0})} \right)_P - \ln \left(\frac{e(u^*, p^{t1})}{e(u^*, p^{t0})} \right)_R = \ln \left(\frac{Q}{P} \right)_P - \ln \left(\frac{Q}{P} \right)_R = \sum_K (1 - \eta_k) (\ln Q_{Pk} - \ln Q_{Rk})$$

Product Group Details and Concordance Tables

Details and concordance tables of the consumer product groups used in the estimations of the paper are provided in the Online Appendix.