ESSAYS ON INTERNATIONAL ECONOMICS AND TRADE

Marco Errico

A dissertation submitted to the Faculty of
the Department of Economics in partial
fulfillment of the requirements for the
degree of Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

April 2023

ESSAYS ON INTERNATIONAL ECONOMICS AND TRADE

Marco Errico

Advised by:

Jaromir Nosal, Ph.D.

James E. Anderson, Ph.D.

Danial Lashkari, Ph.D.

Rosen Valchev, Ph.D.

Abstract

This dissertation comprises three self-contained essays that investigate the determination and transmission of exchange rate fluctuations, as well as the impact of import quality on consumers' gains from globalization.

In the first chapter, "Decomposing the (In)Sensitivity of CPI to Exchange Rate", I examine the role of domestic frictions – distribution costs, variable markups and nominal rigidities – in explaining the low sensitivity of domestic prices to exchange rate fluctuations. I begin by modeling what the sensitivity of CPI to exchange rates is expected to be, given the presence of insensitivity in border prices and domestic frictions. Distribution costs, such as transportation and wholesaling costs, introduce a wedge between the retail price, on one side, and the border price of imports and the domestic producers' costs, on

the other. Similarly, domestic firms do not fully adjust their price to changes in their own cost because of changes in the desired markup or because prices are sticky. These frictions introduce wedges between the change in domestic producers' costs and border prices following an exchange rate shock, and the response of domestic consumption retail prices. Using firm and transaction data from Chile, I document that domestic frictions account for 60% of the overall insensitivity of domestic CPI. Moreover, the presence of domestic frictions also impacts the sensitivity of domestic CPI: contrary to previous literature, most of the sensitivity arises from the direct consumption of imported final goods, rather than through the costs associated to imported inputs in the production of domestic goods. This is because domestic frictions dampen the response of domestically produced goods more significantly. In addition, I quantify a rich heterogeneity in the sensitivity across products, which stems from the interaction of domestic frictions and import exposure. These heterogeneities are relevant for the overall (in)sensitivity, as sectors with higher import exposure face also larger frictions. Overall, my results showcase the importance of domestic frictions and their heterogeneity in studying the response of domestic prices to exchange rate fluctuations, with implications for monetary policy in open economy and redistribution dynamics.

In the second chapter, "Strategic Behavior and Exchange Rate Dynamics", joint work with L. Pollio, I examine the impact of heterogeneous investors with different degrees of price impact on exchange rate behavior. The huge trading volume in the currency markets, about \$6 trillions per day, is highly concentrated among the market-making desks of few large financial institutions. However, models of exchange rate determination assume that investors take the equilibrium price as given, ignoring the presence of a few large investors who recognize the price impact of their decisions and can exert pressure on market prices. We incorporate heterogeneity in price impact, following of Kyle (1989), into a two-country, dynamic monetary model of exchange rate determination. Our theory of exchange rate determination with heterogeneity in price impact reveals that market structure is a key determinant of exchange rate dynamics. Strategic investors recognize their price impact, which leads them to trade less on any information and reduce the information loading factor of the exchange rate (price informativeness). The presence of strategic investors explains the weak explanatory power of macroeconomics variables in predicting exchange rates (exchange rate disconnect puzzle) and the excess volatility of the exchange rate relative to fundamentals (excess volatility puzzle). We also provide empirical evidence that supports our theoretical predictions by using trading volume concentration data from the NY Fed FXC Reports for 18 currencies from 2005 to 2019. We extend our theoretical framework to include another dimension of heterogeneity among investors, information heterogeneity, that provides similar qualitative predictions in terms of exchange rate dynamics. We demonstrate that both dimensions of heterogeneity are quantitatively relevant in explaining the disconnect of exchange rates and their excess volatility.

In the third chapter, “The Quality of US Imports and the Consumption Gains from Globalization", joint work with D. Lashkari, I examine the role of quality improvement in shaping the gains from trade. The existing empirical literature indicates that globalization has offered consumers around the world access to a wider variety of products at cheaper prices. However, since the available data typically lacks detailed information on product characteristics, we may underestimate the value of imports for consumers if the quality of goods within each product rises over time. To overcome this limitation, we propose a novel methodology to estimate demand elasticity and infer unobserved quality using only data on prices and market shares. Our approach builds on the standard framework that models product quality as residual demand. This framework requires estimating price elasticities and the standard approach assumes CES demand and imposes uncorrelated supply and demand shocks. However, the latter assumption is untenable if we associate demand shocks with quality and generates an upward bias in the estimates of price elasticities. Our strategy circumvents this problem by restricting the dynamics of product quality to a Markov process. We apply our new methodology to the US customs data (1989-2006), and find that quality improvements contribute the most to the gains from trade in the US. Quality improvements have lowered the price of US imports relative to the CPI by 17%, with Chinese products contributing the most. In comparison, import prices have fallen by around 11% relative to the CPI and increasing variety has contributed an additional 4%. These findings demonstrate that accounting for quality is essential to better understand and measure the effects of international trade.

Contents

List of Tables

List of Figures

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisors, Jaromir, James, Danial and Rosen, for their unwavering support and guidance throughout my academic journey. Their invaluable insights, encouragement, and constructive criticism have helped me to develop the necessary skills and knowledge to complete this dissertation. I am truly grateful for their patience, firm guidance, and mentorship, which have been essential to my success. Their passion for research and commitment to excellence have been a constant source of inspiration, and I feel incredibly fortunate to have had the opportunity to work with them.

During my time in Boston, I have been incredibly lucky to make meaningful and lasting friendships with many people within and outside the Ph.D. program. A special thank you goes to Luigi, my roommate, classmate and coauthor, for the companionship during this long journey, for all the exquisite meals we made and shared together, for all the brainstorming, discussions, insight, and creativity in our joint projects. Thank you to Marius and Vera, for their genuine friendship and inspiring life choices, and Pietro for the disinterested support. Un ringraziamento a tutti i miei amici italiani de *50 Sfumature di Sambo* e non solo, per essere sempre presenti nonostante la distanza.

Grazie alla mia famiglia – i miei genitori, Maria Irene e Stefano, e mio fratello, Tommaso – per il supporto e l'affetto incondizionato. Il vostro incoraggiamento, sacrifici e presenza costante sono stati fonte di motivazione e forza; questa tesi è stata scritta soprattutto grazie a voi. I would also like to thank my new family in the U.S.—Julie, Arthur, Liz, Robyn, Ella and Ty for supporting me during my Ph.D. and making me feel at home. And finally, there are not enough words to express my gratitude to my partner, Tess, for your constant support, happiness and love.

Per Aspera Ad Astra

Chapter 1.

Decomposing the (In)Sensitivity of CPI to Exchange Rates

I study the relative importance of domestic frictions and border price insensitivity for the response of domestic consumer prices (CPI) to exchange rate fluctuations. Using firm and transaction-level data from Chile, I estimate that the presence of domestic frictions — distribution costs, variable markups and nominal rigidities — reduce the responsiveness of domestic CPI to exchange rate fluctuations by 60% relatively to an economy that abstract away from it. These frictions are quantitatively more important than the insensitivity of border prices. The presence of domestic frictions also matters for the channels of CPI sensitivity: contrary to prior work, most of the sensitivity arises from the change in the price of imported consumption goods. This channel is more important than the costs arising from imported inputs in the production of domestic goods. The reason is that domestic frictions dampen the price sensitivity of domestically produced goods relatively more. Furthermore, the sensitivity varies across products because of the heterogeneity in domestic frictions, import exposure, and consumption shares. The heterogeneity matters for the overall (in)sensitivity as domestic products with higher import exposure face larger frictions and have lower consumption shares. Ignoring the heterogeneity identifies the wrong products from which most of the sensitivity arises, with implications for monetary policy targeting in open economy and redistribution dynamics.

1.1 Introduction

The relationship between domestic prices (Consumer Price Index, CPI) and exchange rates is a central question in international economics, with implications from optimal monetary policy in open economy to domestic redistribution dynamics.[1] Figure 1.1 documents that, on average, CPI changes by 0.76% after a 10% exchange rate change in Chile between 2009 and 2019. Thus, exchange rate changes are only partially transmitted to domestic prices, in line with the extensive evidence documenting that CPI responds weakly to exchange rate fluctuations (Goldberg and Campa, 2010, Gopinath, 2015). In order to rationalize the weak response of CPI, the literature has focused on the low sensitivity of the border price of imported goods with respect to exchange rate fluctuations.[2] In other words, the common assumption is that domestic prices do not change because the price of imported goods is not influenced by exchange rate fluctuations. However, back-of-the-envelope calculations show that the low sensitivity of border prices imply a sensitivity of domestic prices much higher than the estimated

[1]One fundamental aspect for monetary policy trade-offs in open economy is which inflation rate is relevant to policymakers, which, in turn, depends on the exchange rate pass-through into domestic prices (Mishkin, 2008, Benigno and Benigno, 2003, Corsetti et al., 2010). Similarly, exchange rate fluctuations influence domestic redistribution dynamics as firms and consumers use different mixes of domestic and imported products (Cravino and Levchenko, 2017*a*, Jaravel, 2021). Moreover, understanding relationship between CPI and exchange rates, and the factors influencing it has broad implications for the transmission of international shocks, international business cycle comovements and external imbalances (Corsetti et al., 2008, Backus and Smith, 1993).

[2]See Burstein and Gopinath (2014) for a survey.

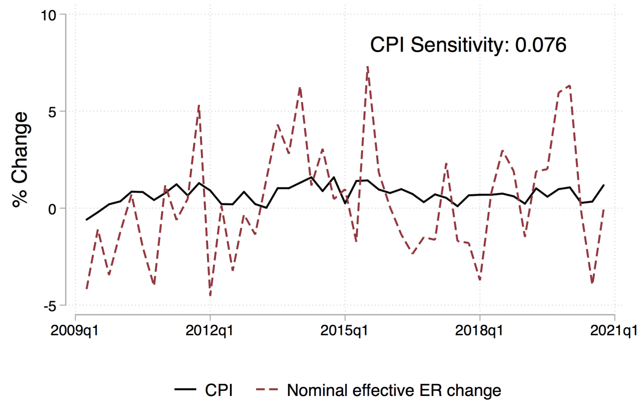
one.[3]

In this paper, I provide extensive empirical results to document that the insensitivity of domestic prices emerges mainly because of the existence of several domestic frictions, instead of border price insensitivity. I start by developing a framework to quantify what the sensitivity of CPI to exchange rates is expected to be, given the existence of insensitivity in border prices and domestic frictions (Goldberg and Campa, 2010). CPI is sensitive to exchange rate fluctuations because of the consumption of imported goods (direct exposure), the use of imported intermediate inputs in the production of domestic goods and the presence of domestic input-output linkages (indirect exposure). The model aims at capturing the role that domestic frictions — distribution costs, variable markups and nominal rigidities — have in the domestic transmission of exchange rate fluctuations to CPI. I compare the importance of domestic frictions to the effect of border price insensitivity, which is taken as given.

The presence of domestic frictions introduces a wedge between the border price of imports and producers' costs, on one side, and the domestic retail price, on the other, dampening the response of the latter to exchange rate changes and making CPI less sensitive. Distribution costs, i.e. transportation, insurance and wholesaling costs represent a substantial component of

[3]For the case of Chile, the estimated incomplete exchange rate pass-through into border prices is about 0.75. Knowing that the share of imported final consumption is 15% and the share of imported intermediate inputs in total production costs is 25%, the sensitivity of domestic prices should be around 0.27, much higher than the 0.076 reported in Figure 1.1.

FIGURE 1.1 – Estimated CPI Sensitivity to Exchange Rates



The figure plots the relationship between the change in domestic CPI (black, solid line) and the trade-weighted measure of nominal exchange rate (red, dashed line). Inflation and exchange rate data are sourced from IMF and Datastream, respectively. Trade shares are computed from the universe of import transactions from 2009 to 2020. The coefficient reported is the contemporaneous CPI sensitivity estimated from Equation (1) in Appendix C.

retail prices (Goldberg and Campa, 2010, Burstein et al., 2003). This reduces the exposure of CPI to exchange rates by reducing the weight of import border prices and domestic producers' cost in CPI. Similarly, the presence of variable markups and nominal rigidities in the domestic economy creates additional wedges between the change in domestic producers' costs following an exchange rate change and the retail price of domestic goods (Klenow and Willis, 2016, Nakamura and Steinsson, 2008). The pass-through rate of marginal cost changes is incomplete because of variable markups. In other words, domestic firms do not fully adjust their price to changes in their own cost because they absorb part of the cost change in their own margins by modifying the markup they charge. Moreover, the price of domestic goods is sticky because domestic firms face nominal rigidities in the spirit of Calvo.

I leverage several, highly disaggregated data sources from Chile to discipline the rich structure of the model and gauge the role of each domestic frictions relative to border price insensitivity. I construct a granular, product-level (180$\times$180) input-output table for the Chilean economy to measure the channels through which exchange rate fluctuations are transmitted to CPI. The input-output table allows me to account for direct and indirect exposure to imports and to capture the transmission of exchange rate changes through the domestic network (Basu, 1994, Rubbo, 2020). I calibrate each domestic friction using micro-level data, allowing me to account for their heterogeneity at the product level. Specifically, I compute distribution costs for each product from the input-output table, differentiating according to their origin (domestic vs imported) and use (intermediate vs final consumption). I estimate markups using state-of-the-art production function estimation methods and firm-level data from Chile to calibrate variable markups and markup elasticities at the sectoral level. Similarly, I calibrate nominal rigidities using micro-level estimates of price adjustment frequencies from Chile. Lastly, I use the universe of import transaction data to calibrate, in reduced form, the exchange rate pass-through into border prices and its heterogeneity across products due to importers' heterogeneity.

The calibrated model including both border price insensitivity and domestic frictions matches the untargeted estimated sensitivity of domestic prices to exchange rate fluctuations (Figure 1.1). Combining domestic frictions and border price insensitivity allows to explain the insensitivity of CPI with respect to exchange rates documented in Figure 1.1 in its entirety. This supports the importance of accounting for domestic frictions, the relevance of the modelling choices and the validity of the calibration strategy, providing a benchmark for future empirical studies on CPI sensitivity to exchange rates.

I find that domestic frictions are more important than the insensitivity of border prices in explaining the insensitivity of domestic prices to exchange rates, Figure 1.1. Relative to an economy where exchange rate changes are passed entirely into import and domestic prices, the presence of domestic frictions reduces the sensitivity of CPI with respect to exchange rates by 60%. On the contrary, accounting for border price insensitivity reduces CPI sensitivity by 40%. Thus, by dampening the domestic transmission of exchange rate fluctuations, the insensitivity of domestic prices emerges mainly because of the existence of several domestic frictions. Moreover, each individual friction substantially contributes to the overall insensitivity of domestic prices. Distribution costs, variable markups and nominal rigidities reduce the sensitivity of CPI by approximately 35%, 20% and 15%, respectively, suggesting the importance of jointly modelling these frictions.

I gauge the implications for domestic prices quantifying the relative importance of domestic frictions and insensitivity of border prices during the depreciation of the Chilean peso triggered by the *"Estallido Social"* in 2019.[4] Following the 10% depreciation of the Chilean peso between 2019Q3 and 2020Q1, the price of imported goods rose, fueling higher domestic inflation. Through the lens of the calibrated model, the presence of domestic friction insulated domestic inflation, reducing the domestic inflation rate by 50% (0.6 p.p. lower at the quarterly level), twice as much as the contribution of border price insensitivity (0.3 p.p. lower).

Accounting for domestic frictions provides novel insights also on the dominant channel for the sensitivity of CPI to exchange rate fluctuations. In contrast to previous literature, I find that the presence of domestic friction implies that the dominant channel for the sensitivity of CPI is through the presence of imported goods in the final consumption basket, also known as direct exposure. This is in contrast to previous quantification exercises showing that direct exposure is as relevant as indirect exposure, where the latter instead arises from the use of imported intermediate inputs in the production of domestic goods (Goldberg and Campa, 2010).[5] The conflicting evidence can be rationalized by the presence of domestic frictions. Domestic frictions not only reduce the sensitivity of all prices, but make the price of domestically produced goods relatively more insensitive than the price

[4]The "Estallido Social " (social outburst) refers to a series of massive and severe riots in Chile between October 2019 and March 2020. The riots triggered a major devaluation of the Chilean peso against all major currencies until the Central Bank of Chile intervention in late November.

[5]Goldberg and Campa (2010) focuses on a group of OECD economies. Chile's exposure to imports is quantitatively similar to the average exposure of OECD countries.

of imported goods. One of the reasons is that domestic frictions dampen the spillover effects of the domestic input-output network, reducing the role of indirect import exposure.

Calibrating the model at product-level unveils a rich heterogeneity in the sensitivity to exchange rates across products, with implications for inflation targeting and redistribution. The sensitivity varies across products because of the heterogeneity in domestic frictions, import exposure, consumption shares, and border price sensitivity. These different sources of heterogeneity matter for the overall (in)sensitivity as domestic products with higher import exposure in production face larger distribution costs, larger real rigidities and have lower consumption shares. Moreover, the identity of the products transmitting the exchange rate fluctuations the most varies when I take into account different subsets of frictions. Ignoring any friction or their heterogeneity has implications for inflation targeting and redistribution: optimal policy requires knowing what products are contributing the most and therefore what prices to target (Pasten et al., 2020, Rubbo, 2020). Similarly, consumers and firms are differentially exposed to exchange rate fluctuations since they use different mixes of imported and domestic goods (Jaravel, 2021).

Incomplete pass-through into border price explains part of the low sensitivity of CPI to exchange rate fluctuations and part of its quantitative role arises because of importers' heterogeneity. I show that importers' heterogeneity in terms of age, size and market power, and presence of trade relationships matters for the sensitivity of border and domestic prices. Specifically, I measure these dimensions with a measure of importers' experience and find that importers with longer experience have larger market shares and face a lower pass-through rate of exchange rate fluctuations into border prices. Importers' heterogeneity reduces CPI sensitivity by 20%. Moreover, the rise in importers' experience accounts for 40% of the decline in CPI sensitivity to exchange rates over the period 2009-2019 (Campa and Goldberg, 2005, Camatte et al., 2021, Georgiadis et al., 2020).

Prior Work: This paper is related to several strands of literature. First, it contributes to the literature studying the low sensitivity of domestic inflation to exchange rate fluctuations. Goldberg and Campa (2010) quantify CPI sensitivity accounting for the effects of import exposure and distribution costs for a set of OECD economies, and document that the main channel for CPI sensitivity is through the costs arising from imported input used in goods production (indirect exposure), as opposed to imported final consumption (direct exposure). In my analysis, I extend their framework to include a more accurate and comprehensive characterization of the domestic economy and its (heterogeneous) frictions. By accounting for domestic frictions, the main channel for CPI sensitivity changes as imported goods directly consumed are more important than imported input use in goods production. Burstein et al. (2003) and Corsetti and Dedola (2005) also show that distribution costs dampen the response of import and consumer prices to exchange rate changes, but fall short in combining them with other leading frictions or accounting for their heterogeneity and interactions.

My work is connected to the vast literature studying the incomplete pass-through rate into border prices and its determinants.[6] Gopinath and Itskhoki (2011) show that both nominal and real rigidities are necessary to quantitatively account for the response of border prices to exchange rates. I complement their work by showing that the effects of these frictions are not limited to border prices but are relevant also for the response of domestic price to exchange rates. In addition, I document that incomplete pass-through into border prices is not the main driver of the low sensitivity of domestic prices, as domestic frictions account for 60% of the insensitivity of CPI.

Prior work focuses on the firm-level determinants of incomplete pass-through into border prices, such as firm size and market share (Berman et al., 2012, Atkeson and Burstein, 2008), imported inputs (Amiti et al., 2014), strategic complementarities (Amiti et al., 2019), product quality (Chen and Juvenal, 2016) and bargaining and buyer market power (Drozd and Nosal, 2012, Heise, 2019, Alviarez et al., 2021, Errico, 2022).[7] I contribute to this literature by quantifying the aggregate relevance of micro-level determinants

[6]See Burstein and Gopinath (2014) and Goldberg and Hellerstein (2008) for recent surveys.

[7]Other related papers are Neiman (2010), which focuses on the effect of intra-firm and arm-length relationships, and Gopinath et al. (2010) and Chen, Chung and Novy (2022), that study the effect of invoicing choices on pass-through.

of heterogeneous pass-through rates, as I account for the heterogeneity in border price pass-through due to importers' experience.

My work is related to the literature that focuses on production networks, heterogeneity in frictions and propagation of shocks.[8] Rubbo (2020) and Pasten et al. (2020) show, in closed economy, that heterogeneity in price rigidity is key for the transmission of monetary shocks, whereas I focus on different heterogeneous domestic frictions, their interactions and their role for the transmission of exchange rate changes. Dhyne et al. (2021) quantify the propagation of foreign demand shocks using domestic firm-to-firm transactions. Using Chilean data, Huneeus (2018) focuses on the effects of foreign demand shocks in a model with endogenous network. Relative to these papers, I combine input-output tables and product-level frictions to describe the domestic economy and study the transmission of exchange rate changes into domestic prices. Di Giovanni et al. (2017), Cravino and Levchenko (2017*b*) and Di Giovanni and Levchenko (2010) study the role of multinational firms and international input-output linkages for the transmission of productivity and inflation shocks across borders. My analysis complements theirs in focusing on the domestic transmission of exchange rate changes.

Finally, my paper is related to the literature documenting a long-run decline in domestic price sensitivity to exchange rate fluctuations. Auer

[8]See Carvalho and Tahbaz-Salehi (2019) for a recent survey.

et al. (2019), Camatte et al. (2021) and Georgiadis et al. (2020) use aggregate global input-output table to show that CPI sensitivity to exchange rates decreases as global value chain (GVC) participation and trade openness rise. My work is complementary to theirs as I use micro-level data to quantify the aggregate effects of importers' experience, which relate to prolonged participation in international markets and GVC. Consistent with the literature, I find that a substantial part of the long-run decline can be explained by rising importers' experience.[9]

The rest of the paper is structured as follows. In Section 1.2, I present my modelling approach, beginning with a price aggregator and then presenting the model of pass-through, with particular attention to the role of leading domestic frictions. Section 1.3 discusses the calibration strategy of the model in detail and Section 1.4 presents the main results on the decomposition of the (in)sensitivity of domestic prices to exchange rates. Section 1.5 concludes.

1.2 A Model of Exchange Rate Pass-Through into CPI

In this section, I derive a set of measurement equations for the pass-through of exchange rate fluctuations into domestic prices (CPI) to decom-

[9]The quantitative importance of importers' experience and GVC participation is of relevance also for the missing inflation puzzle: Heise et al. (2022) show that global factors, like imported products and import competition, account for part of the growing disconnect between domestic inflation and unemployment.

pose the role that domestic forces and border price response play for the sensitivity of CPI.

The focus of the modelling approach is characterizing the domestic transmission of exchange rate fluctuations. I describe a theoretical framework that formalizes the domestic channels and frictions influencing the domestic transmission of exchange rate fluctuations into the CPI. I account for incomplete and heterogeneous pass-through into border prices, but I abstracts away from any micro-foundation and directly disciplined it using import transaction data.

I propose a parsimonious, one-period, partial-equilibrium, multi-product framework in the spirit of Goldberg and Campa (2010). I combine and extend several elements that affect the domestic transmission of (exchange rate) shocks previously studied in the literature, such as distribution costs (Burstein et al., 2003, Corsetti and Dedola, 2005), variable markups (Goldberg and Verboven, 2001), imported inputs in the production of domestic products (Goldberg and Campa, 2010) and roundabout production (Basu, 1994), and nominal rigidities (Gopinath and Itskhoki, 2011). The model allows to outline the key components influencing the sensitivity of domestic prices to exchange rate fluctuations, linking the behavior of border prices to the dynamics of domestic CPI, and perform an accurate calibration exercise to quantitatively assess their individual role.

1.2.1 *Set up*

The section introduces the assumptions about preferences, production, and frictions. I then derive a measurement equation for the pass-through rate of exchange rate fluctuations into domestic inflation.

Price Aggregator. The preferences of the domestic representative household are given by

$$W(C, L) = U(C) - V(L), \tag{1}$$

where C and L represent the household's final consumption and total labor supply, respectively.[10] I assume domestic households consume N sectoral goods $i \in \{1, \cdot, N\}$.[11] Specifically, the final consumption basket of the household, C, is given by a homogeneous of degree one consumption aggregator $\mathcal{C}$ of the individual sectoral goods, $C = \mathcal{C}(c_1, \cdot, c_N)$. The household's utility maximization problem is subject to a standard budget constraint given by:

$$PC \equiv \sum_{i=1}^{N} p_i c_i \leq wL + \sum_{i=1}^{n_D} \int_0^1 \pi_{ik} dk, \tag{2}$$

where P is the nominal price index of the final consumption bundle; wL is the labor income; and the last term captures the dividends from owning the domestic firms.

[10] Typical regularity conditions are imposed on U and V: strictly increasing, twice differentiable, and $U'' < 0$, $V'' > 0$ and the Inada conditions are satisfied.

[11] I use i to indicate both the good and the industry that produces the good.

I assume that C takes the form of a Cobb-Douglas aggregator as follows:

$$C(c_1, \cdot, c_N) = \prod_{i=1}^{N} \left(\frac{c_i}{\beta_i} \right)^{\beta_i}, \qquad \text{with} \sum_{i=1}^{N} \beta_i = 1 \tag{3}$$

where c_i is the amount of good i consumed and the constants $\beta_i \geq 0$ capture the share of each good in the household's final consumption.

The utility-based final consumption price index, which is the model-implied measure of CPI, is then given by:

$$P(p_1, \cdot, p_n) = \prod_{i=1}^{N} p_i^{\beta_i}, \tag{4}$$

where p_i is the retail price of the good of industry i.

Therefore, the pass-through of exchange rates into CPI (the elasticity of CPI to changes in nominal exchange rates, e), $\eta^{P,e}$, is given by:

$$\eta^{P,e} \equiv \frac{d \log P}{d \log e} = \boldsymbol{\beta} \times \boldsymbol{\eta}^{\mathbf{p},e}, \tag{5}$$

where $\boldsymbol{\beta}$ refers to the $N \times 1$ vector of expenditure shares, $(\beta_1, \cdot, \beta_N)$, and $\boldsymbol{\eta}^{\mathbf{p},e}$ to the $N \times 1$ vector of price elasticities, $(\eta^{p_1,e}, \cdot, \eta^{p_N,e})^T$.

The pass-through of exchange rate movements into CPI is a weighted average of the pass-through rates into the prices all goods consumed in the final consumption basket. Given the Cobb-Douglas specification in Equation (3), the relative weights correspond to the expenditure shares in total consumption, $\beta_i = \frac{p_i c_i}{PC}$.

I assume that a subset n_F ($n_D = N - n_F$) of sectoral goods are imported (produced domestically).[12] In this way, I can disentangle the effects of direct and indirect import exposure. The former refers to the presence of imported final consumption goods, while the latter accounts for the use of imported intermediate inputs in the production of domestic goods. Highlighting this decomposition, Equation (5) can be rewritten as:

$$\eta^{P,e} = \boldsymbol{\beta} \times \boldsymbol{\eta}^{\mathbf{p},e} = \underbrace{\boldsymbol{\beta}^{D} \times \boldsymbol{\eta}^{\mathbf{p}^{D},e}}_{\substack{\text{Indirect exposure} \\ \text{(Imported Intermediate Inputs)}}} + \underbrace{\boldsymbol{\beta}^{F} \times \boldsymbol{\eta}^{\mathbf{p}^{F},e}}_{\substack{\text{Direct exposure} \\ \text{(Imported Final Consumption)}}} , \tag{6}$$

where $\eta^{\mathbf{p}^{D},e}$ ($\eta^{\mathbf{p}^{D},e}$) is the vector of pass-through rates into the retail price of a domestically (imported) sectoral goods.

In the following paragraphs, I first characterize the sensitivity of domestically produced goods, $\eta^{\mathbf{p}^{D},e}$ in Equation (6), by introducing several elements that influence the transmission of exchange rate fluctuations. I then elaborate further on the sensitivity of imported goods, $\eta^{\mathbf{p}^{D},e}$.

Production and Price Elasticity of Domestic Goods, $\eta^{\mathbf{p}^{D},e}$. I assume that each domestic sectoral good, $i \in n_D$, is produced by a local competitive distributor by aggregating a mass of sectoral varieties, La'O and Tahbaz-Salehi (2022). In turn, sectoral varieties are produced by a continuum of domestic monopolistically competitive firms, indexed by $k \in [0, 1]$.

The competitive distributor of industry $i \in n_D$ aggregates the mass of differentiated varieties into an homogeneous sectoral good, y_i, using an ho-

[12] I label a sectoral good $i \in n_F$ ($i \in n_D$) as "imported" ("domestic").

mothetic Kimball aggregator, Kimball (1995):

$$\sum_k A_i \mathcal{K}_i \left(\frac{y_{ik}}{y_i} \right) = 1, \tag{7}$$

where y_{ik} is the consumption of variety k in industry i, and A_i is a demand shifter; $\mathcal{K}(\cdot)$ is such that $\mathcal{K}(\cdot) > 0$, $\mathcal{K}'(\cdot) > 0$, $\mathcal{K}''(\cdot) < 0$ and $\mathcal{K}(1) = 1$. The distributor's VES technology represents the demand schedule that monopolistically competitive firms face. In the quantitative analysis in Section 1.3, I adopt the common Klenow and Willis (2016) formulation for the Kimball aggregator. In this case, Marshall's weak second law is satisfied and implies that, as firms lower their prices, their demand becomes more inelastic and their markup increases. Thus, larger monopolistically competitive firms will have higher markups, higher markup elasticity and lower pass through rate of cost shocks (Burstein et al., 2003, Kimball, 1995).

The distributor sells the homogeneous sectoral good y_i incurring in distribution costs. Distribution costs represent the per-unit service inputs required to bring the homogeneous industry goods to consumers and firms, e.g. transportation, wholesales and retail services, marketing, etc (Burstein et al., 2003, Corsetti and Dedola, 2005). I follow Burstein et al. (2003) and assume that distribution services are combined with one unit of sectoral homogeneous good using a Cobb-Douglas technology and that distribution services are produced using only labor. Thus, the retail price of good i, p_i, is:

$$p_i = \widetilde{p}_i^{\,1-\phi_i} w^{\phi_i} \qquad \text{with } \phi \leq 1, \tag{8}$$

where $\widetilde{p}_i$ is the price of the aggregate homogeneous good i and ϕ_i the cost share of distribution services in the retail price of good i. I assume that distribution costs are heterogeneous across industries, as denoted by the i-specific weights in the production technology.

The monopolistically-competitive firms within each domestic industry $i \in n_D$ are symmetric and use a common constant return to scale production function. Domestic and imported sectoral goods can be used as inputs in the production of domestic varieties, together with labor. Indirect exposure arises from both the direct use of imported inputs and the presence of domestic input-output linkages.[13] The production function of firm k is given by:

$$y_{i,k} = F_i(l_{i,k}, x_{i1,k}, \cdot, x_{iN,k}), \tag{9}$$

where $y_{i,k}$ is firm k's output, $l_{i,k}$ is the labor input and $x_{ij,k}$ is the amount of good j used as input by firm k in sector i. I assume that firms employ the same Cobb-Douglas technology:

$$y_{i,k} = F_i(l_{i,k}, x_{i1,k}, \cdot, x_{iN,k}) = \zeta_i l_{i,k}^{\alpha_{i,l}} \prod_{j=1}^{N} x_{ij,k}^{\alpha_{i,j}} \qquad \text{with } \alpha_{i,l} + \sum_{j=1}^{N} \alpha_{i,j} = 1. \tag{10}$$

I assume that $\alpha_{i,l} > 0$, i.e. that labor is an essential input for the produc-

[13]In other words, a firm's production cost is directly exposed to imported intermediate inputs when the firm is directly using imported inputs in production. However, the firm is potentially exposed even when it does not use any imported input. This happens through the links to other domestic firms that make use of imported inputs. The latter is captured by domestic input-output linkages.

tion of all varieties, in the sense that $F_i(0, x_{i1,k}, \cdot, x_{iN,k}) = 0$. $\alpha_{i,j}$ denotes the share of good j in industry i's production technology.[14] ζ_i is a sector-specific normalization constant.

Given the assumption on the distributor's aggregating technology, monopolistically competitive producers charge a variable markup over the marginal cost:

$$\widetilde{p_{ik}} = \mu_i mc_i \qquad \text{with } mc_i = w^{\alpha_{i,l}} \prod_{j=1}^{N} p_j^{\alpha_{i,j}}, \tag{11}$$

where $\widetilde{p_{ik}}$ is the price paid by the distributor for variety k, μ_i is the markup charged and the expression for the marginal cost, mc, comes from the specific production function assumed in Equation (10). The markup charged by monopolistically competitive firms increases in firm sales and becomes more sensitive to cost shocks, which implies a lower pass through rate.

I assume that monopolistically competitive producers are subject to Calvo-style nominal rigidities: a fraction δ_i of firms in each sector i can adjust prices to changes in sectoral marginal costs $d \log \mathrm{mc}_i$. I consider a one-period framework, Rubbo (2020). The timing is as follow: before the world begins, firms set prices based on their marginal cost, Equation (11); then the exchange rate change is realized; because of price rigidities, firms are allowed to adjust their price after observing the realized change in their marginal cost with probability δ_i; the world ends after production and

[14] I assume that $\alpha_{i,j} \geq 0$ or, in other words, that industry i may rely on the goods produced by other (domestic or imported) industries as intermediate inputs.

consumption take place.

I now derive an expression for a change in the retail price of a domestic sectoral good following a change in exchange rate, which feeds into domestic prices through imported intermediate inputs and input-output linkages. I focus on the direct effect of exchange rate, Burstein and Gopinath (2014): I consider a partial-equilibrium response of domestic prices, not accounting for changes in the wage rate or the response of firms to changes in sectoral price indices.

A change in the price of domestic goods $i \in n_D$, π_i^D, is:

$$\pi_i^D \equiv d\log p_i^D = \underbrace{(1-\phi_i)}_{\text{Distribution Costs}} \quad \underbrace{\delta_i}_{\text{Nominal rigidities}} \quad \underbrace{\frac{1}{1+\Gamma_i}}_{\text{Real rigidities}} \quad d\log mc_i \tag{12}$$

$$\underbrace{d\log mc_i}_{\text{Change in mc}} = \underbrace{\sum_{j=1}^{n_D} \alpha_{i,j}\pi_j^D}_{\substack{\text{Exposure } via \\ \text{IO linkages}}} \quad + \quad \underbrace{\sum_{j'=1}^{n_F} \alpha_{i,j'}\pi_{j'}^F (d\log e)}_{\text{Import Exposure}}. \tag{13}$$

A change in the retail price of a domestic good, π^D, follows a change in the marginal cost - last term in Equation (12). The latter, in turn, originates from a change in input prices, Equation (13). The second summation captures the change in the price of imported inputs (π^F) while the first summation represents the change in the price of domestically sourced inputs. Crucially, the former depends directly on the (log) exchange rate change, $d\log\varepsilon$. The latter instead captures the indirect effects that exchange rate changes have through the domestic production network and the indirect exposure to imported inputs. Notice that the relevant input prices are the retail prices set by the distributors, which include distribution services.

A change in marginal cost is not passed completely into the retail price of domestic goods because of the presence of several frictions in the economy. Equation (12) shows that the change in marginal cost is attenuated by the presence of distribution costs, variable markups and nominal rigidities. The presence of nominal rigidities allows only a fraction δ_i of firms to change prices, i.e. those firms touched by the Calvo fairy.

Even if the firm is able to adjust its price, real rigidities due to variable markups make firms reluctant to change their price relative to other firms' prices. The presence of variable markups allows firms to incompletely pass the change in marginal cost into prices by adjusting its markups and partially absorbing the change in costs. The pass-through rate inversely depends on how much the markup is sensitive, i.e. on the markup elasticity $\Gamma_i = \frac{\partial \mu_i}{\partial p_i} > 0$: the more the markup is sensitive, the lower the pass-through of cost shocks to prices. The ratio $\frac{1}{1+\Gamma_i} < 1$ in Equation (12) formally captures the incomplete pass-through due to variable markups.

Lastly, the presence of distribution costs in Equation (8) reduces the sensitivity of retail prices to changes in the production cost as the latter accounts only for a share $1 - \phi_i$ of the retail price. By reducing the sensitivity of prices to changes in marginal costs, these frictions ultimately dampen the transmission of exchange rate fluctuations.

Because of round-about production and input-output linkage, domestic prices can change because of indirect exposure. Let $\pi^D = (\pi_1, \cdot, \pi_{n_D})^T$ be the $n_D \times 1$ vector of domestic price changes. Combining Equations (13) and (12) and rearranging, the vector of changes in domestic prices becomes:

$$\pi^D = \underbrace{(I - \Phi\Delta\Gamma S_d)^{-1}}_{\substack{\text{Adjusted} \\ \text{Leontief Inverse}}} \underbrace{\Phi}_{\substack{\text{Matrix of} \\ (1-\phi_i)}} \underbrace{\Delta}_{\substack{\text{Matrix of} \\ \delta_i}} \underbrace{\Gamma}_{\substack{\text{Matrix of} \\ \frac{1}{1+\Gamma_i}}} \underbrace{S_m}_{\substack{\text{Imported intermediate} \\ \text{input shares}}} \pi^F(d\log e). \tag{14}$$

A change in the price of foreign inputs, $\pi^F = (\pi_1, \cdot, \pi_{n_F})^T$, is transmitted to domestic prices through the shares of imported intermediate inputs, captured by the matrix S_m.[15] However, the resulting change in marginal costs is attenuated by the presence of distribution costs, variable markups and nominal rigidities, captured respectively by the diagonal matrices Φ, Δ and Γ. Lastly, the first term represents the Adjusted Leontief Inverse matrix, that captures the effects of domestic round-about production. Namely, the matrix quantifies the amplifying effect of domestic input-output linkages on the transmission of cost changes. The Leontief matrix $(I - S_d)^{-1}$, with S_d being the input-output matrix of domestic input shares, captures the total expenditure of sector i on good j.[16] The adjusted matrix accounts for the fact that marginal cost changes are not fully passed into prices because of

[15] In other words, the matrix S_m collects all the input shares $\alpha_{i,j}$ where $j \in n_F$.

[16] Similarly to S_m, S_d captures all the input shares $\alpha_{i,j}$ where i, j are both domestic products.

the presence of domestic frictions, ultimately capturing the effective total elasticity.

It follows immediately that the price elasticity of domestic goods in Equation (6), $\boldsymbol{\eta}^{\mathbf{p}^D,e}$, is:

$$\boldsymbol{\eta}^{\mathbf{p}^D,e} = \underbrace{\left(I - \Phi\Delta\Gamma S_d\right)^{-1} \Phi\Delta\Gamma}_{\substack{\text{Domestic network} \\ \text{\& frictions}}} \times \underbrace{S_m}_{\text{Import Exposure}} \times \underbrace{\boldsymbol{\eta}^{\mathbf{p}^F,e}}_{\substack{\text{Elasticity of} \\ \text{imported inputs}}}, \tag{15}$$

where $\boldsymbol{\eta}^{\mathbf{p}^F,e}$ is the vector of price elasticities of imported goods. Equation (15) shows that the sensitivity of domestic goods to exchange rate fluctuations depends not only on how the retail price of imported goods reacts to exchange rate fluctuations ($\boldsymbol{\eta}^{\mathbf{p}^F,e}$) and how much domestic production is directly exposed to imported inputs (S_m), but also on the features (frictions and network) of the domestic economy.

Price Elasticity of Imported Goods, $\eta^{\mathbf{p}^F,e}$. The sensitivity of CPI to exchange rates depends directly on how the price of imported good changes after an exchange rate shock, $\eta^{\mathbf{p}^F,e}$, as part of the final consumption bundle is imported from abroad (direct exposure). Similarly, CPI indirect exposure also depends on $\eta^{\mathbf{p}^F,e}$ as imported inputs are used in the production of domestic goods.

I specify the sensitivity of the retail price of imported goods, $\eta^{\mathbf{p}^F,e}$, assuming that imported goods are produced abroad and purchased by a local distributor, which combines imported goods with local distribution services and determines the retail price of imported goods, p_i. I also assume that the domestic economy is small (small open economy assumption) and rule out international input-output linkages. In this case, changes in domestic prices do not affect the foreign production costs of imported goods.

As in Equation (8), the retail price of imported goods is given by:

$$p_i = \left(\widetilde{p}_i(e)\right)^{1-\phi_i} w^{\phi_i} \qquad \text{with } \phi_i \leq 1, \tag{16}$$

where $i \in n_F$ and $\widetilde{p}_i$ is the border price of the imported good, which is determined by the foreign producer and depends on the exchange rate.

Given the specific focus on the role of domestic frictions and domestic transmission, I abstract away from any micro-foundation of the production process of imported goods and discipline directly how border prices react to exchange rate fluctuations. I assume that the pass-through rate of exchange rate fluctuations into border prices is incomplete, i.e. $\Psi_i = \frac{\partial \log \widetilde{p}_i}{\partial \log e} < 1$, consistently with extensive evidence (Burstein and Gopinath, 2014, Gopinath, 2015). In the quantitative analysis, I use the universe of import transactions to discipline the behavior of border prices at the product level in a reduced form.

Following the same reasoning for domestic prices, the change in the retail price of imported goods following an exchange rate shock, $d\log e$, is:

$$\pi_i^F \equiv d\log p_i^F = \underbrace{(1-\phi_i)}_{\text{Distribution Costs}} \quad \underbrace{\Psi_i}_{\substack{\text{Heterogeneous} \\ \text{Border Pass-Through}}} \quad d\log e,$$

where Ψ_i captures the incomplete pass-through rate into border prices. It follows that the price elasticity of imported goods appearing in Equations (6) and (15) is:

$$\eta^{p^F,e} = \underbrace{\Phi}_{\substack{\text{Matrix of} \\ (1-\phi_i)}} \underbrace{\Psi}_{\substack{\text{Matrix of Heterogeneous} \\ \text{Border Pass-Through}}} . \tag{17}$$

The sensitivity of imported good retail prices decreases the larger is the share of distribution services included (ϕ_i), and the lower is the sensitivity of border prices (Ψ_i). In Section 1.3, I calibrate the sensitivity of border prices at the product level, Ψ_i, using import transaction data and leveraging heterogeneity across importers. In this regard, I assume that Ψ_i depends on importers' characteristics such as importers' size and experience since a large literature points to the role of customer accumulation, buyer market power and firm-to-firm trade relationships on pricing and pass-through dynamics (Atkeson and Burstein, 2008, Berman et al., 2012, Drozd and Nosal, 2012, Alviarez et al., 2021, Heise, 2019, Errico, 2022).

The sensitivity of the retail price of imported goods, $\eta^{p^F,e}$ in Equation (17), together with (6) and (15), fully characterizes all elements determining the transmission of exchange rate fluctuations into CPI. The three measurement equations jointly provide a decomposition of the major forces impacting the sensitivity of domestic prices. Section 1.3 shows how to calibrate in detail each component.

Discussion of Model Assumptions. I close this section with a discussion on the assumptions and caveats made in the description of the domestic economy and the sensitivity of CPI to exchange rate changes.

I derive the pass-through of exchange rate into CPI, Equation (5), focusing on the direct effect of exchange rates into prices (Burstein and Gopinath, 2014). In other words, I abstract away from the effect of exchange rate changes on domestic wages, sectoral prices and quantities. While such partial-equilibrium assumption is a simplification, most of the exchange rate fluctuations at quarterly level are relatively small and changes in aggregate variables like wages are likely to occur in response to larger devaluations or over long horizons. Thus, the quantitative analysis in Section 1.4 can be interpreted as a short-run quantification. Moreover, general equilibrium dynamics require additional structure in terms of wage determination and taking a stance on the dynamics of exchange rates and sectoral prices for a careful characterization of the dynamics of domestic prices in the presence of Calvo rigidities.

The second key assumption is that the production and consumption specifications are Cobb-Douglas. The main implication for the analysis carried out in here is that expenditure switching forces are low as relative consumption and input shares remain constant.[17] This is consistent with the

[17] In addition to sales reallocation, non-linearities and second-order effects can be relevant in an frictional production network like the one considered here (Hulten, 1978, Baqaee and Farhi, 2020). Exploring these elements in a general equilibrium setting is left for future research.

short-run analysis on the effects of exchange rate fluctuations carried out in the paper. Expenditure switching forces are likely to occur in response to larger devaluations or over long horizons. Values of the elasticity of substitution in the range of 1-2 are chosen to describe aggregate import demand in the macroeconomic real business cycle literature, Ruhl et al. (2008). Low values of the elasticity of substitution are appropriate as relative price shocks due to exchange rate fluctuations are transitory and, thus, demand-side responses are likely to be limited.[18] Moreover, the product categories in the input-output tables are relatively aggregated compared to the standard disaggregation levels in trade data, making substitution across products relatively low.[19]

A key assumption is the reduced form treatment of the exchange rate pass-through into border prices. The reason is twofold: on one side, the aim and focus of the model are the description of domestic frictions and forces influencing the transmission of exchange rate fluctuations; on the other hand, the richness of the data available allows to directly and carefully disciplining the behavior of border prices. The main implications of accounting for incomplete pass-through into border price in a reduced form is the assumption of separability between the interactions with domestic and

[18] In the international real business cycle literature, matching the terms of trade volatility and the negative relationship between terms of trade and trade balance generally require low values of the trade elasticity, Hillberry and Hummels (2013).

[19] Reinert and Roland-Holst (1992) show that trade elasticities are particularly low across manufacturing sectors in the US, ranging between 0.25 to 3.5. More disaggregated data like those used in the international trade literature estimate a much larger trade elasticity, between 4 and 15, Hillberry and Hummels (2013).

with foreign suppliers. In other words, there are no strategic interactions between domestic and foreign suppliers.

Lastly, I take a stance on how the leading domestic frictions included are micro-founded. I followed Burstein et al. (2003) in modelling distribution costs. Compared to Corsetti and Dedola (2005), which use additive distribution costs, the qualitative implications on pass-through are the same but the calibration is immediate as the shares ϕ_is can be computed directly from the input-output tables. I also assume that distribution services are paid in labor and the distribution sector is competitive. The former implies that the share does not react to exchange rate changes, given the focus on the direct effects of exchange rates into prices. The latter implies that distributors do not charge markups, abstracting from double marginalization and additional incomplete pass-through due to variable markups.[20] Similarly, the micro-foundation of variable markups nominal rigidities follows standard choices in the macro and international economics literature and are compatible with the data available. Notice that abstracting away from nominal rigidities makes the effect of variable markups vanish because monopolistically competitive firms are symmetric. If nominal rigidities are absent, the change in price is identical for all firms. Thus, relative prices do not change

[20]Goldberg and Campa (2010) provide a raw estimate of the sensitivity of distribution services to exchange rate. They show that distribution margin slightly decreases following an exchange rate depreciation. However, an accurate product level calibration is difficult due to data limitations. The estimated effect of distribution costs can be considered as an upper bound.

and the effect of variable markups disappears.[21]

1.3 Calibration

A detailed calibration of the domestic economy is one of the goals and contributions of this paper. The measurement equations (15), (17) and (5) testify how different channels and frictions determine the sensitivity of CPI to exchange rates. Each element (distributions margins, variable markups and nominal rigidities, trade exposure and the granularity of the production network, and incomplete border pass-through rates) and their heterogeneity across products are carefully disciplined using a variety of micro-level data. The key ingredients are the 2013 "make" and "use" tables from the Central Bank of Chile, data from the survey of manufacturing from 2000 to 2007 (ENIA, *Encuesta Nacional Industrial Anual*) compiled by the Chilean National Statistical Agency (INE, *Instituto National de Estadisticas*), and the universe of Chilean import transactions from 2009 to 2019 from the Chilean Customs Agency (*Aduanas*).[22]

I now discuss in details the data and the strategy I use to calibrate each element of the main measurement equations and additional information is

[21] Departing from the symmetric firms case implies that firm-level pass-through depends on the covariance between markup elasticity and the cost shock, Amiti et al. (2019). Expanding the analysis to introduce within sector heterogeneity across firms requires additional firm-level data to discipline the covariance, representing a valuable venue for the future.

[22] I also use additional macroeconomic variables such as inflation rates, sectoral deflators, GDP growth rates, exchange rates from IMF, OECD or Central Bank of Chile.

provided in Appendix A. In the following Section, I show how a granular representation of the domestic economy and heterogeneity in frictions are key to accurately gauge the transmission of exchange rate fluctuations into domestic prices. This suggest that the strategy and the data I use can provide the basis for future calibrations and quantitative analyses.

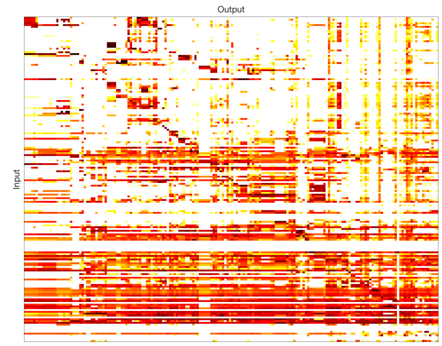
Domestic Network: S_m, S_d and β. I construct the input-output matrices for the Chilean economy combining the 2013 "make" and "use" tables provided by the the Central Bank of Chile. The tables consist of two basic national accounting tables: the "make" table shows the production of commodities by industry while the "use" table shows the use of commodities by intermediate and final users. The Central Bank of Chile also provides information on international flows, allowing the construction of international make (for imports) and use (for exports) tables. The tables are very disaggregated and include 180 products and 110 industries.[23]

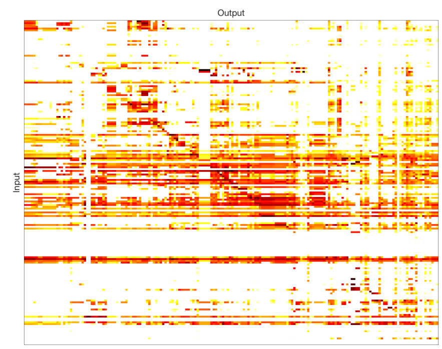
I combine the make and use tables under the industry technology assumption to construct a (180×180) product-by-product input-output matrix.[24] Each matrix quantifies how much of each product (row) is used in the production of other products (column). I also use the input-output tables to compute the share of each product in final consumption. This allows

[23] As a comparison, commonly used input-output tables as the WIOD or the OECD tables have around 30 to 40 industries. Pasten et al. (2020) shows that the granularity of the input-output table plays a central role in the quantification of the real effects of monetary policy, as less granular input-output tables tend to underestimate its effects.

[24] Appendix II provides details on the technical assumptions for the construction of the IO matrices.

FIGURE 1.2 – Domestic and International Leotief Matrices





The left (right) panel plots the domestic (international) input-output matrix of the Chilean economy in 2013. The matrices are computed using the make and use table under the industry technology assumption. Each row (column) represents an input (output). The intensity of the coloring shows how much one product is used as input in the production of other products: the darker (lighter) the color, the higher the input share. Log input shares smaller than -10 are censored.

me to calibrate the S_m and S_d matrices and the vector $\boldsymbol{\beta}$. Figure 1.2 reports the domestic and international input-output tables, S_d and (left) S_m (right) respectively, where a more intense color refers to a higher share of a certain input in the production of a given product. Importantly, domestic network is highly sparse and trade exposure is heterogeneous across products. Both elements play an important role in shaping the response of aggregate variables, Pasten et al. (2020).

Distribution margins: Φ. The distribution margin is computed as the ratio of the value of trade and transport margins to the value of total supply of that product at purchasers' prices:

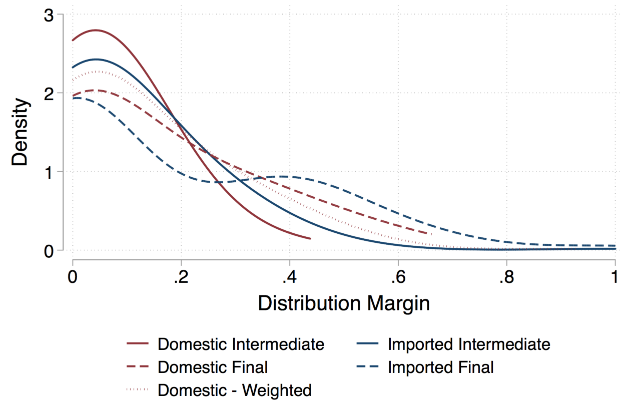
$$\phi_i \equiv \frac{\text{Retail + wholesale + Transportation costs}}{\text{Value at purchaser prices}} \equiv \frac{\text{Value at purchaser prices - value at basic prices}}{\text{Value at purchaser prices}} \tag{1}$$

Following Goldberg and Campa (2010), I use the input-output matrices for the Chilean economy to compute the value of trade and transport margins as the difference between the cost of supply (basic price) and the purchaser price.[25] The richness of the data allows me to compute not only heterogeneous margins across products but also across use (final vs intermediate consumption) and origin (imported vs domestic). In the model in Section 1.2, the price of domestic goods is the same independently of their use, final consumption vs intermediate input. Therefore, it is not possible to use the corresponding distribution margins. For each domestic product, I calibrate the common distribution margin as the expenditure-weighted average of the distribution margin for final and intermediate use. The same issue does not arise for imported products.

Figure 1.3 and Table 3.8 in Appendix II report the density distribution for different class of products (domestic vs imported, intermediate vs final). On one hand, domestically produced products tend to have lower distri-

[25] The Central Bank of Chile provides the make and use tables both at basic and purchaser prices. The latter is defined as the cost of supply plus retail, wholesale, transportation costs, and net taxes.

FIGURE 1.3 – Density of Distribution Margins



The figure plots the density distribution of the distribution margins across products. The distribution margins are computed according to Equation (1). I differentiate products depending: on their use, final vs intermediate use (solid vs dashed lines, respectively); on their origin, imported vs domestically produced (blue vs red lines, respectively). The dotted line shows the density distribution of the expenditure-weighted average of the distribution margin for final and intermediate domestic products.

bution margins compared to imported goods, consistent with the fact that internationally sourced goods are subject to larger transportation costs. On the other hand, intermediate goods also tend to have lower distribution margins. This suggests that lower pass-through due to distribution costs potentially arises at the end of the production chain, when products reach final consumers.

Markup elasticity: Γ. I use the Annual National Industrial Survey (ENIA) from 2000 to 2007 to estimate markup elasticities at the 3-digit industry level.[26] The theoretical model in Section 1.2 assumes a Kimball VES technology. For the main quantitative exercise, I further specify Equation (7) as-

[26] I match the estimated 3-digit industry level parameters with the product classification in the IO tables. It is possible that the same estimated markup elasticity is used for more than one product. For missing products, mostly in services, I use the estimated aggregate markup elasticity.

suming that the Kimball aggregator takes the form of a Klenow and Willis (2016) aggregator. I follow Gopinath et al. (2010) and Amiti et al. (2019) and calibrate the steady-state value of the markup elasticity:

$$\Gamma_i = \frac{\epsilon_i}{\sigma_i - 1}, \tag{2}$$

where the markup elasticity depends on two parameters, the industry-specific elasticity of demand, σ_i, and the super-elasticity of demand, ϵ_i.[27]

For each industry, I calibrate the elasticity of demand to match the revenue-weighted average estimated markup, $\bar{\mu}_i$, $\sigma_i = \frac{\bar{\mu}_i}{\bar{\mu}_i - 1}$. ENIA provides information on sales, inputs expenditures, employment and wage bill, investment, industry code (ISIC rev 3), for approximately 5000 plants per year with more than 10 employees. I estimate production functions and firm-level markups using state-of-the-art techniques and best practices, Levinsohn and Petrin (2003), Ackerberg et al. (2015) and De Loecker and Warzynski (2012). As robustness, I consider alternative measures of markups: I estimate markups using different definitions of variable input (cost of good sold vs labor only) and using the alternative cost share approach (Autor et al., 2020, De Loecker et al., 2016). Appendix III provides

[27]The markup elasticity of variety k in industry i takes the form of $\Gamma_{ik} = \frac{\epsilon_i}{\sigma_i - 1 + \epsilon_i \log\left(\frac{\widetilde{p_{ik}}}{\widetilde{p_i}}\right)}$, with $\widetilde{p_{ik}}$ and $\widetilde{p_i}$ being the price of variety k and the industry price index, respectively. Both Gopinath et al. (2010) and Amiti et al. (2019) calibrate it under the assumption that $\widetilde{p_{ik}} = \widetilde{p_i}$. Under this assumption, the markup elasticity can be interpreted as the steady-state markup elasticity, Gopinath et al. (2010), or the markup elasticity for an average firm, Amiti et al. (2019).

additional details on the estimation of production function and markups.

I follow Edmond et al. (2018) in estimating the super-elasticity parameter ϵ using the within-industry relationship between markups and market shares implied by the Klenow and Willis (2016) specification:

$$\frac{1}{\mu_{ik}} + \log\left(1 - \frac{1}{\mu_{ik}}\right) = a_i + b_i \log \texttt{share}_{ik}, \qquad b_i = \frac{\epsilon_i}{\sigma_i}, \tag{3}$$

where $\texttt{share}_{ik}$ is the market share of firm k in industry i. I estimate the slope coefficient b_i for each industry introducing firm and year fixed effects. Fixed effects are meant to control for unobserved productivity and quality (Edmond et al., 2018, Errico and Lashkari, 2022). I retrieve the superelasticity, ϵ_i, given the estimated demand elasticity.

Table 1.1 reports the estimated sectoral parameters (markup elasticity, demand elasticity and superelasticity) and summary statistics of the sectoral markup distributions. Estimated average and median markups are reasonable and in line with previous results from Chile, Levinsohn and Petrin (2003) and Garcia-Marin et al. (2019).[28] Importantly, the implied steady-state markup elasticities are in the range of values previously used in the literature and show substantial heterogeneity across sectors.[29] More-

[28]Figure 3.9 in Appendix III plots the distribution of markups across firms for each industry.

[29]Gopinath et al. (2010) vary the super-elasticity ϵ between $[0, 8]$, implying a Γ varying between $[0, 2]$, given a $\sigma = 5$. Consistent with the the chosen Kimball specification, the right panel of Figure 3.10 in Appendix III shows that the positive relationship between average markup and markup elasticity holds also across industries. The left panel of Figure 3.10 in Appendix III shows that there is no relationship between the average markups and the estimated superelasticity across industries.

TABLE 1.1 – Markup and Markup Elasticity

	Markup				Implied Parameters		
	Mean	Median	StD	Weighted Mean	σ	ϵ	Γ
Food Beverages and Tobacco	1.343	1.302	0.226	1.415	4.098	2.281	0.479
Textile and Apparel	1.274	1.262	0.186	1.301	4.266	1.672	0.498
Wood Paper and Printing	1.289	1.257	0.201	1.377	3.643	1.712	0.646
Petroleoum and Chemical Products	1.392	1.275	0.410	1.420	3.521	1.139	0.434
Plastic Rubber and Construction	1.292	1.262	0.209	1.391	3.930	2.546	0.578
Fabricated Metal	1.165	1.101	0.263	1.295	4.939	0.810	0.226
Machinery and Equipment	1.201	1.177	0.188	1.152	8.122	1.595	0.380
Motor Vehicle	1.088	1.119	0.265	1.047	13.18	7.582	0.486
Forniture	1.244	1.227	0.172	1.275	4.641	2.283	0.627
Aggregate	1.274	1.237	0.247	1.408	3.453	1.093	0.446

The table reports summary statistics of the estimated markups aggregated at the 2-digit sectoral level. Weighted-mean reports the average markup weighted by revenue. Markups are estimated using the survey of manufacturing (ENIA) from 2000 to 2007 and state-of-the-art production function estimation, Ackerberg et al. (2015) and De Loecker and Warzynski (2012). The table reports also the average implied demand elasticity (σ), super-elasticity (ϵ) and markup elasticity (Γ). Demand elasticity is calibrated to match the estimated revenue-weighted average markup. I follow Edmond et al. (2018) to estimate the demand super-elasticity leveraging the within-industry relationship between markups and market shares implied by the Klenow and Willis (2016) specification. Markup elasticity is defined as in Equation (2). Appendix A provides additional information on data and empirical specifications.

over, markups and the implied parameters are very similar independently of the markup estimation approach or variable input used.

Calvo probability: Δ. Due to lack of disaggregated domestic pricing data, I calibrate a common probability of price adjustment (Calvo parameter), δ, across all products.[30] I set the average monthly frequency of price adjustment to 30%, following the micro-level estimates of Aruoba et al. (2022) from confidential daily transaction data from the Chilean Tax Authority.[31]

[30] As shown in the following Section, heterogeneity in frictions is key in determining which products are the most important contributors to the transmission of exchange rate fluctuations. At this stage, the role of price rigidities cannot be fully explored and is left to future research.

[31] The frequency of price adjustment is slightly higher compared to the estimated value of $\approx 20\% - 25\%$ for the US, Nakamura and Steinsson (2008) and Pasten et al. (2020).

This implies an average quarterly probability of adjustment of 65%, with an average duration of about 2.8 months.

Pass-through into Border Prices: Ψ. Differently from domestic frictions, the model of exchange rate pass-through in Section 1.2 captures the role of (heterogeneous) incomplete pass-through into border prices in a reduced form, *via* Ψ_i. I use transaction-level import data from the Chilean Custom Agency (*Aduanas*) and follow previous work to discipline directly the pass-through into border prices, accounting for importers' heterogeneity.

Specifically, Ψ_i is disciplined at the product level accounting for the heterogeneity due to importers' experience. The aim is to capture the role that firm level determinants such age, size, market power and the presence of trade relationships have in shaping the pass-through rate of exchange rates into border prices. Alviarez et al. (2021), Juarez (2022) and Errico (2022) show that importers exert market power on their supplier and pay a markdown on the price they pay. This gives room to adjust the markdown following an exchange rate changes, keeping prices stable and lowering the exchange rate pass-through. Similarly, Dasgupta and Stiglitz (1988) and Heise (2019) show that relationship capital is accumulated as trade relationships grow older, influencing pricing and pass-through behavior. Moreover, an extensive literature on exporters' dynamics points to the role of market share, size and productivity in influencing pass-through rate into export prices (Atkeson and Burstein, 2008, Alessandria, 2009, Berman et al., 2012, Drozd and Nosal, 2012, Amiti et al., 2014). I document that importers with longer experience have larger market shares and face a lower pass-through rate into border prices.[32] I calibrate Ψ combining this empirical evidence.

The universe of import transactions provided by the Chilean Customs Agency includes, for each import transaction, standard information such as the importer's unique identifier (*importer*), the 8-digit HS product code (*product*), the date of the transaction, the country of origin (*origin*), the FOB and CIF values, the quantity shipped. I use data from 2009 to 2019; additional information on cleaning and summary statistics are reported in Appendix I.

I measure importers' experience constructing a measure of importing tenure at firm-product-origin level. I define the tenure of an importer-product-origin triplet as the number of quarters the importer has been consecutively importing a certain HS8 product from a given origin.[33] Importers with longer tenure are firms that have been consistently engaging in importing activities for longer periods of time.

Table 3.7 in Appendix I provides information on the distribution of im-

[32] Errico (2022) rationalizes this findings with through an open economy model of oligopsony that delivers consistent qualitative predictions. As importers grow older and larger, they gain experience in foreign markets which allows to exert stronger market power on their foreign supplier.

[33] As robustness, in Appendix B, I relax this definition of tenure and consider the number of quarters the importer has been importing a given product, dropping the consecutive requirement. I also consider the cumulative imported quantity for each firm-product-origin triplet.

porting tenure and the number of observations along different dimensions. Import flows are dispersed across firms, products and countries of origin, in line with previous literature (Eaton et al., 2021, Piveteau, 2021). The median importing firm records four flows per quarter, concentrated in one product or a couple of countries of origin. The second half of the table shows that the sparsity appears also along the time dimension. Importing is not a long-lasting activity as the median importing tenure across firm-product-origin triplets is one quarter. These statistics provide an overview of the prevalence of short import spells, and this is true using both definitions of tenure.

Fact I: Responsiveness of Border Prices. I augment a standard exchange rate pass-through regression to quantify the effect of importing tenure on the transmission of exchange rate fluctuations into border price (Heise, 2019, Errico, 2022). Let f index an importing firm, p an HS8 product category, o the country of origin, and t the quarter. The pass-through is estimated at quarterly frequency to be consistent with the Calvo probability Δ, also calibrated at the quarterly level. The baseline specification is:

$$\Delta \log p_{fpot} = \beta_1 \Delta \log e_{ot} + \beta_2 \log \text{Tenure}_{fpot} \times \Delta \log e_{ot} + \beta_3 X_{fpot} + \eta_{fop} + \nu_t + \varepsilon_{fpot}, \tag{4}$$

where $\Delta \log p_{fpot}$ is the price change of product po imported by firm f between quarter t and $t-1$, and $\Delta \log e_{ot}$ is the change in the Chilean peso-country o exchange rate between quarter t and $t-1$. Tenure$_{fpot}$ is the importing tenure at quarter t, defined as described in the previous section. In the main specification, I use the log of tenure to reduce the impact of the positive skewness in the distribution of tenure. I include time fixed effects and importer-product-origin fixed effects, meaning that the effect of importing tenure on the pass-through of exchange rate fluctuations, β_2, is estimated using the variation within the same import relationship over time.

X_{fpot} is a set of controls that includes the average size of the importer-product-origin triplet and an index of competitor price change. The former is used to control for differences in size and productivity, as larger firms may exhibit lower pass-through rates because of their size or stronger pricing to market behavior, Amiti et al. (2014) and Berman et al. (2012). The latter controls for strategic complementarities across importers. Following Amiti et al. (2019), I construct an index of competitor price change as a weighted average of the price changes of all other importers of the same product p:

$$\Delta \log p_{-ft} = \sum_{j \in F_p} \frac{S_{jt}}{1 - S_{ft}} \Delta \log p_{jt}, \tag{5}$$

where F_p refers to the set of importers purchasing product p from any origin. The shares S_{jt} are defined for each product p across all origins in terms of quantity. Given the potential endogeneity in the change of competitors' prices, I instrument the competitor price changes with movements in the bilateral exchange rates. As in standard pass-through regression, I control for the inflation rate in the origin country to control for changes in the production cost, Burstein and Gopinath (2014) and Goldberg and Campa (2010).[34]

TABLE 1.2 – Effect of Importing Tenure on ERPT into Border Prices

	(1)	(2)	(3)	(4)	(5)
$\Delta \log e$	0.2546	0.2711	0.3376	0.3880	0.3868
	(0.098)	(0.109)	(0.116)	(0.115)	(0.115)
Log Tenure X $\Delta \log e$			-0.0408	-0.0342	-0.0350
			(0.013)	(0.012)	(0.013)
Average Size X $\Delta \log e$				-0.0096	-0.0092
				(0.003)	(0.004)
Strategic $\Delta \log p_{-f}$					0.2718
					(0.313)
Time	Yes	Yes	Yes	Yes	Yes
Importer X Product X Country	No	Yes	Yes	Yes	Yes
Observations	2,568,634	2,368,422	2,368,422	2,368,422	2,365,619

Coefficients for terms in levels (log tenure, average size and inflation of origin country) and left and right censorship dummies are omitted. Standard errors clustered at country level. Tenure is defined as the number of quarters the importer has been consecutively importing a Product X Origin pair. Average size is defined as log average quantity traded at the Importer X Product X Origin level. Strategic is constructed according Equation (5).

Table 1.2 presents the estimates of the key coefficients of interest. Column (1) reports the estimated exchange rate pass-through rate from a standard regression, with no controls except for time fixed effects. The magnitude is comparable to Heise (2019) but falls short relative to standard estimates in the literature, which does not control for time fixed effects.[35] Using

[34]The macroeconomic variables used in the empirical analysis, such as inflation rates and exchange rates, are obtained from additional sources like the IMF, the OECD or the Central Bank of Chile.

[35]The average estimated magnitude in the literature is around 0.75 (Amiti et al., 2014, Gopinath et al., 2020). The discrepancy with the literature is explained by the presence of time fixed effect in the main specification in Equation (4). Table 3.10 in Appendix IV shows that removing time fixed effect provides an estimated pass-through rate of approximately 0.75, depending on the type of variation used, in line with the previous estimates from the literature. In addition, the specification without time fixed effects - Table 3.10 in Appendix IV - estimates a higher effect of importing tenure on the exchange rate pass-through than the one reported in Table 1.2. For this reason, the effects of heterogeneous border price pass-through in Section 1.4 should be considered as a conservative lower bound.

within importer-product-origin variation, Column (3) shows that each additional quarter in importing tenure reduces the sensitivity of border price. The estimated effect implies that an increase in importing tenure from the bottom quartile (25th percentile) to the top quartile (75th percentile), approximately 3-4 years difference in 2019, reduces the pass-through rate by approximately 5-6%, a substantial drop. Column (4) and (5) introduce additional controls. The qualitative and quantitative effect of importing tenure on pass-through is unaltered. In line with previous results in the literature, own average size reduces pass-through rates, Amiti et al. (2014). Similarly, the index of competitor price change shows the presence of strategic complementarity among importers.[36]

In Table 3.11 Appendix B, I analyze the sensitivity of the results in Table 1.2. In the first column, I run the baseline specification in Equation (4) using the preferred definition of tenure in levels. I show that results are quantitatively similar, and, as expected, a larger implied pass-through for lower values of tenure. In the second and third columns, I run the baseline specification in Equation (4) using alternative measures of importing tenure. I replace my conservative measure of tenure with the number of quarters since the first time the firm imported a specific HS8 product-country of origin pair. Alternatively, I use the cumulative quantity traded up to that quarter within each importer-product-origin triplet. Both measures have

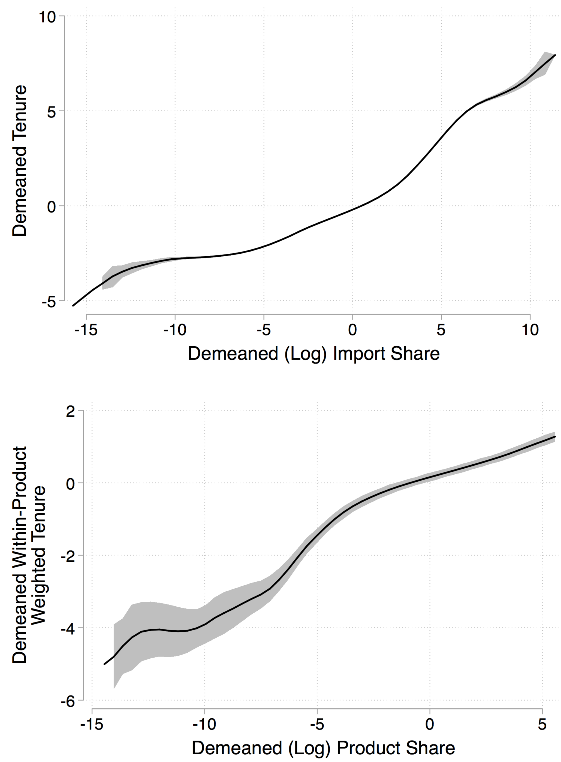
[36] Amiti et al. (2019) show that strategic complementarities are significant only for larger firms. This could explain why the average effect of strategic complementarities in Table 1.2 is not significant.

the same qualitative effects on the pass-through of exchange rate shocks. In the fourth column, I identify the effect of tenure on pass-through using the variation coming from different importing experience across different origins, within a firm-product pair. I use a combination of origin-product and firm-product fixed effects to substitute for the firm-product-origin fixed effects. Also in this case, the qualitative effect of importing tenure on pass-through is preserved. The remaining columns examine the sensitivity of my results with respect to the set of controls used in Equation (4). I run the baseline regression using different measures to control for the heterogeneity in firm size. I replace the average quantity of the importer-product-origin triplet and use the size of the importer computed as the total quantity traded across all product-origin pairs throughout the entire dataset or the quantity traded in each given quarter at the importer-product-origin level. Lastly, I construct alternative competitor price indices to control for strategic complementarities. I reconstruct the index in Equation (5) where the shares are computed using transaction values, rather than quantities. In addition, I use a more conservative definition of competitor and redefine the set of competitors of each importer at the product-origin, F_{po}, which includes all importers purchasing product p from origin o. In all these cases, I find similar results to the baseline specification.

Fact II: Market Share. Figure 1.4 shows non-parametrically that, at each point in time, products that are imported more intensively are also those where firms have, on average, longer importing tenure. The left panel uses the whole sample, defining tenure and market shares at the firm-product-origin-quarter level. The right panel aggregates the data, defining a product category at the 3-digit SITC level. In the latter, for each product, I compute the expenditure-weighted average tenure across all firms importing in that product category. Similarly, market shares now refer to the overall market share of the product category. Independently of the level of aggregation, I demean all variables at the quarter level to avoid the mechanical increase in tenure as time passes and make it comparable over time.

Figures 3.11 in Appendix B shows that the positive relationship between market shares and tenure is robust to different measures of tenure, variations and subsamples. Panel a) uses the less conservative measure of importing tenure, which is defined as the number of quarters since the first time the firm imported a specific HS8 product-origin pair. In panel b) and c), I demean the variables at the quarter and quarter-firm-product level, respectively. Finally, panel d) uses only the second half of the sample to avoid possible mechanical increases in average tenure. Similarly, aggregating tenure and market shares at the product level, Figure 3.12 shows that the relationship is robust to i) the measure of tenure used (panel a); the aggregation weighting (panel b uses simple averages across firm-origin pairs); the subsample considered (panel c uses the second half of the sample only); the

FIGURE 1.4 – Relationship Market Share - Importing Tenure



The left panel plots the non-parametric relationship between the (log) market share and the tenure in the whole sample. Market shares and tenure are defined at the firm-product-origin-quarter level. Products are defined at the 8-digit level. Variables are demeaned at the quarter-firm level. The right panel plots the non-parametric relationship between the (log) market share of a product and the expenditure-weighted average tenure across all firms importing that product. Products are defined at the 3-digit SITC level. Share and average tenure are computed at the quarterly level. Variables are demeaned at the quarter level. The panels show the 99% confidence intervals.

aggregation level (panel d aggregates at the 5-digit level). In all these cases, I find similar results to the baseline relationship documented in Figure 1.4.

Disciplining Ψ. I combine the empirical facts documented above to discipline the heterogeneous sensitivity of border prices, Ψ_i. Fact I and II imply that imported products with higher market shares are also those with lower exchange rate pass-through rates into border prices because importers with longer importing tenure are relatively more active.

I calibrate a baseline incomplete pass-through rate to be 0.75. The value is estimated using the Customs data and the regression in Equation (4) after dropping the time fixed effects. The estimated magnitude is in line with previous estimates from the literature and reported in Table 3.10 in Appendix IV. This value represents the exchange rate pass-through into border price of a product that exhibits zero importing tenure.

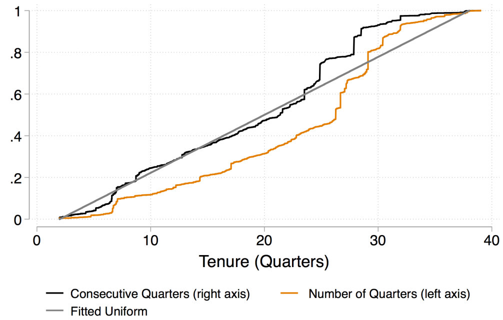
I calibrate heterogeneous pass-through rates across imported products using the estimates on the effect of importing tenure on the pass-through of exchange rate fluctuations (Fact I - Table 1.2). Figure 1.5 shows that, in 2019, the preferred measure of importing tenure aggregated at the 5-digit product level ranges between 1 a 40 quarters.[37] Given the estimated effect of importing tenure, this implies an heterogeneous pass-through ranging between 0.6 and 0.75.[38] Figure 1.5 shows that the cumulative distribution of importing tenure across 5-digit products closely resembles a uniform distribution. Thus, I evenly distribute product-level pass-through rates in the rage $[0.6, 0.75]$.[39]

[37]My analysis focuses on the effect of tenure on pass-through across products, not dynamically. I choose the distribution of importing tenure in 2019 interpreting it as the stationary distribution of importing tenure across products. In addition, choosing 2019 makes the quantification of the event study - the 2019 "*Estallido Social*" - more accurate.

[38]In order not to underestimate the effects of high levels of importing tenure, I use the estimated coefficient for the effect of tenure in level, column (1) in Table 3.11, rather than in logs (Table 1.2). This implies that the lowest pass-through rate is $0.75 - 0.0038 \times 40 \approx 0.75 - 0.15 = 0.6$. Using the coefficients in logs delivers a slightly higher lower bound.

[39]The cumulative distribution is very close to a uniform distribution except for very high value of importing tenure. Assuming a uniform distribution slightly overestimates the effect of products with high tenure.

FIGURE 1.5 – Cross-product Distribution of Tenure



The figure plots the cumulative distribution of average importing tenure at the product level. I consider 5-digit SITC product categories. The average importing tenure for each product is computed as the expenditure-weighted average tenure across all firm-origin pairs. The black (orange) line plots the most preferred (alternative) definition of importing tenure, as defined in Table 3.7 in Appendix I. The solid gray line represents a uniform distribution over the range of importing tenure. The figure uses data from 2019 only.

I leverage the positive relationship between market share and importing tenure (Fact II - Figure 1.4) to allocate the heterogeneous pass-through rates across imported products. Imported products with larger market shares are those with higher average importing tenure and, therefore, with lower pass-through.

In Section 1.2, I assume that sectoral goods are used for both final consumption and as intermediate inputs. The same price elasticity applies to both direct exposure (final consumption) in Equation (6) and indirect exposure (intermediate inputs) in Equation (15). In the empirical quantification, the imported sectoral goods used both as final consumption and intermediate inputs are considered separately, calibrating two different pass-through rates depending on their use.

1.4 Empirical Results

I quantify the importance of domestic frictions and border price dynamics for the sensitivity and insensitivity of domestic prices to exchange rate fluctuations.

I show that domestic frictions are quantitatively more relevant than border price sensitivity in explaining the insensitivity of domestic prices. Moreover, I find that all domestic frictions are individually relevant for the low responsiveness of domestic prices. I quantify the relevance of domestic frictions and incomplete border price pass-through during the sharp depreciation of the Chilean peso following the "*Estallido Social*" event in 2019, showing that the former (latter) insulated domestic prices reducing inflation by 0.6 (0.3) p.p. at quarterly level.

Similarly, domestic frictions determine the sources of sensitivity of domestic prices. Contrary to previous results in the literature, I find that most of the CPI sensitivity arises through changes in the price of imported final goods (direct exposure) because domestic frictions dampen relative more the response of domestically produced goods. Moreover, the interaction between the heterogeneity in frictions, import exposure and consumption share influences the overall response of CPI and the contribution of individual products.

The low sensitivity of border prices still plays a substantial role in explaining the low sensitivity of CPI to exchange rate fluctuations, even after accounting for domestic frictions. Using back-of-the-envelope calculation and my estimates from Section 1.3, I quantify the aggregate effects of micro-level determinants of exchange rate pass-through into border prices. I show that the increase in average importing tenure from 2009 to 2019 can account for 40% of the decline in the aggregate sensitivity of domestic prices to exchange rate fluctuations.

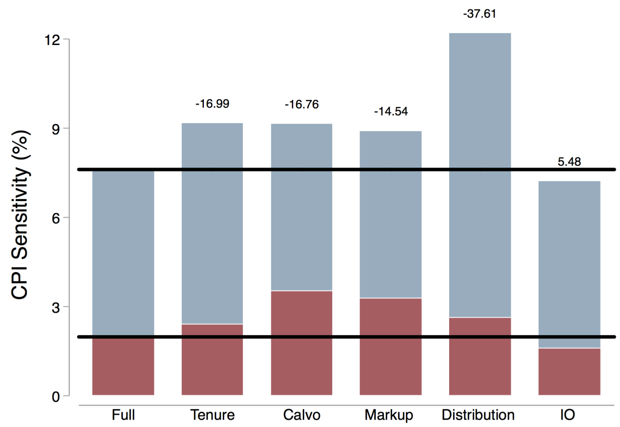
I conclude discussing what these results imply for inflation targeting and monetary policy in open economy, and future modelling and calibration exercises.

1.4.1 Role of Individual Mechanisms

I now present the first quantitative result: all mechanisms operating in Equations (15) and (17) but the presence of domestic input-output linkages are quantitatively relevant in shaping the response of domestic prices to exchange rate fluctuations.

I proceed by studying the response of domestic prices and CPI to a positive change in the exchange rate (depreciation of the Chilean peso). The baseline economy is a fully calibrated economy in which all mechanisms - distribution costs, variable markups and nominal rigidities, domestic input-output linkages and heterogeneous border price sensitivities - are active at

FIGURE 1.6 – Role of Individual Mechanisms



The figure plots the CPI sensitivity to a one percent depreciation in the exchange rate for different cases. The first bar (Full) refers to the fully calibrated model which includes incomplete and heterogeneous pass-through, input-output linkages and all domestic frictions (distribution costs, variable markups and nominal rigidities). All the other bars refer to an economy that abstracts away from one element at the time. For instance, the bar "Calvo" represents a fully calibrated model that omits the role of nominal rigidities. I scale all the numbers by 100. Notice that all scenarios use the same input and consumption shares to be as comparable as possible. The red (blue) part of each bar accounts for the part of sensitivity arising from indirect (direct) exposure as defined in Equation (6). The horizontal lines refer to the Full model implied sensitivities. The numbers on top of each bar represents the difference between the fully calibrated model and each alternative scenario.

the same time. I then assess the importance of each individual mechanism shutting down one mechanism at the time and quantifying the response of domestic CPI when abstracting away from it.

Each mechanism considered (Ψ, Δ, Γ and Φ) substantially dampen the response of domestic prices after a depreciation. Figure 1.6 reports the sensitivity of domestic CPI in the fully calibrated (Full) and in the five different economies in which one mechanism is shut down. Abstracting away from distribution costs implies the larger departure from the full model as domestic CPI is 37% less responsive. Variable markups and nominal rigidities equally insulate domestic CPI from exchange rate fluctuations, respectively 17% and 15% lower. Distribution costs play a larger role than variable markups and nominal rigidities because they affect the retail price of both imported and domestically produced goods, while variable markups and nominal rigidities influence only the latter. Domestic prices in the Full model are 17% less responsive than in an economy that abstracts away from heterogeneous border price sensitivity and experienced importers (Tenure). The quantitative relevance of importing tenure shows the importance of adjusting import exposure for the presence of experienced importers, as the latter influence the sensitivity of import price.

Lastly, the presence of domestic input-output linkages increases CPI sensitivity as shocks are propagated through the domestic network by round-about linkages, but the effect is negligible. Figure 1.6 shows that the amplification mechanism increases CPI response by 5% only. The amplifying role is dampened by the presence of multiple frictions in the domestic network and has key implications for the sources of CPI sensitivity, as explored in Section 1.4.5.

1.4.2 Decomposing CPI Insensitivity

How sensitivity is CPI in a frictionless world where all costs shocks are passed entirely into prices? Answering this question gives us a benchmark to understand how insensitive is domestic CPI to exchange rate fluctuations and provide additional information on the relative importance of domestic frictions and border price insensitivity.

I again proceed by quantifying the response of CPI to a one percent depreciation of the exchange rate across different scenarios. I calibrate six different cases in which I add one channel at a time to develop step-wise intuition. Table 1.3 lists the different combinations of pass-through into import prices and domestic frictions I study. The benchmark economy is a frictionless economy (i.e. no distribution costs, variable markups and nominal rigidities) that includes input-output linkages and in which the pass-through rate into border prices is complete. On the contrary, Case V considers a fully calibrated economy that includes all frictions, input-output linkages and in which the pass-through rate into border prices is incomplete and heterogeneous.

Figure 1.7 shows that the fully calibrated model predicts a CPI sensitivity extremely close to the estimate for the period from 2009 to 2019 while a frictionless benchmark economy largely overestimates it. The full model implies a sensitivity that falls in the range of estimated sensitivities for Chile, supporting the validity of the measurement equation in Section 1.2 and showing that its simplicity is not coming at the expenses of quantitative performance. The implied sensitivity in the benchmark economy is four times larger than the estimated one (29.3% vs 7.62%). As expected, abstracting away from all elements that dampen the transmission of costs shocks increases the sensitivity of domestic prices.

TABLE 1.3 – Overview of Calibration Cases

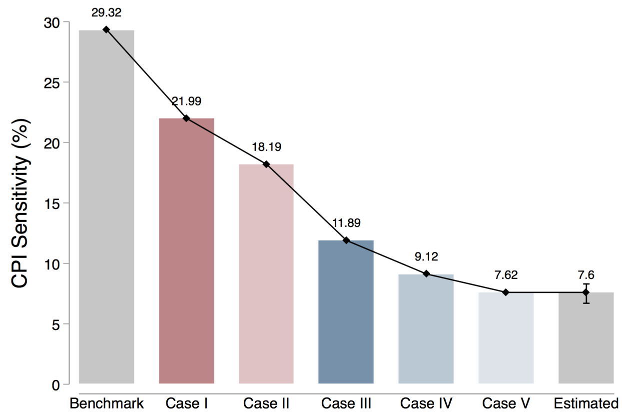
	Pass-through into Import Prices		Domestic Frictions			
	Average Ψ	Heterogeneous Ψ (Tenure)	Φ	Γ	Δ	IO Linkages
Benchmark	Complete					✓
Case I	Incomplete					✓
Case II	Incomplete	✓				✓
Case III	Incomplete	✓	✓			✓
Case IV	Incomplete	✓	✓	✓		✓
Case V ("Full")	Incomplete	✓	✓	✓	✓	✓

The table details the assumptions on pass-through into border prices, importing tenure, domestic frictions and input-output linkages for the different cases considered in the calibration. Notice that all scenarios use the same input and consumption shares.

Figure 1.7 shows that most of the insensitivity of CPI is due to domestic factors, i.e. mechanisms that do not operate on border prices. Including homogeneous incomplete pass-through rate into import prices reduces the sensitivity of domestic prices by 25% (22/29.3). Accounting for heterogeneity in border price sensitivity further reduces domestic price sensitivity by another 18%, 18.2/22. However, the effect on border prices falls short in matching the estimated CPI sensitivity as less than 50% of the gap between the estimated value and the benchmark economy is closed. Applying the same reasoning to the domestic frictions considered — distribution costs, variable markups and nominal rigidities — reduces domestic price sensitivity by approximately 35%, 25% and 17%, respectively.[40] All together,

[40]Figure 1.7 also provides additional evidence on the relative importance of each individual mechanism considered. Consistently with Figure 1.6, all channels considered contribute substantially to the overall aggregate insensitivity of domestic prices and the relative importance is qualitatively the same. In Appendix C, I show that the specific order does not changes the qualitative predictions of the relative importance of each mechanism.

FIGURE 1.7 – Decomposing CPI Sensitivity



The figure plots the aggregate CPI sensitivity to a one percent depreciation in the exchange rate for different cases. See Table 1.3 for a description of the different cases. I scale all the numbers by 100. The last column, "Estimated", reports the estimated CPI sensitivity to exchange rate estimated at the quarterly level from 2009 to 2019 (also scaled by 100). The bands refer to the range of estimated CPI sensitivity across different specifications in terms of lags and controls. Appendix C provides additional details on the estimation.

domestic frictions are quantitatively more relevant in dampening the sensitivity of CPI than incomplete pass-through to border prices. This shows how the response of border prices and the presence of domestic frictions need to go hand in hand to fully characterize the response of domestic CPI to exchange rate fluctuations.

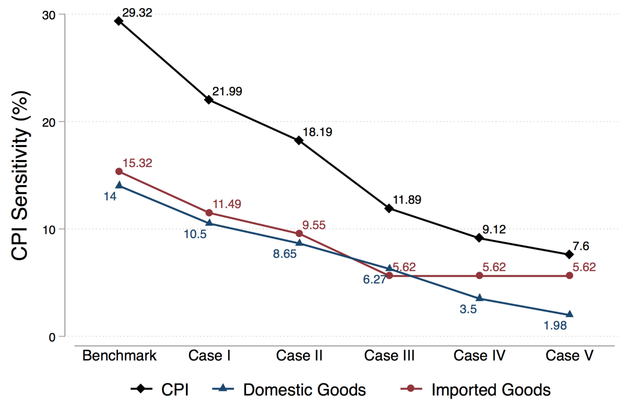
1.4.3 Direct vs Indirect Exposure and Input-Output Linkages

In contrast with previous work, Figure 1.6 documents that, in the fully calibrated economy and across all the scenarios considered, the bulk of the CPI response to a depreciation shock comes from the direct exposure of CPI to exchange rates, Equation (6). Direct exposure, i.e. imported final consumption (blue area in Figure 1.6), accounts for approximately 75% of the overall sensitivity in the fully calibrated case even though imported consumption represents only 15% of the total final consumption basket.[41] This results is at odds with previous work, that tends to assign the same importance to direct and indirect exposure (Goldberg and Campa, 2010, Burstein et al., 2003, Gopinath, 2015). I now investigate the conflicting results on the role of direct and indirect exposure and argue that standard quantification exercises tend to overestimate the contribution of imported intermediate inputs because they abstract away from a careful calibration of (heterogeneous) domestic frictions. In doing so, I also explore the role of input-output linkages as determinant of indirect exposure.

The importance of direct exposure is usually overestimated by the omission of domestic frictions that mainly alter the response of domestically produced goods. Figure 1.8 shows that, as more domestic frictions are considered, not only CPI becomes less sensitive to exchange rate fluctuations, but the sensitivity of CPI is increasingly driven by imported final consumption goods ("direct exposure"). In a frictionless economy ("Benchmark"), direct and indirect import exposure equally contribute to the overall price change. Introducing (heterogeneous) incomplete pass through into border prices does not alter the relative importance of the two types of exposure. However, the relative importance changes when domestic frictions are introduced (Case IV and V) as they influence only the sensitivity of domesti-

[41] In comparison, imported inputs in the production of domestic goods account for 25% of total inputs.

FIGURE 1.8 – Decomposing Aggregate CPI Sensitivity



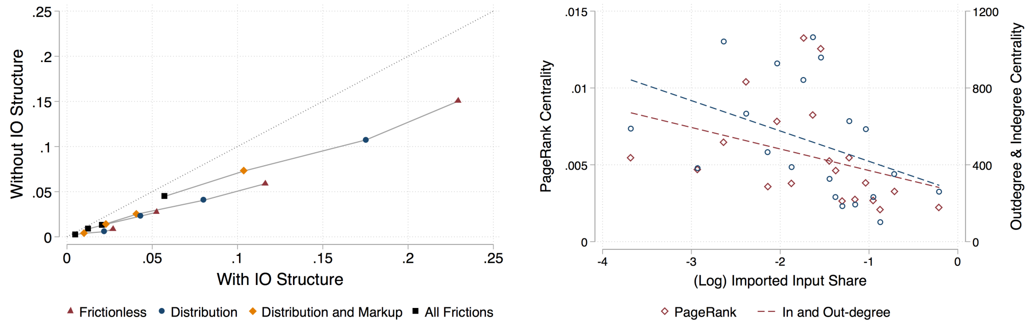
The figure plots the aggregate CPI sensitivity to a one percent depreciation in the exchange rate and its decomposition into imported final consumption ("Imported"), i.e. direct exposure, and domestic final consumption (”Domestic”), i.e. indirect exposure, for different cases. See Table 1.3 for a description of the different cases. I scale all the numbers by 100.

cally produced goods. Standard practices do not account for the presence of domestic frictions, and quantify direct exposure in frameworks comparable to the frictionless economy case (“Benchmark").

Indirect exposure originates also from the domestic input-output production network. Even though a domestically produced good does not make direct use of imported intermediate inputs, the domestic inputs used in its production could be exposed to imports. Figure 1.6 shows that the contribution of roundabout production is actually modest, as abstracting away from input-output linkages reduces CPI sensitivity to exchange rate by only 5%. The presence of domestic frictions and the centrality of import exposure are key to understand the small role of input-output linkages.

As more frictions are included in the domestic economy, the amplification generated by the presence of input-output linkages shrinks (Basu,

FIGURE 1.9 – Role of IO network



The left panel compares the evolution of the price of domestic products in an economy that includes input-output linkages (x-axis) to the evolution in an economy that abstracts away from input-output linkages (y-axis), as more domestic frictions are considered. Each series plots the median price change in each quartile of the distribution. I consider the following scenarios: "Frictionless" refers to the absence of domestic frictions; "Distribution" includes only distribution margins; "Distribution and markups" includes both distribution and variable markups; "All Frictions" includes distribution, variable markups and Calvo frictions. In all scenario, pass-through into import prices is incomplete and heterogeneous due to importing tenure. The dotted line shows the 45 degree line. Table 3.12 in Appendix C reports the CPI sensitivity for all scenarios considered in the presence of and abstracting away from input-output linkages. It also reports the decomposition between direct (imported final consumption) and indirect exposure (imported intermediate inputs). The right panel shows the relationship between the centrality of a product in the domestic production network and the share of imported inputs in its production. I consider the PageRank centrality measure (left axis) and the average of the in-degree and out-degree measures (right axis). Centrality is measured weighting the edges according to the input-output linkages. The share of imported inputs is computed over total costs from the IO tables. The dashed line shows a linear fit. Table 3.15 in Appendix C reports the corresponding coefficient. Section 1.3 and Appendix A provide additional details on the IO tables.

1994, Pasten et al., 2020). The left panel of Figure 1.9 compares the sensitivity of domestic prices in the case of roundabout production (x-axis) and without roundabout production (y-axis). I show that the median change in domestic prices in each quartile of the distribution is higher when roundabout production is considered as shocks are amplified through the network (Acemoglu et al., 2016). However, propagation diminishes as frictions are introduced in the economy. The intuition is that domestic frictions reduce price responsiveness and, thus, downstream propagation at any point in the network (Carvalho and Tahbaz-Salehi, 2019).[42]

Moreover, the right panel of Figure 1.9 shows that imported inputs are not central in the production network of domestic goods. I measure the centrality of a product in the domestic input-output network using both the PageRank centrality measure and the average between the In-degree and Out-degree measures. Centrality measures are used to assess the relative importance of each node in networks.[43] Products that are more central rely relatively less on imported inputs, therefore reducing amplification forces.

These results suggest that evaluating the role of import exposure for the transmission of exchange rate fluctuations to domestic prices requires both incorporating domestic frictions and detailed production networks. Common practice in calibrating aggregate models is to compute import exposure as the sum of direct and indirect exposure, where the latter is commonly computed from dense input-output tables (Burstein et al., 2003, Gopinath, 2015, Pasten et al., 2020). However, omitting domestic frictions results in overestimating the role of indirect exposure and, thus, CPI sensitivity.

[42] The (adjusted) Leontief inverse matrix captures direct and indirect downstream propagation. Abstracting away from domestic frictions implies using the Leontief inverse matrix rather than the adjusted one in Equation (15), where the former implies a stronger amplification.

[43] In-degree (out-degree) centrality counts the number of ties directed to (from) the node, quantifying the relevance of a node in the immediate vicinity. As standard practice I take the average of the two. PageRank centrality is a variant of eigenvector centrality, which weights the linked nodes by their centrality. In my sample, the two measures are highly correlated (65%). In both cases, edges are weighted according to the input shares forming the input-output tables (see Appendix II). No frictions are considered in the weighting. Figure 3.19 in Appendix C graphically represents the production network, the centrality and import intensity of each node.

1.4.4 *The 2019 "Estallido Social".*

The "*Estallido Social*" (social outburst) refers to a series of massive and severe riots originated in Chile between October 2019 and March 2020. From the perspective of my analysis, the riots triggered a major devaluation of the Chilean peso against all major currencies and make the event a natural laboratory to study the effects of domestic frictions on domestic prices.

Figure 3.16 in Appendix C documents the timing and the evolution of the shock using the Google index for protests: riots do not constitute an expected event and is short-lived.[44] Following the social outburst, the Chilean peso sharply depreciates with respect to all major foreign currencies. Political and social tensions increase uncertainty and risk, putting pressure on the value of the Chilean peso. The three-month depreciation rate of the Chilean rate peaks at 12% in mid November, right before the Central Bank of Chile intervention on the currency market to stabilize the value of the currency.[45]

I use the model to gauge the response of domestic prices to the sharp depreciation triggered by the shock, assessing the insulating effect of domestic frictions and border price insensitivity. I first quantify the implied

[44] The protests started in the capital, Santiago, on October 6 after subway fares rose by 4%. The increase in subway fares was the trigger of the protests, but high costs of living and socio-economic inequality represent the deeper roots of the social outburst. The riots quickly escalated and spread across the entire country, though with different levels of intensity (Aruoba et al., 2022). `https://www.bloomberg.com/news/articles/2021-07-06/investors-look-abroad-amid-political-tensions-chile-market-chat`

[45] The Central Bank of Chile used around $24bn in open market operations in the period between 2019Q3 and 2020Q1.

rise in domestic prices following the depreciation of the Chilean peso using the fully calibrated economy. I compare the prediction from the fully calibrated model to two counterfactual scenarios: one economy that includes only domestic frictions; another economy that accounts for incomplete border price pass-through only. I consider three different scenarios in measuring the quarterly depreciation rate of the Chilean peso (column (1) in Table 1.4). In the most conservative scenario, I consider the average quarterly depreciation in the last quarter of 2019 with respect to the third quarter of the same year, which is 5.6% ("*Average*"). Alternatively, I consider the peak depreciation rate during the last quarter of 2019, which is about 12% ("*Peak*"). Finally, to account for lagged response of the exchange rate and domestic prices, I consider also the cumulative depreciation of the Chilean peso over the 2019Q4-2020Q1 period with respect to the third quarter of 2019 ("*Cumulative*").

Domestic frictions insulate domestic prices more than the insensitivity of border prices during the depreciation of the Chilean peso. The fully calibrated model predicts an increase in domestic prices which accounts for about 30% to 90% of the actual inflation rate in Chile during the time period considered, depending on the scenario (column 2 and column 3).[46] In the "*Average*" scenario, a counterfactual economy with incomplete and heterogeneous pass-through into border prices but without domestic frictions

[46] These numbers are sensible considering that the average quarterly inflation in the previous 4 quarters was 0.5%. Additional inflationary forces in the economy can explain the remaining part.

TABLE 1.4 – "*Estallido Social*" and Counterfactual

			Full	W/out Domestic Frictions		Complete Border Price PT	
	Depreciation	Actual π	Imported π	$\widehat{\pi}$	% Change	$\widehat{\pi}$	% Change
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Average	5.61	1.02	0.43	1.61	58.2	1.28	25.3
Peak	11.8	1.02	0.90	2.27	122.9	1.56	53.4
Cumulative	10.8	2.32	0.82	3.46	49.3	2.82	21.4

The table reports back-of-the-envelope calculations on the relative importance of domestic frictions and incomplete border price pass-through on domestic inflation during the 2019 "*Estallido Social*" in Chile. Each row corresponds to a different scenario in terms of Chilean peso depreciation rate. Column 1 shows the depreciation rate corresponding to each scenario. Column 2 reports the actual quarterly inflation rate (in %) corresponding to each scenario. Column (3) quantifies the implied inflation (in %) following the depreciation of the Chilean peso using the fully calibrated model. Column 4 (6) quantifies the counterfactual domestic inflation (in %) in an economy that includes incomplete pass-through into border prices (domestic frictions) and abstracts away from domestic frictions (incomplete pass-through into border prices). Column 5 (7) quantifies the percentage difference between the counterfactual inflation rate in column 4 (6) relative to the actual inflation rate in column 2.

(columns 4 and 5) predicts the inflation rate to be 0.6 p.p. higher (approximately 50% higher) than the actual inflation rate, a sizeable difference at the quarterly level. Domestic inflation is 0.3 p.p. higher (approximately 25%) in an economy with domestic frictions but complete pass-through into border prices (columns 6 and 7), half as much as the effects of domestic frictions. As expected, domestic inflation have stronger insulating effects than incomplete pass-through into border prices. This confirms the importance of both border prices dynamics and domestic transmission for the response of domestic prices to exchange rate changes.

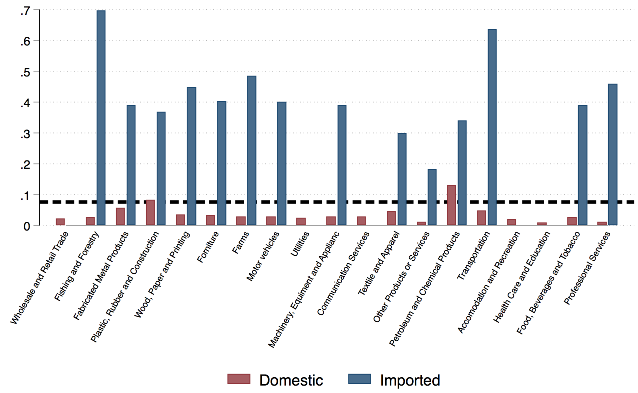
1.4.5 Heterogeneity across Products and Identity Effects

Focusing on the cumulative effects masquerades substantial heterogeneity across products. Moreover, the interactions of different dimensions of heterogeneity, such as heterogeneity in domestic frictions, import exposure and consumption share, play a crucial role for both the aggregate response and the relative contribution of different products to the CPI response.

Figure 1.10 graphically illustrates substantial heterogeneity in the sectoral response to the common exchange rate depreciation shock in the fully calibrated economy. Crucially, imported final goods are more sensitivity than domestically produced goods, consistent with the fact that direct exposure accounts for the bulk of the sensitivity of CPI. Moreover, within each category - domestic and imported goods - sectoral goods exhibit very different patterns in terms of sensitivity. For instance, among domestically produced goods, accommodation and service sectors are insensitive to exchange rates compared to the chemical and rubber sectors, as the latter are more exposed to imported inputs.

Figure 1.11 documents that not only the presence of frictions, but also their heterogeneity, is relevant to understand the low sensitivity of domestically produced goods. The left (right) panel shows a positive correlation between the share of imported inputs in production and the markup elas-

FIGURE 1.10 – Sectoral Heterogeneity

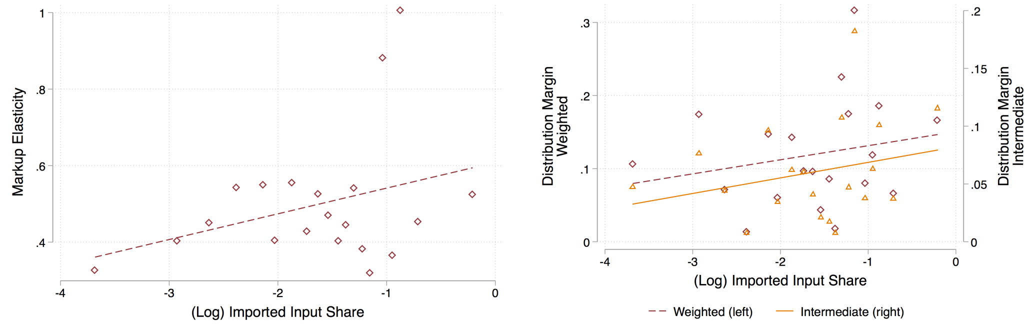


The figure plots sensitivity of prices across different 2-digit industries. Price sensitivity is computed in the fully calibrated model. I distinguish between imported final consumption (blue bars) and domestic final consumption (red bars). For each sectors, I compute the expenditure-weighted average sensitivity across products. Sectors are in ascending order (left to right) in terms of consumption shares. The dashed line presents the sensitivity of CPI.

ticity (distribution costs). The heterogeneity in frictions and their positive correlation with imported inputs make the role of frictions even more relevant for the overall response of CPI: the dampening effects are stronger for those products that are more relevant for the transmission of exchange rate fluctuations. Similarly, Figure 3.20 in Appendix C shows that ignoring heterogeneous consumption shares matters for aggregate sensitivity (Chen, Devereux, Shi and Xu, 2022). Domestic products that have larger consumption shares are also those that are less sensitive to imports and, thus, to exchange rate fluctuations.

The identity of the most relevant products for the overall sensitivity changes when different dimensions of heterogeneity are considered. The heterogeneity in sensitivity across domestically produced goods arises because of the heterogeneity in the exposure to imported inputs, domestic fric-

FIGURE 1.11 – Import Exposure and Friction Heterogeneity

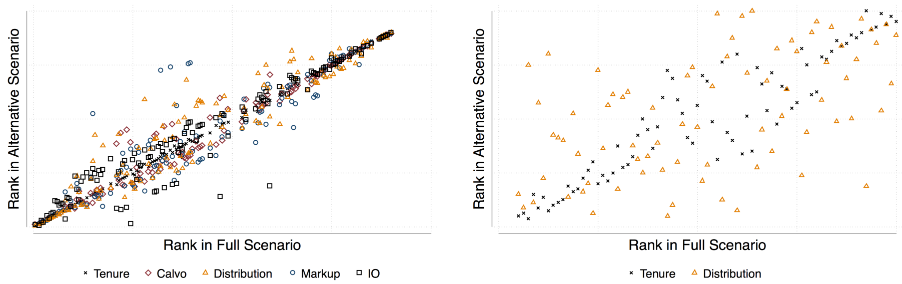


The left panel plots the relationship between the share of imported inputs in production and the markup elasticity for the set of domestically produced goods. The share of imported inputs is computed as the ratio between the total expenditure on all imported goods used in production and the total costs of production. The right panel plots the relationship between the share of imported inputs in production and the distribution margin for the set of domestically produced goods. The distribution margin is computed for domestic intermediate inputs only or as a weighted average between domestic intermediate inputs and final consumption goods. The dashed lines show linear fit. Table 3.15 in Appendix C reports the corresponding coefficients. Section 1.3 and Appendix A provide additional details on how import shares, distribution margins and markup elasticities are computed. Log imported input shares smaller than -10 are dropped.

tions and border price sensitivity. Interacting different dimensions of heterogeneity translates into different relative contributions across products. Figure 1.12 shows how the ranking of the products contributing the most to the overall CPI sensitivity changes depending on the frictions considered.[47] Compared to the fully calibrated model, shutting off one dimension of heterogeneity can substantially alter the ranking across products. The effect is pronounced i) when omitting distribution costs and ii) for imported final consumption (right panel). Table 3.14 in Appendix C shows that the changes in ranking are not correlated across scenarios, suggesting that different dimensions of heterogeneity impact each product in different ways.[48]

[47] I consider the scenarios of Figure 1.4.1, by shutting down one element at the time between distribution cost, variable markups and nominal rigidities, IO linkages and hetero-

FIGURE 1.12 – Ranking of Products



The figure compares the ranking of the products contributing the most to the overall CPI sensitivity in the fully calibrated model (x-axis) to the ranking in an alternative scenario (y-axis). For domestically produced goods (left panel), I consider the following alternative scenarios: a fully calibrated economy that omits, one at the time, the role of the heterogeneity in border price sensitivity, nominal rigidities, distribution costs, variable markups, and input-output linkages. For imported goods (right panel), I consider the following alternative scenarios: a fully calibrated economy that omits, one at the time, the role of the heterogeneity in border price sensitivity, and distribution costs.

1.4.6 Heterogeneous Pass-through into Border Prices

While domestic frictions are important for the (in)sensitivity of domestic prices, incomplete pass-through into border prices still plays a substantial role in explaining the low sensitivity of CPI to exchange rate fluctuations. In this section, I extend the analysis on the role of heterogeneous pass-through into border prices and show that the rise in average importing tenure accounts for 40% of the decline in domestic price sensitivity over the period 2009-2019.

A growing literature documents a decline in the sensitivity of domestic prices to exchange rate fluctuations across several advanced economies

---

geneous border price sensitivity.

[48]Table 3.15 in Appendix C shows that centrality and individual frictions (variable markups and distribution costs) do not mutually exclude each other and have comparable correlations with import exposure, suggesting that jointly accounting for all these elements is key to quantify CPI sensitivity.

since the late 1980s. Several papers consider the rise of global value chains and the stability of international trade relationships as possible explanations for the decline in exchange rate pass-through (Campa and Goldberg, 2005, Camatte et al., 2021, Georgiadis et al., 2020). The effect of importing tenure on exchange-rate pass-through into border and domestic prices points in the same direction.[49]

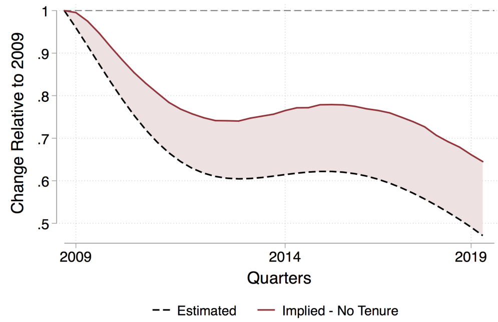
I find that the sensitivity of domestic prices to exchange rate changes decreases in Chile, complementing the recent evidence from advanced economies (Camatte et al., 2021). The dash line in Figure 1.13 plots the estimated trend from 2007 to 2020 using a 5-year rolling window. The pass-through into CPI decreases by 50% relative to 2009.[50]

Using back-of-the-envelope calculations based on my estimates, I compute the contribution of the rise of importing tenure to the decline in CPI sensitivity. Relative to the beginning of 2009, when importing tenure is normalized to one quarter, Figure 3.17 in Appendix C shows that the expenditure-weighted average importing tenure increased to 18 quarters. I quantify the change in CPI sensitivity driven by the increase in importing tenure using the estimated effect of importing tenure on border prices (Ta-

[49] Figure 1.6 shows that CPI is 16% less sensitive in a fully calibrated economy compared to an economy that abstracts away from the effects of importing tenure. Table 3.13 in Appendix shows that its contribution is quantitatively similar across multiple combinations of alternative frictions.

[50] Figure 3.15 in Appendix C shows that the decline is just part of a long-run negative trend started in the 70s. CPI sensitivity to exchange rates is initially around 0.35%, and reaches a value of 0.07-0.1% in the last decade. I estimate a trend because exchange rate pass-through rates at quarterly level are particularly noisy. Appendix C provides additional details on the estimation.

FIGURE 1.13 – Trend in ERPT and Contribution of Tenure



The figure plots the estimated trend in CPI sensitivity to exchange rates (dash line) and the counterfactual trend in CPI sensitivity to exchange rates abstracting away from the rise in importing tenure. The trend is estimated using a polynomial approximation of the series of estimated exchange rate pass-through rates into CPI. Exchange rate pass-through rates are estimated using a 5-year rolling window from 2007 to 2020 at the quarterly level. Appendix C provides additional details on the estimation. The counterfactual trend is computed subtracting the effect due to the rise in the average importing tenure, documented in Figure 3.17 in Appendix C. The effect of importing tenure is computed multiplying tenure by its effect on the pass-through into border prices (Table 4) and scaled by its contribution to domestic price sensitivity (Table 3.13).

ble 1.2) and the fact that omitting importing tenure increases CPI sensitivity by approximately 20%.[51] Figure 1.13 shows that the counterfactual trend in CPI sensitivity (solid line) decreases 40% less relative to the estimated one.[52] This confirms the importance of micro-level determinants of border price pass-through and their evolution in explaining aggregate dynamics like the trend in domestic price sensitivity to exchange rates.[53]

[51]A tenure of 18 quarters implies a pass-through rate into import price 0.10 lower ($\log(18) \times 0.035$), given the estimates in Table 1.2. I then multiply it by 20% to get the effect on domestic prices, which is approximately 0.025%. CPI sensitivity declines from 0.117% to 0.055%. Omitting the role of tenure, the end point is 0.0755%, approximately 35% higher.

[52]As robustness, Figure 3.18 in Appendix C shows the counterfactual trends using different measures of tenure and different estimates for the marginal effect of tenure on border price pass-through rate. The counterfactual trend decreases at least 20% less than the estimated one.

[53]The rise in average importing tenure and, more generally, international market participation in the period starting from 2009 could be driven by the formation of new international relationship following the Great Trade collapse in 2008, Heise (2019). Expanding the analysis to include the years of the Great Recession and/or around Covid with a focus on business-cycle dynamics is an interesting avenue for future work (Di Giovanni et al., 2022,

1.4.7 *Taking Stock and Policy Implications*

My empirical analysis establishes a number of important facts. Taken together, these results show that accurately accounting for the role of domestic frictions is key to understand both the insensitivity and the sensitivity of domestic prices to exchange rate fluctuations. Moreover, heterogeneity in friction and import exposure is essential to determine which sectors are the most important contributors to the (in)sensitivity of domestic prices. I now elaborate on the broad policy implications of the results presented as domestic price sensitivity to exchange rates is key for the transmission of international shocks, monetary policy and domestic redistribution dynamics.

One fundamental aspect for monetary policy trade-offs in open economy is which inflation rate is relevant to policymakers, that, in turn, depends on the exchange rate pass-through (ERPT). On one hand, ERPT is related to inflation stabilization in open economy, exchange rate misalignment and the so-called "fear of floating" (Calvo and Reinhart, 2002). Incomplete ERPT partially insulates domestic prices to exchange rate fluctuations, reducing the cost of floating and volatile exchange rates. On the other hand, ERPT is also related to the transmission and the absorption of shocks, and terms of trade imbalances. Incomplete ERPT limits expenditure switching forces, trade and capital adjustments, reducing the effectiveness of ex-

Antràs, 2020).

change rates as shock absorber and policy instrument. The inflation rate central banks should target crucially depends on the degree of exchange rate pass-through into domestic prices: PPI (CPI) targeting is optimal in case of low (high) pass-through rates of exchange rate fluctuations (Corsetti et al., 2010, Chen, Devereux, Shi and Xu, 2022). Abstracting away from domestic friction implies a substantially higher sensitivity of domestic prices and, thus, a potentially different optimal monetary policy target in open economy.

Moreover, exchange rate fluctuations are transmitted heterogeneously to domestic products. Policymakers should weight different components of domestic inflation depending on their frictions and exposure, not necessarily coinciding with CPI weights, resembling the closed-economy long-standing inflation targeting debate (Bernanke and Woodford, 2005). The heterogeneous sensitivity is also relevant for the transmission of international shocks and domestic redistribution dynamics: domestic households and firms might be differentially exposed to exchange rate changes depending on the consumption and input mixes used (Cravino and Levchenko, 2017*a*, Jaravel, 2021).

Lastly, the role of importers' characteristics showcases that micro-level determinants of heterogeneous incomplete pass-through into border prices matter for aggregate dynamics and long-run trends. These findings suggest that, in a globalized and interdependent economy, it is important to learn about micro-level forces that influence the transmission of shocks across borders and how they interact with aggregate dynamics and policy conduct (Di Giovanni and Levchenko, 2010, Heise et al., 2022).

1.5 Conclusion

In this paper, I have explored the role of domestic frictions for the (in)sensitivity of domestic prices to exchange rate fluctuations. I find that domestic frictions such as distribution costs, variable markups and nominal rigidities account for 60% of the overall insensitivity of domestic CPI, relatively more than incomplete pass-through into border prices. The presence of domestic frictions impacts also the channels of domestic price sensitivity: contrary to previous literature, most of the sensitivity arises from direct exposure (imported final consumption) because domestic frictions dampens relatively more the response of domestically produced goods (indirect exposure).

The extensive use of micro-level data allows to quantifies a rich heterogeneity in sensitivity across products, originating from the interaction of heterogeneous domestic frictions, direct and indirect exposures and incomplete pass-through rates. Importantly, the identity of the products contributing the most to the transmission of exchange rate fluctuations depends on the subset of heterogeneity considered. This testifies the importance of jointly accounting for the frictions and mechanisms included in the analysis. In this regard, the model and the calibration strategy used can guide future research on the relationship between domestic prices and exchange rates.

Chapter 2.

Strategic Behavior and Exchange Rate Dynamics

with Luigi Pollio

The huge trading volumes in the foreign exchange rate markets are highly concentrated among few financial players. We develop a monetary model of exchange rate determination featuring heterogeneous investors with different degrees of price impact. We show that the presence of price impact leads to the amplification (dampening) of non-fundamentals (fundamental) trade on the exchange rate, reducing its informativeness. Thus, investors' price impact provides a rationale for the exchange rate disconnect and the excess volatility puzzles. Further, we provide empirical evidence in line with our theoretical predictions using trading volume concentration data from the US FX market for 18 currencies from 2005 to 2019. Finally, we extend our framework to accommodate for another dimension of heterogeneity across investors, information dispersion: we show that 25% of the disconnect and 60% of the excess volatility due to investors' heterogeneity is due to heterogeneity in price impact.

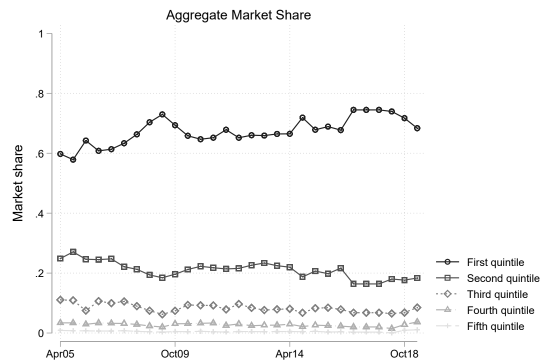
2.1 Introduction

Two of the most established puzzles in international economics are the poor explanatory power of macroeconomic fundamentals in explaining exchange rate fluctuations (Exchange rate determination puzzle) and the excess volatility of exchange rates relative to fundamentals (excess volatility puzzle) (Meese and Rogoff, 1983, Obstfeld and Rogoff, 2000).[54] Recent evidence from the microstructure approach to exchange rates suggests that investor heterogeneity is key to understand exchange rate dynamics and determination. In particular, Lyons et al. (2001) shows that exchange rate behavior is related to order flow, which in turn is associated to investors heterogeneity. Similarly, Bacchetta and Van Wincoop (2006) show that both puzzles can be explained by the presence of information heterogeneity resulting in rational confusion.

This paper examines the impact on exchange rate behavior of the presence of heterogeneous investors with different degrees of price impact. The huge trading volume in the currency markets, about $6 trillions per day, is highly concentrated among the market-making desks of few large financial institutions. Figure 2.1 shows that the market share of the top quintile of financial institutions in the foreign exchange rate market in New York ac-

[54]Meese and Rogoff (1983) finds that macroeconomic models have a lover predictive power than a random walk. Similarly, Obstfeld and Rogoff (2000) shows that exchange rates fluctuate much more than information about their fundamentals.

FIGURE 2.1 – Market Concentration – NY OTC Foreign Exchange Market



Notes: The figure shows the investors' market share in the New York OTC foreign exchange market by quintile. Market share are computed in terms of total transactions across all currencies. Data are from the NY Fed Biannual FXC report, from 2005 to 2020. Appendix D provides additional information on the data used.

counts for around 70% of the overall turnover. [55] Models of exchange rate determination assume that investors take the equilibrium price as given, abstracting away from the presence of few large investors that recognize the price impact of their decisions and can exert pressure on market prices.[56]

We embed heterogeneity in price impact into a two-country, dynamic monetary model of exchange rate determination. Investors face an international portfolio choice model with noise shocks. We depart from the standard assumption of price-taking investors assuming the presence of a continuum of investors that differ in their degree of price impact. A fraction of

[55] In 2019, the average daily global volume the foreign exchange market was about $6.6 trillion. 80% of all transactions took place in the six major markets, UK, USA, Japan, Singapore, Switzerland and Hong Kong. Within each market, 75% of tradings is concentrated in the hand of four-five financial players. Source: BIS Triennial Survey of Foreign Exchange Markets, 2019; NY Fed FX report, 2019.

[56] Anecdotal evidence of manipulation in the exchange rate market also support the assumption of non-zero price impact. In June 2013, Bloomberg News trilled that "*traders at some of the world's biggest banks colluded to manipulate the benchmark foreign-exchange rates used to set the value of trillions of dollars of investments in Pensions Funds and money managers globally*". After extensive investigations, banks pleaded guilty and paid more than $10 billion in fines.

investors are atomistic and competitive, taking prices as given. The complementary fraction is populated by a finite number of strategic investors with non-zero mass acting oligopolistically and internalizing the effects of their trading decision on equilibrium prices.

Our theory of exchange rate determination with heterogeneity in price impact makes market structure a key determinant of exchange rate dynamics. The exchange rate is determined as a weighted average of fundamental (interest rate differential) and noise components. Strategic investors recognize their price impact, which makes them trade less on any information. Thus, the presence of strategic investors leads to the amplification (dampening) of the impact of noise (fundamental) shocks on the exchange rate. By reducing the information loading factor of the exchange rate (price informativeness), i.e. the information content of exchange rate about fundamentals, strategic investors explains the weak explanatory power of macroeconomics variables in predicting exchange rates. Moreover, since fundamentals are less volatile than noise shocks, strategic behavior rationalizes the excess volatility of the exchange rate relative to fundamental by increasing the relevance of the noise component in the dynamics of exchange rate.

We use a panel of 18 currencies from 2005 to 2019 to provide evidence supporting the key predictions of our model. We combine daily exchange rate data together with currency-level concentration data from the New York FED FXC biannual reports. Consistently with our theoretical framework, we show that currencies that are exchanged in more concentrated markets are more disconnected from fundamentals and are more volatility.[57]

Lastly, we quantify the impact of strategic behavior for the dynamics of the exchange rate and compare its contribution to a competing dimension of investors' heterogeneity, information heterogeneity (Bacchetta and Van Wincoop, 2006). We extend our theoretical framework to include dispersed information in the spirit of Nimark (2017) and Bacchetta and Van Wincoop (2006). Information heterogeneity also provides a rationale for both puzzles: since investors do not know whether changes in the exchange rate are driven by noise or fundamental shocks (rational confusion), they always revise their expectations, amplifying (dampening) the effects of noise (fundamental) shocks. Using analysts' expectations on future exchange rates from the ECB Professional Forecasters survey from 2002 to 2020 to calibrate information dispersion, we solve the dynamic infinite regress problem following Nimark (2008) and show that exchange rates are, on average, 30% more volatile and 40% more disconnected relative to an economy populated only by homogeneous investors (price-taking investors with full information). Moreover, approximately 25% of the extra

[57] We also show that traded volumes are lower for more concentrated currencies, in line with the fact that strategic investors invest less because they internalize the effects of their trading. Moreover, the presence of strategic investors has implications for the excess predictability puzzle: since strategic investors invest less, a larger risk premium is required to absorb the supply of foreign assets, generating stronger UIP deviation. This prediction is also confirmed in the data: more concentrated currencies are also those that are more predictable.

disconnect and 60% of the extra volatility can be attributed to heterogeneity in price impact, suggesting that both dimensions of heterogeneity are quantitatively important for the dynamics of the exchange rate.[58]

2.1.1 Related literature

We contribute to the microstructure approach to exchange rates focusing on investors' heterogeneity in the degree of price impact. Lyons et al. (2001) shows that exchange rate behavior is related to order flow, which in turn is associated to investors heterogeneity. Bacchetta and Van Wincoop (2006) focuses on information heterogeneity, showing that the exchange rate determination and the excess volatility puzzles can be explained by the rational confusion originating due to dispersed information. However, despite extensive evidence that foreign exchange rate markets are highly concentrated and atomic, price-taking investors are hardly realistic, the literature has ignored the potential heterogeneity in price impact. In our modeling approach, we follow Kyle (1985) and Kyle (1989), which have not been studied in the context of exchange rate markets.

This paper contributes to the literature that studies exchange rate determination in the presence of frictions. The literature has focused on different forms of frictions: informational frictions (Evans and Lyons, 2002, Bac-

[58]Notice that the two dimensions of heterogeneity interact non-linearly. The effects of information heterogeneity are amplified by the heterogeneity in price impact: strategic investors trade less, reducing the informativeness of the exchange rate and making prices more dispersed for any level of information heterogeneity.

chetta and Van Wincoop, 2006), infrequent portfolio adjustment (Bacchetta and Van Wincoop, 2010, 2019), imperfect and frictional markets (Gabaix and Maggiori, 2015). To the best of our knowledge, our work is the first to focus on this specific feature of the market structure – the presence of strategic investors – for the determination of the exchange rate.

Our theoretical analysis also relates to the vast literature trying to rationalize major puzzles in international economics. We contribute here providing a new rationale based on strategic behavior and price impact for the failure of macroeconomic fundamentals to predict exchange rate and the large volatility of the exchange rate relative to fundamentals (Meese and Rogoff, 1983, Obstfeld and Rogoff, 2000). We also show the interactions between strategic behavior and UIP violations (Fama, 1984). We do not propose a novel explanations for UIP deviations but the presence of strategic investors can explain currency level differences in UIP deviations. In this regard, we use a panel of 18 currencies to show that different market structures can explain cross-currencies differences in exchange rate puzzles and dynamics, which is relatively unexplored.

The remainder of the paper is organized as follow. Section 2 presents the theoretical framework and the basic mechanism of strategic behavior. In section 3, we present the key implications for the dynamics of the exchange rate and the evidence supporting the theoretical predictions. Section 4 extends the basic framework to include information heterogeneity and quantify the relative contribution of each mechanism. Section 5 concludes. Any omitted proofs, derivations and robustness analysis are in the Appendixes.

2.2 A Monetary Model with Strategic Investors

We propose a framework that embeds strategic behavior in the spirit of Kyle (1989) and Kacperczyk et al. (2018) in a standard two-country, discrete time, general equilibrium monetary model of exchange rate determination (Mussa, 1982). To provide the key intuition on the basic mechanism, we first present a version of the model that assumes that agents have rational expectations on the dynamics of the exchange rate. In Section 2.4, we extend the model to incorporate dispersed information following Bacchetta and Van Wincoop (2006) to conduct our quantitative decomposition.

2.2.1 Basic Set-up

There are two economies – Home and Foreign – that produce the same good so that purchasing power parity holds:

$$p_t = p_t^\star + s_t,$$

where s_t is the log nominal exchange rate, p_t ($p_t^\star$) the log price level in the Home (Foreign) country.[59] The exchange rate is defined as the value of the

[59]Variables referring to Foreign are indicated with a star.

foreign currency in term of domestic currency, so that an increase in the exchange rate reflects an appreciation (depreciation) of the foreign (domestic) currency. There are three assets: one-period nominal bonds of both countries with interest rates i_t and $i_t^\star$ and a technology with fixed real return r. The latter is infinitely supplied while bonds are in fixed supply in their respective currency. We follow Bacchetta and Van Wincoop (2010) and assume asymmetric monetary rules across countries. The Home central bank commits to a constant price level $p_t = 0$ so that $i_t = r$ while the monetary policy in Foreign is stochastic, $i_t^\star = -u_t$ where

$$u_t = \rho_u u_{t-1} + \sigma_u \epsilon_t^u \qquad \epsilon_t^u \sim N(0,1) \tag{1}$$

is the Foreign monetary policy structural shock. Thus, the interest rate differential is

$$i_t - i_t^\star = u_t + r,$$

implying that only the Foreign country influences the dynamics of the exchange rate through its monetary policy.[60] In our model, we refer to a shock in the Foreign monetary policy as a fundamental shock.

There is a continuum of investors of mass one. We assume there are overlapping generations of investors that live for two periods and make

[60]Bacchetta and Van Wincoop (2010) specify a simplified Wicksellian rule of the form $i_t^\star = \psi(p_t^\star - \bar{p^\star}) - u_t$ where ψ is set equal to zero, consistent with the low estimates of ψ reported by Engel and West (2005). Bacchetta and Van Wincoop (2010) shows that an exogenous interest rate rule, as in our case, does not compromise the existence of a unique stochastic steady state for the exchange rate.

only one investment decision.[61] Investors in both countries a re born with an exogenous endowment ω and can invest in the nominal bonds and the risk free technology. We assume that Foreign country is infinitesimally small so that the market equilibrium is entirely determined by the market participants located in the Home country. There are two type of investors: strategic (S) and competitive (C). A mass $1-\lambda$ of investors is composed by standard atomistic price-takers agents. The complementary segment of size λ is composed by a finite number N of strategic investors, each with positive mass λ_i ($\sum_i^N \lambda_i = \lambda$). Importantly, strategic investors internalize their effect on asset prices, acting as an oligopoly.

Investor j maximizes mean-variance preferences over next period wealth w^j_{t+1} by allocating their initial endowment (ω) between domestic and foreign bonds:

$$\max_{b^j_t} \quad E^j_t(w^j_{t+1}|\Omega^j_t) - \frac{\rho}{2} Var^j_t(w^j_{t+1}|\Omega^j_t) \tag{2}$$

$$\text{s.t. } w^j_{t+1} = (\omega - b^j_t)i_t + (i^\star_t + s_{t+1} - s_t)b^j_t, \tag{3}$$

where b^j_t defines the foreign bond holdings, ρ the rate of risk aversion and Ω^j_t represents the information set of investor j at time t. i_t and $i^\star_t + s_{t+1} - s_t$ are the log-linearized returns of domestic and foreign bonds, respectively. Notice that, under the monetary policy assumptions and PPP, we have that $p^\star_t = -s_t$ and both returns are expressed in real terms. The

[61] We abstract away from saving decisions by assuming that investors derive utility only from their wealth at the end of life (Bacchetta and Van Wincoop, 2006, 2010).

only difference between the two assets is that the return on foreign bonds is stochastic.[62] We now assume that agents have symmetric rational expectations on the dynamics of the exchange rate ($\Omega_t^j = \Omega_t$) and include dispersed information in Section 2.4.

Investors' demand schedule and portfolio allocation depend on their type. Strategic investors internalize the effects that their demand has on equilibrium prices (more precisely, on the equilibrium exchange rate), while competitive investors do not. Appendix B shows that the optimal demand for foreign bonds by investor j is:

$$b_t^C = \frac{E_t(s_{t+1}) - s_t + i_t^\star - i_t}{\rho\sigma_t^2} \qquad \text{if } j = C \tag{4}$$

$$b_t^S = \frac{E_t(s_{t+1}) - s_t + i_t^\star - i_t}{\rho\sigma_t^2 + \frac{\partial s_t}{\partial b_t^S}} \qquad \text{if } j = S, \tag{5}$$

where investors' demand of foreign bonds positively on the expected excess return, $q_{t+1} \equiv E_t(s_{t+1}) - s_t + i_t^\star - i_t$, negatively on its variance, σ_t^2, and investors' risk aversion. We focus on a stochastic steady state where the variance σ_t^2 is time-invariant. Importantly, strategic behavior, captured by investors' own price impact $\frac{\partial s_t}{\partial b_t^S}$, reduces investors' demand of foreign bonds for every level of excess return.[63] Given a total supply of foreign bond B, the price impact of strategic investor i,

$$\frac{\partial s_t}{\partial b_t^{S,i}} = \frac{\lambda_i \rho\sigma_t^2}{B\rho\sigma_t^2 + (1-\lambda)} > 0 \tag{6}$$

[62] $p_t = 0$ implies $i_t = r$. Similarly, $p_t^\star = -s_t$ implies that the foreign bond return $i_t^\star + s_{t+1} - s_t$ is expressed in real terms as well.

[63] See Appendix B for the derivation of the analytic expression of the price impact.

is positive, increasing in the mass of the investor, λ_i, and decreasing in the fraction of atomistic investors $1 - \lambda$. In the case strategic investors have the same mass, $\lambda_i = \frac{\sum_i \lambda_i}{N} = \frac{\lambda}{N}$, the individual price impact becomes $\frac{1}{N} \frac{\lambda \rho \sigma_t^2}{B\rho\sigma_t^2 + (1-\lambda)}$.[64] The structure of the market determines the magnitude of the price impact, thus the relevance of strategic behavior: the magnitude of the individual price impact depends negatively on the number of strategic traders, N, and positively the size of the strategic segment, λ. Thus, the price impact is larger in more concentrated markets (lower N and/or higher λ).[65]

In addition to the agents described above, we introduce noise traders. As standard in this class of models, their presence allows to match exchange rates moments in the data such as exchange rate volatility, disconnect and UIP deviations (Kyle, 1989, Bacchetta and Van Wincoop, 2006, 2010). We follow Bacchetta and Van Wincoop (2010) and assume that the noise traders' demand for foreign bonds is exogenous and given by:

$$X_t = (\bar{x} + x_t)\bar{W},$$

[64]In what follows, we focus on the case of symmetric strategic investors because comprehensive investor-level market share data are not available. The NY Fed FX Reports, used for our calibration and empirical analysis, provide information on the aggregate market share of each quintile and the number of investors in total. Notice that qualitative predictions are not altered by our assumption.

[65]Notice that, in our international portfolio model, strategic traders have a lower price impact on the equilibrium price of an asset than in a closed-economy version because the supply of assets is subject to valuation effects. Strategic traders account also for variations in the total value of the supply of assets when they internalize the effect that their demand has on the exchange rate. This explains the presence of B, the total supply of foreign assets, at the denominator, reflecting a peculiarity of a portfolio of international bonds. See Appendix B for additional details.

where $\bar{W}$ is the steady state aggregate financial wealth in the Home economy, $\bar{x}$ is a constant and x_t follows an exogenous process:

$$x_t = \rho_x x_{t-1} + \sigma_x \epsilon_t^x \qquad \epsilon_t^x \sim N(0,1)$$

The demand of foreign assets absorbed by noise traders in the stochastic steady state is equal to $\bar{x}\bar{W}$. Any deviation from the steady state driven by x_t is interpreted as a noise shock (orthogonal to the fundamental shock u_t in Equation (1)). Positive shocks to x_t increase the desirability of the foreign assets leading the foreign currency to appreciate without movements in the interest rate differential.

Equilibrium and Basic Mechanism We derive an expression for the equilibrium exchange rate combining investors' demand schedules and the market clearing condition of the foreign bond market:[66]

$$(1-\lambda)b_t^C + \sum_i^N \lambda_i b_t^{S,i} + X_t = Be^{s_t}, \tag{7}$$

where the left hand side is the total demand of foreign bonds from competitive investors, strategic investors and noise traders, and the right hand side is the (constant) supply of foreign bonds denominated B denominated in domestic currency.

[66] The market clearing for the domestic bond is not relevant because the bond is perfectly substitutable with the risk free technology, which is infinitely supplied. Similarly, a monetary model would also require a market clearing condition for the money market. Bacchetta and Van Wincoop (2006) and Bacchetta and Van Wincoop (2010) assume that investors generate a money demand (independent of their portfolio decision) and that money supply accommodates it under the exogenous rule for interest rates. We do not explicitly model a money market in order to limit notation, leaving it in the background.

We define the concept of equilibrium in our model as follow: for an history of shocks $\{\varepsilon_t^x, \varepsilon_t^{\Delta i}\}_{t=0}^{-\infty}$, an equilibrium path is a sequence of quantities $\{b_t^C, \{b_t^{S,i}\}_{i=1}^N\}$ and foreign currency (asset) price $\{s_t\}$ such that investors optimally choose their portfolio and market clearing condition holds.

The model allows to derive an explicit solution for the exchange rate s_t from the market clearing condition in Eq. (7):

$$s_t = \underbrace{(1-\mu)\left(\frac{\bar{x}}{b}-1\right)}_{\text{constant}} + \underbrace{\mu\left(E_t s_{t+1} + i_t^\star - i_t\right)}_{\text{fundamental}} + \underbrace{(1-\mu)\frac{1}{b}x_t}_{\text{noise}}, \tag{8}$$

where $b = \frac{B}{W}$ and $\mu = \frac{1}{1+\Phi(\lambda,N)}$ with $\Phi(\lambda, N) = \frac{B\rho \operatorname{Var}_t(s_{t+1})\left(1+B\rho \operatorname{Var}_t(s_{t+1})-\lambda\frac{N-1}{N}\right)}{\left(1+B\rho \operatorname{Var}_t(s_{t+1})-\lambda\frac{N-1}{N}\right)-\frac{\lambda^2}{N}}$. The exchange rate follows a forward looking auto-regressive process with drift where the constant term depends on a set of parameters and the stochastic component depends on future fundamental and noise shocks. Manipulating Eq. (8) further, it can be shown that the exchange rate s_t can be written as follows:

$$s_t = \mu \sum_{k=0}^{\infty} \mu^k \left(i_{t+k}^\star - i_{t+k}\right) + \frac{1-\mu}{b} \sum_{k=0}^{\infty} \mu^k \left(x_{t+k}\right). \tag{9}$$

The exchange rate is a weighted average of current and future fundamental shocks ($i_{t+k}^\star - i_{t+k}$) and noise shocks (x_{t+k}). The weight, μ, represents the informativeness of the exchange rate and quantifies how much information about fundamental is carried by the exchange rate. Importantly, the informativeness of exchange rate decreases when strategic investors operate in the foreign bond market (higher λ or lower N imply higher Φ and, thus, lower μ). As the market becomes more concentrated (lower N and/or higher λ), demand from optimizing investors declines because of the larger price impact. Thus, noise traders' demand becomes relatively more important in the determination of the exchange rate.[67] [68]

We now calibrate and simulate the basic model to illustrate the mechanism discussed above. We consider 18 exchange rate pairs, all defined against the USD, from 1993 to 2019 at daily frequency.[69] Without loss of generality, we set $\bar{r} = 0$, so that the $i_t - i_t^{\star} = u_t$. The interest rate differential is defined as the difference between the 1-month forward and the spot exchange rate, assuming covered interest rate parity holds. The volatility and the persistence of the fundamental shock, σ_u and ρ_u, are calibrated to match the cross-currency average of the estimated AR(1) parameters. This yields $\sigma_u = 0.012$ and $\rho_u = 0.8$.[70] The perceived variance of the excess return, σ_t^2, is assumed to be time-invarying and homogeneous across investors. We approximate it to the average variance of the one-period exchange rate change, which is $\sigma(\Delta s_{t+1}) = 0.029$ in the data. The parameters controlling the mag-

[67]The key intuition is based on Kyle (1989): when traders recognize that the residual supply curve is upward-sloped, quantities are restricted and also less elastic. Prices are then less informative. The same intuition applies here.

[68]Our price informativeness parameter μ relates to the magnification factor in Bacchetta and Van Wincoop (2006). In their case, information dispersion across investors reduces the information content of exchange rates by amplifying the effects of noise traders. As in their paper, the behavior of the informativeness index is key for the amplification mechanism analyzed here.

[69]The currencies are: Euro, Japanese Yen, Argentinian Peso, Brazilian Real, Canadian Dollar, Swiss Franc, Australian Dollar, Chilean Peso, Indian Rupee, Mexican Peso, British Pound, South African Rand, Russian Ruble, Swedish Krona, Turkish Lira, New Zealand Dollar, Singapore Dollar, Norwegian Krone.

[70]We use daily data averaged at monthly frequency.

nitude of the strategic behavior are λ and N. We use the NY Fed Biannual FCX Reports from 2005 to 2019 to calibrate $\lambda = 0.7$ to match the market share of the top quintile of investors in the NY FX market and $N = 4$ as the number of investors in the top quintile. The process of noise demand x_t, which cannot be observed, is calibrated to match exchange rate dynamics: the persistence of the noise shock, ρ_x, is set high enough such that the exchange rate behavior is sufficiently close to a random walk; the volatility of the process is chosen to match the volatility of the one-period change in exchange rate, $\sigma(\Delta s_{t+1})$. Given the benchmark values for λ and N, σ_x is set equal to 0.0998.[71] We set b, the inverse home bias measure, equal to 0.33, meaning that foreign assets account for one third of the total domestic financial wealth. This is an approximate average obtained from the IMF IIPS dataset as in Bacchetta and Van Wincoop (2019).[72] Lastly, the rate of relative risk aversion ρ is set to 50 following Bacchetta and Van Wincoop (2019).[73] The parametrization, summarized in Table 2.1, uses values in line with previous literature.[74]

The main implication of the presence of strategic investors is that the

[71] $\bar{x}$ is calibrated such that the value of the exchange rate in the stochastic steady state is zero. Our choice excludes any trend in the dynamics of exchange rate but does not affect the results of our model.

[72] For simplicity, the supply of foreign assets, B, is normalized to one. To consistently close the model, we set ω, the initial endowment of each investor, equal to 3. This comes from the fact that $b = \frac{B}{\bar{W}}$. Calibrating b and normalizing B mean that $\bar{W} = 3$. Total financial wealth in equilibrium is equal to the initial endowment.

[73] In the model, risk aversion is the only source of currency premia, which would be very small for standard rates of risk aversion. Our results are nevertheless qualitatively robust to different values of different risk aversion coefficients. Moreover, notice that ρ and B are multiplicative in the model, and $\rho = 50$ could be different if B was normalized differently.

[74] See Appendix A for additional details on data.

TABLE 2.1 – Benchmark Parametrization

	Value	Moment - Target
λ	0.7	Share transactions top investors – (1st quintile) in NYFXC
N	4	Number of top investors – (1st quintile) in NYFXC
ρ_u	0.8	Average persistence AR(1) Δi_t
σ_u	0.012	Average StD innovation AR(1) Δi_t
$\sigma(\Delta s_t)$	0.04	(Average) StD FX change
σ_x	0.0998	$\sigma(\Delta s_t)$
ρ_x	0.9	FX Random Walk/Average R^2
b	0.33	Home Bias
ρ	50	

Notes: The table summarizes the parametrization used in the basic framework. For each parameters, we report the value used in the model, the corresponding moment and data used to calibrate, and, if applicable, the target moment used to estimate it. Appendix D provides additional information on the data used.

response of the exchange rate to noise (fundamental) shock is amplified (dampened) relative to an exchange rate market without strategic investors ($\lambda = 0$ or $N \to \infty$, labeled "Competitive").[75]

Figure 2.2 plots the impulse response functions to one standard deviation shock in fundamental and noise shock in the presence of strategic investors and abstracting away from them ("Competitive" markets). The bottom row considers a positive noise shock x_t, which can be interpreted either as a positive demand shock or a negative supply shock of foreign assets. Either way, the residual demand of foreign assets decreases, increasing the price of the foreign assets and the exchange rate without any change in fundamentals. The excess return falls below the steady state because the

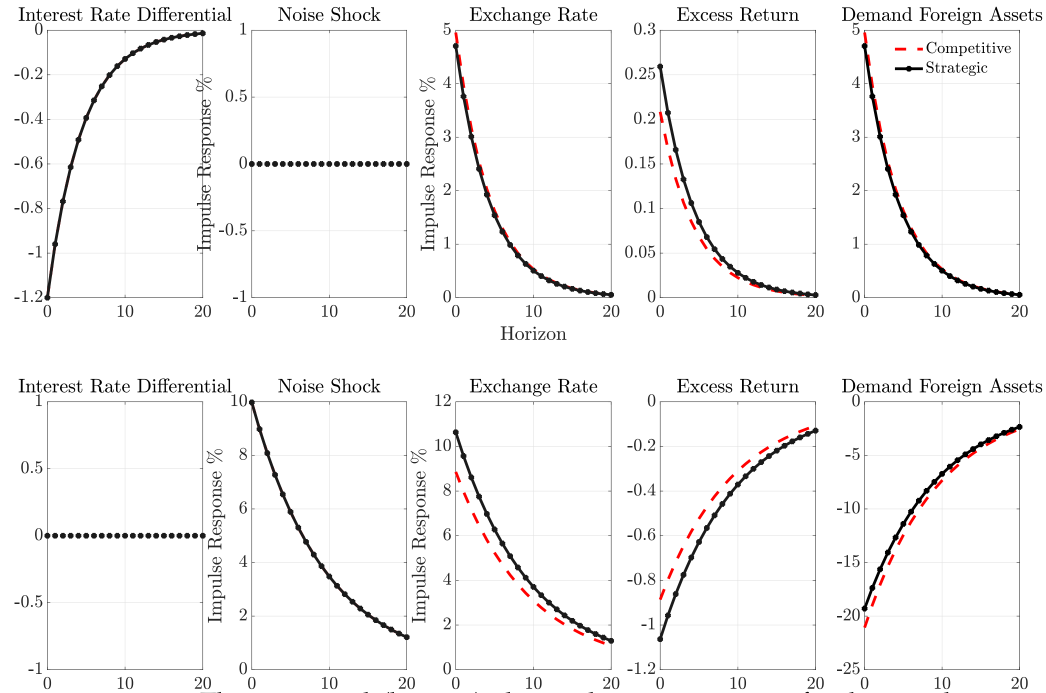
[75]We show in Appendix B that the result is independent of the parameterization of the model.

exchange rate increases. Lower excess return pushes investors to purchase less foreign assets, re-balancing in favor of domestic ones. However, in a world where investors are strategic (solid), the demand of foreign assets decreases less than in a competitive market (dashed) as investors internalize the negative impact of their tradings on the price. This makes total demand less sensitive and amplifies the effect of the noise shock on the exchange rate.[76] A smaller decline in investors' demand exerts additional upward pressure on the price of the foreign bonds and on the exchange rate. Given the lower sensitivity of the demand of foreign bond, the response of the excess return is dampened relative to a competitive market in order for the market to clear.

The top row of Figure 2.2 considers the response to a shock in fundamental u_t. A contraction in monetary policy in the foreign country leads to a drop in the interest differential, increasing the excess return and, thus, the investors' demand of foreign assets. This results in the appreciation of the foreign currency. In a world in which investors are strategic (solid), investors increase their holdings of foreign assets relatively less as their price impact makes their demand less sensitive. As a consequence, the price of foreign assets increases relatively less than in a competitive market, damp-

[76]The dynamics of total demand are the results of the compositional forces. Both competitive and strategic investors' demands, b_t^C and $b_t^{S,i}$, drop when the excess return falls. However, conditional to the same change in excess return, the reaction of $b_t^{S,i}$ is smaller because of their price impact. The smaller response of total demand for larger concentration is then explained by the fact that, with $\lambda > 0$, more weight is given to the demand of strategic traders.

FIGURE 2.2 – Impulse Response to Exogenous Shocks



Notes: The top panel (bottom) shows the response to a fundamental (noise) shock. The first and the second columns show the dynamics of a one standard deviation shock in fundamental and noise, respectively. The third column shows the dynamics of the exchange rate. Column four shows the response of the realized excess return, defined as $q_{t+1} = s_{t+1} - s_t + i_t^\star - i_t$. The last column shows the response of the total demand of foreign assets, defined as $(1-\lambda)b_t^C + \sum_i^N \lambda_i b_t^{S,i}$, where b_t^C and $b_t^{S,i}$ are defined according to Equations 4 and 5, respectively. The solid black line shows the response in the benchmark parametrization with strategic investors, $\lambda = 0.7$. The red dashed line shows the response in an economy without strategic investors, $\lambda = 0$. Remaining parameters are common across scenarios, see Table 2.1.

ening the effect of fundamental shock on the exchange rate.

2.3 Implications for Exchange Rate Dynamics

We use the calibrated model to illustrate and discuss the implications of strategic behavior for exchange rate dynamics in terms of exchange rate volatility, exchange rate disconnect and UIP deviations. In particular, we show that the presence of strategic investors increases the volatility of the exchange rate and its disconnect from fundamentals increase. Moreover, strategic behavior does not generate excess predictability *per se* but makes UIP deviations larger.

Exchange Rate Disconnect One of most robust empirical evidence on exchange rate dynamics is their disconnect from fundamentals (Meese and Rogoff, 1983). The standard measure used to evaluate the disconnect of exchange rates is the R-squared or any other measure of explanatory power of the following regression:

$$s_{t+1+k} - s_t = \alpha + \beta_k (i_t - i_t^{\star}) + \varepsilon_{t+k+1}, \tag{1}$$

where $i_t - i_t^{\star}$ represents the fundamental driver of the exchange rate change $s_{t+1+k} - s_t$, with k representing the horizon of interest.

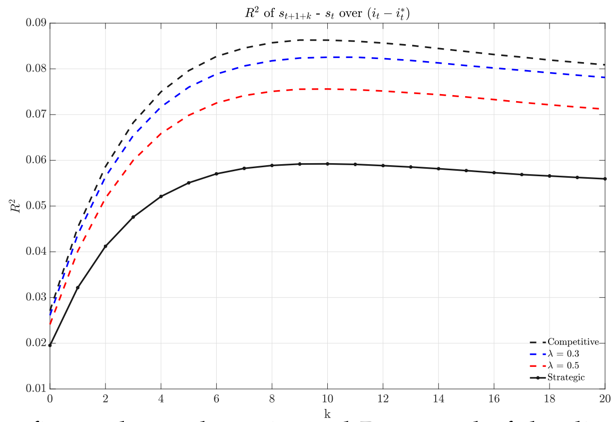
We show how market structure and the presence of strategic behavior helps explaining the poor explanatory power of standard theories of nominal exchange rates. We simulate the model, estimate Equation (1) and study how the explanatory power (R-squared) of the disconnect regression changes when the economy is increasingly populated by strategic investors.[77] Figure 2.3 shows the R-squared at different horizons (k up to 20) and for different degrees of strategic behavior (different levels of λ). In-

[77] We run 3000 simulations and, for each iteration, the model runs for 1000 periods with 4000 burn-in.

dependently of the presence of strategic investors, the model predicts that the puzzle is less acute for long-run exchange rate movements, consistently with the literature (Obstfeld and Rogoff, 2000). A competitive market predicts that current fundamentals explains from 4% to 12% of the fluctuations in the exchange rate in the short-run and in the long-run, respectively. Our benchmark calibration reduces the explanatory power of fundamental by 50% (2% in the short-run and 5% in the long-run) because the presence of strategic investors reduces the informativeness of exchange rates about the underlying fundamental, reducing the explanatory power of Equation (1). This is consistent with the fact that the volatility of the exchange rate explained by noise shocks is higher in the presence of strategic investors relative to a competitive market (left panel of figure 2.4). Under our benchmark parametrization, the variance explained by noise is approximately 84% of the total variance at the impact. The share of the total variance in exchange rate explained by noise traders decreases by 10% (approximately 76%) when abstracting away from strategic behavior.

Exchange Rate Volatility Extensive evidence show that exchange rates are more volatile than fundamentals, the so-called excess volatility puzzle (Obstfeld and Rogoff, 2000). We show how the presence of strategic behavior contributes to the excess volatility of the exchange rate relative to fundamentals by exacerbating the relevance of noise traders.

FIGURE 2.3 – Exchange Rate Disconnect



Notes: The figure shows the estimated R-squared of the disconnect regression in Equation 1 for different horizons (1 to 20 periods) using simulated data. We run 3000 simulations and, for each iteration, the model runs for 1000 periods with 4000 burn-in. Data are simulated in four different scenarios in terms of strategic behavior: solid black line for our benchmark parametrization ($\lambda = 0.7$); dashed red line with $\lambda = 0.5$; blue dashed line with $\lambda = 0.3$; black dashed line for a competitive economy ($\lambda = 0$). Remaining parameters are common across scenarios, see Table 2.1.

Manipulating Equation 9, we can derive an expression of the unconditional variance of the exchange rate as a combination of the variances of both fundamental and noise shocks:

$$\mathrm{Var}(s) = \frac{\mu^2}{(1-\mu\rho_u)^2}\left[\frac{1}{1-\mu^2} + \frac{\rho_u^2}{1-\rho_u^2}\right]\sigma_u^2 + \frac{(1-\mu)^2}{(1-\mu\rho_x)^2 b^2}\left[\frac{1}{1-\mu^2} + \frac{\rho_x^2}{1-\rho_x^2}\right]\sigma_x^2. \tag{2}$$

The right panel of Figure 2.4 shows that unconditional variance of the exchange rate is increasing in the presence of strategic behavior.[78] Relative to the volatility of the fundamental, $\frac{\sigma_u}{\sqrt{1-\rho_u^2}} \approx 0.02$, the exchange rate is an order of magnitude more volatile. Importantly, strategic behavior makes the

[78] Appendix B shows that, theoretically, the effect of strategic behavior is not monotonic. The monotonicity is satisfied given standard parametrizations. On this regard, our calibration is very conservative: higher values of ρ_x, and lower values ρ_u or b would all strengthen presence of monotonicity. Further details are available in Appendix B.

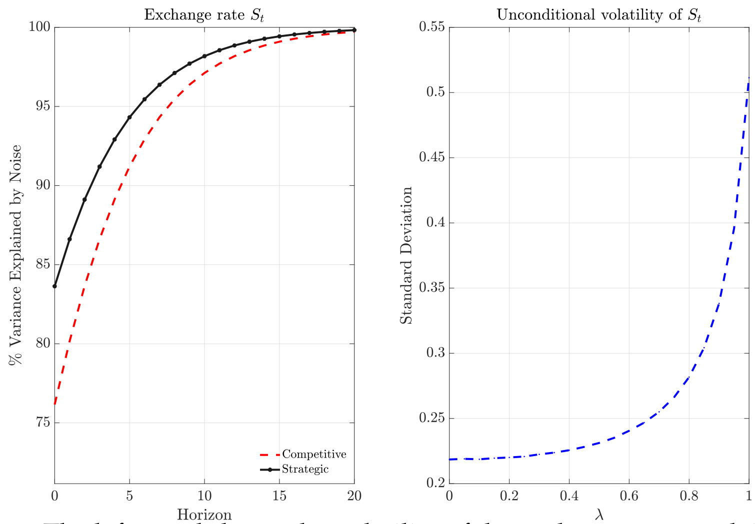
exchange rate 20% more volatility relative to a competitive market.[79] The presence of strategic investors reduces the informativeness of the exchange rate attributing relatively more weight to the noise component, which is relatively more volatile than the fundamental component, contributing to the excess volatility of the exchange rate.

Notice that accounting for the presence of strategic investors in the underlying market structure reduces the implied volatility of noise traders required to match exchange rate dynamics (exchange rate volatility) because strategic investors amplify their effects. Figure 3.24 in Appendix G shows that there exists a negative relationship between the magnitude of strategic behavior (N and λ) and σ_x, for a given target value of $\sigma(\Delta s_{t+1})$. In the case of competitive markets, the volatility of the noise component should be $\sigma_x = 0.12$ in order to match the same volatility of the exchange rate, 20% higher than in our benchmark calibration. This highlights the importance of accounting for the underlying market structure and, in particular, the presence of strategic investors, as noise traders are not as noisy as previously calibrated.

Excess Return Predictability - UIP Another empirically robust evidence is the predictability of excess return, commonly defined as deviations from the Uncovered Interest Parity (UIP), which entails a positive correlation

[79]The standard deviation of the exchange rate is 0.22 in a competitive market – $\lambda = 0$ – and increases to 0.27 in our benchmark calibration.

FIGURE 2.4 – Exchange Rate Volatility



Notes: The left panel shows the volatility of the exchange rate explained by noise shocks at different horizons using simulated data from our model. We run 3000 simulations and, for each iteration, the model runs for 1000 periods with 4000 burn-in. Horizon k goes from 1 to 20. The solid black line shows the volatility explained by noise shocks in the benchmark parametrization with strategic investors, $\lambda = 0.7$. The red dashed line shows the volatility explained by noise shocks in an economy without strategic investors, $\lambda = 0$. Remaining parameters are common across scenarios, see Table 2.1. The right panel shows the estimated unconditional standard deviation of the exchange rate defined according Equation 2 for different levels of strategic behavior using simulated data from our model.

between exchange rate appreciation and interest rates (Fama, 1984). Our model predicts systematic deviations from UIP due to a non-zero net supply of foreign assets, independently of the presence of strategic investors. However, strategic behavior generates larger UIP deviations relative to a competitive market.

Through the lens of our model, the one-period excess return, $q_{t+1} = s_{t+1} - s_t - (i_t - i_t^*)$, can be expressed as follow from Equation (8):

$$E_t q_{t+1} = \frac{\Phi}{B}\left(Be^{s_t} - X_t\right), \tag{3}$$

where the right-hand side represents the deviation from UIP which can interpret as the risk premium required by investors for holding a foreign asset and clear the market. The risk premium depends on two components: the net supply of foreign assets in parenthesis (supply net of noise traders), and the market structure captured by Φ, which is increasing in λ or decreasing in N. Our model predicts that UIP does not hold even in a fully competitive market, when λ is zero. Moreover, UIP deviations are larger as the market is increasingly populated by strategic investors. The total demand of foreign assets becomes more insensitive when investors are strategic. Thus, a larger risk premium is required to absorb the net supply of foreign assets relative to a competitive market, making the excess return more predictable.

We use the calibrated model and simulated data to estimate a standard one-period Fama regression:

$$q_{t+1} = \alpha + \beta(i_t - i_t^\star) + \epsilon_t. \tag{4}$$

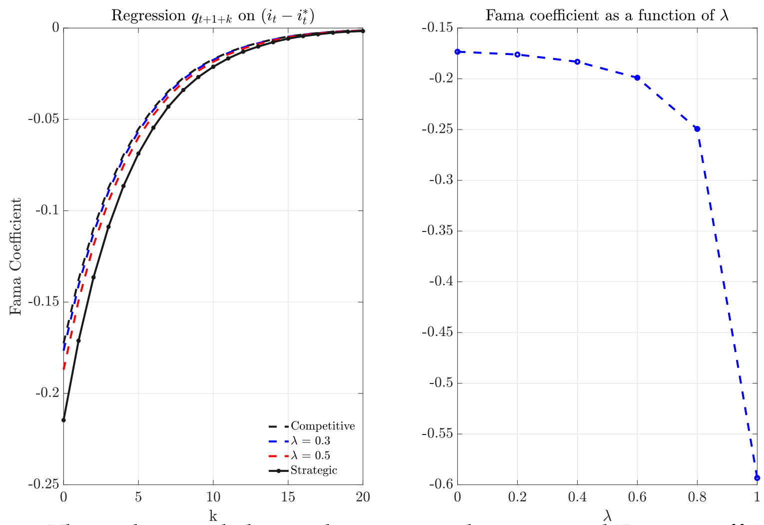
where q_{t+1} is the realized excess return. While UIP implies that the Fama coefficient, β, is zero, empirical evidence typically finds a negative number. Our model predicts that β is given by:[80] [81]

$$\beta = -(1-\mu)\frac{1}{1-\mu\rho_u} < 0,$$

[80]See Appendix E for derivations.

[81]Interestingly, β is equal to zero if the supply of asset is constant when denominated in domestic currency, that is, B is not multiplied by e^{s_t}. In this particular case, the excess return depends only on the noise component X_t, which is orthogonal to fundamental shocks. Therefore, β is equal to zero even if there are systematic deviations in UIP. In other words, risk premium is still positive (UIP does not hold) but it is not predictable ($\beta = 0$).

FIGURE 2.5 – Excess Return Rredictability



Notes: The right panel shows the estimated one-period Fama coefficient using Equation 3 and simulated data from our model for different levels of strategic behavior. We run 3000 simulations and, for each iteration, the model runs for 1000 periods with 4000 burn-in. The left panel shows the Fama coefficients at multiple horizons, estimated using simulated data and the following specification $q_{t+k+1} = \alpha + \beta(i_t - i_t^\star) + \epsilon_t$., where $q_{t+k} = s_{t+k+1} - s_{t+k} - (i_{t+k} - i^*_{t+k})$ is the k-period ahead excess return. Horizon k goes from 1 to 20. Data are simulated in four different scenarios in terms of strategic behavior: solid black line for our benchmark parametrization ($\lambda = 0.7$); dashed red line with $\lambda = 0.5$; blue dashed line with $\lambda = 0.3$; black dashed line for a competitive economy ($\lambda = 0$). Remaining parameters are common across scenarios, see Table 2.1.

which is decreasing in the presence of strategic investors. The right panel of Figure 2.5 plots the estimated excess return predictability coefficient β for different levels of λ. As expected, the coefficient is negative and in line with the estimates from the literature, and its magnitude is monotonically increasing in the presence of strategic investors.[82]

[82] Appendix G generalizes the result showing that the k-period ahead Fama coefficient, β_k. The k-period ahead Fama coefficient is estimated regressing the k-period ahead excess return, $q_{t+k} = s_{t+k+1} - s_{t+k} - (i_{t+k} - i^*_{t+k})$, on current interest rate differentials $(i_t - i_t^\star)$. The left panel of Figure 2.5 shows that predictability is monotonically increasing in k and approaching zero for $k = \infty$. On one side, this is not consistent with the predictability reversal puzzle, which refers to the fact that there is a reversal in the sign of expected excess returns at longer horizons (Bacchetta and Van Wincoop, 2010, Engel, 2016). At the same time, it is not surprising considering the absence of any friction, such as infrequent portfolio adjustment (Bacchetta and Van Wincoop, 2010, 2019).

Testing predictions The model delivers four clear testable relationships between exchange rate dynamics and the extent of strategic behavior in financial markets: (i) exchange rate disconnect increases in λ (ii) exchange rate volatility increases in λ; (iii) excess return predictability decreases in λ; ; (iv) trading volume decreases in λ.[83] We use the disaggregate, currency-level, information provided by the New York Fed Biannual FXC Report to test the implications delivered by our theory.

We use the share of total transactions in the FX intermediated by the top first quintile of investors reported by the NY Fed FXC as our measure of strategic behavior (λ) in the exchange rate market. We consider the same of currencies used to calibrate the model.[84] We consider a shorter period of time, from 2005 to 2019, as the FXC reports are available only since April 2005. The FXC report is published in April and October of each year and contains information relative to those months about aggregate turnover, transactions and concentration. We match these data with measures of exchange rate volatility, excess return predictability and exchange rate disconnect. We use data on exchange rate at daily frequency to compute the standard deviation of each currency in a 8-week window around April and October of each year.[85] Similarly, using data on interest rate differentials,

[83] In our framework, the demand of strategic investors is lower due to the internalization of their price impact. Thus, total volume decreases as markets are increasingly populated by strategic investors, in line with standard market microstructure arguments (Foucault et al., 2013).

[84] To check the robustness of our results, we exclude Euro, British Pound and Japanese Yen as the small open economy assumption could not hold for those countries. Results are qualitatively robust.

[85] As robustness, we construct different measures of volatility considering also 3- and

we compute exchange rate disconnect, R^2, and Fama coeffcients, β, in a 6-month windows around April and October of each year.[86]

Table 2.2 provides empirical evidence consistent with the prediction of our theoretical framework. Columns (1) and (2) document a strong positive and statistically significant relationship between our measure of strategic behavior in the financial markets and exchange rate volatility. Columns (3) and (4) show that, as markets become increasingly populated by strategic investors, currencies become more disconnected to fundamentals and more predictable, respectively. In the case of excess return predictability, although the coefficient of interest is not statistically significant, the estimated relationship is quantitatively close to the one implied by the model simulated data (Figure 2.5), reassuring on the relevance of our mechanism. Lastly, Column (5) supports the prediction that the presence of more strategic investors reduces volumes traded in the market. We estimate the effect of strategic behavior on total trading volume introducing currency and year fixed effect, and also controlling for the number of total transaction executed in the market for each currency pair as a measure of market deepness and liquidity in order to mitigate endogeneity concerns.

5-months windows. Results are qualitatively the same.

[86] Using daily data, each β and R^2 coefficient is estimated using around 120 observations. βs and R^2s are not affected if we use different windows.

TABLE 2.2 – Testable Predictions

	StDev		R^2	β^{UIP}	Total Volume
	(1)	(2)	(3)	(4)	(5)
Strategic Investors (λ)	0.035	0.023	-0.275	-1.574	-20201.9
	(0.006)***	(0.009)**	(0.099)***	(2.341)	(11575.90)*
	[0.013]**	[0.011]**	[0.045]***	[2.733]	[9038.05]**
Transactions					1.266
					(0.06)***
					[0.08]***
Year FE	No	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	No	Yes	Yes
N	325	325	325	325	325

Notes: The table reports the relationship between λ and the variables of interest. λ measures the share of transactions intermediated by the top quintile of investors operating in the New York FX Market. Variable of interest are: exchange rate standard deviation (Columns (1) and (2)); R-squared (Column (2)); Fama coefficient (Column (3)); Share of volume traded (Column (4)). λ is measured in April and October of each year from 2005 to 2019 using New York FX Market Report. The standard deviation is computed using daily exchange rates in a 6-month window around April and October of each year. The Fama coefficient and the disconnect R-squared are estimated using Equations 3 and (1), respectively, using daily exchange rates and interest rate differentials in a 6-month window around April and October of each year. The share of volume traded for each currency is computed using data from the New York FX Market Report in April and October of each year. Column (4) controls also for the number of transaction executed by the top quintile of investors, measured using data from the New York FX Market Report in April and October of each year. Standard errors in parenthesis are clustered at the currency level. We report also the Driscoll-Kraay standard errors in squared brackets. Significance level:* $p<0.10$, ** $p<0.05$, *** $p<0.01$ Currencies considered are: Argentinian Peso, Brazilian Real, Canadian Dollar, Swiss Franc, Australian Dollar, Chilean Peso, Indian Rupee, Mexican Peso, South African Rand, Russian Ruble, Swedish Krona, Turkish Lira, New Zealand Dollar, Singapore Dollar, Norwegian Krone. Appendix D provides additional information on the data used.

2.4 Strategic Behavior vs Dispersed Information: A Quantitative Assessment

The recent microstructure approach to exchange rate shows that investors' heterogeneity is key to understand exchange rate determination. One of the most commonly studied dimension of heterogeneity is dispersed information (Bacchetta and Van Wincoop, 2006, Evans and Lyons, 2002). We extend the basic framework in Section 2.2 by relaxing the symmetric rational expectation assumption and introducing information heterogeneity following Nimark (2017) in order to quantify the relative importance of strategic behavior and information heterogeneity for the dynamics of exchange rates.

2.4.1 Relaxing the Rational Expectation Assumption

The model contains all standard elements of an exchange rate monetary model together with the strategic behavior presented in 2.2. Differently from the basic framework, we assume that investors have imperfect knowledge of the shocks that hit the economy, resulting in dispersed information. The remaining structure of the economy remains the same.

The main implication is that investor j's optimal demand of foreign bonds at time t now depends on their individual information set $\Omega_t(j)$:

$$b_t^C = \frac{E_t(s_{t+1}|\Omega_t(j)) - s_t + i_t^\star - i_t}{\rho\sigma_t^2} \qquad \text{if } j = C \tag{1}$$

$$b_t^S = \frac{E_t(s_{t+1}|\Omega_t(j)) - s_t + i_t^\star - i_t}{\rho\sigma_t^2 + \frac{\partial s_t}{\partial b_t^S}} \qquad \text{if } j = S. \tag{2}$$

where the excess return, $q_{t+1} = E_t(s_{t+1}|\Omega_t(j)) - s_t + i_t^\star - i_t$, and the variance of s_{t+1}, σ_t^2, are now conditional to the time t information set $\Omega_t(j)$. Differently from the basic framework, we assume that σ_t^2 is now endogenous but common to all investors.[87] Notice that the main implication of strategic behavior still holds, as the own price impact reduces strategic investors' demand for any level of excess return.

Information Structure Information structure is inspired by Nimark (2017), a generalization of Singleton (1987) and Bacchetta and Van Wincoop (2006). Investors form expectation about the future price of the foreign bond (exchange rate) by observing their private signal about the fundamental and the history of the exchange rate. Formally, investors' information set is given by:

$$\Omega_t(j) = \{f_{t-T}(j), s_{t-T} : T \geq 0\},$$

where

$$f_t(j) = \Delta i_t + \eta_t(j) \text{ where } \eta_t(j) \sim N\left(0, \sigma_\eta^2\right)$$

represents the private signal about fundamentals. Thus, investors have imperfect knowledge about the history of shocks that hit the economy because they observe an unbias signal $f_t(j)$ about Δi_t with an idiosyncratic measurement error $\eta_t(j)$. Investors do not observe perfectly the path of the foreign interest rate, and are not able to back out the fundamental component from observing the exchange rate due to the unobserved transitory noise shock x_t (Admati, 1985). The private signal, $\eta_t(j)$, implies that investors have different expectations about foreign Central Bank's operating procedures and,

[87]By imposing that the conditional variance of s_{t+1} is common among investors, we implicitly assume that investors have the same capacity to process information.

therefore, the need to 'forecast the forecasts of others' (infinite regress problem) arises because of information dispersion.[88]

Equilibrium and Mechanism We extend the definition of equilibrium of the basic framework in Section 2.2 to accommodate the presence of dispersed information: for an history of shocks $\{\varepsilon_t^x\}_{t=0}^{-\infty}$ and signals about fundamentals $\{f_t(j)\}_{t=0}^{-\infty}$, an equilibrium path is a sequence of quantities $\{b_t(j)\}$ and foreign currency (asset) price $\{s_t\}$ such that investors optimally choose their portfolio and market clearing condition holds.

Combining the market clearing condition with investors' demand schedules, we can derive the following expression for the exchange rate:

$$s_t = (1-\mu)\left(\frac{\bar{x}}{b} - 1\right) + \mu\left(\int E\left[s_{t+1} \mid \Omega_t(j)\right] dj\right) - \mu\left(i_t - i_t^{\star}\right) + (1-\mu)\frac{1}{b}x_t, \tag{3}$$

where μ and Φ are defined as in the basic framework, with the former decreasing in the presence of strategic investors (decreasing in λ and increasing in N).[89] However, a closed form solution for the exchange rate is not available when information is dispersed as it depends on the higher-order

[88]The key distinction with Singleton (1987) and Bacchetta and Van Wincoop (2006) is that private signals are not short-lived, i.e. innovations to the fundamental process are not perfectly and publicly observed after a finite number of periods. Short-lived private information allows to derive a finite dimensional state representation, overcoming the infinite regress problem. The solution method proposed by Nimark (2017)and used here allows to solve our model while relaxing Singleton (1987)'s assumption.

[89]The same intuition on the effect of strategic behavior of the basic framework applies here.

expectations about the fundamental:

$$s_t = \mu \sum_{k=0}^{\infty} \mu^k \left[i_{t+k} - i^{\star}_{t+k}\right]^{(k)}_t + \frac{1-\mu}{b} x_t \tag{4}$$

where $\left[i_{t+k} - i^{\star}_{t+k}\right]^{(k)}_t$ denotes the average expectation in period t of the average expectation in period t+1, and so on, of the average expectation in period t+k-1 of k period ahead fundamentals, that is, $\left[i_{t+k} - i^{\star}_{t+k}\right]^{(k)}_t = \underbrace{\int \mathbb{E}_t \ldots \left[\int \mathbb{E}_{t+k-1}}_{k} \left(i_{t+k} - i^{\star}_{t+k}\right) dj\right] \ldots dj$ for any integer k > 0. Thus, when information is dispersed, the price informativeness parameter, μ, captures the weight that higher order expectations on future fundamentals have on exchange rate dynamics.

We solve the model following Nimark (2008) and Nimark (2017), defining the following state-space representation of the endogenous variable s_t:

$$s_t = \mathbf{v}_0 X^{(0:k)}_{t|t},$$

where $X^{(0:k)}_{t|t}$ is the vector of the average expectations on the exogenous state variables of any order from zero through k. The vector of average expectations about the exogenous state variables $X^{(0:k)}_{t|t}$ is assumed to follow a VAR model of order one:

$$X^{(0:k)}_{t|t} = \mathbf{M} X^{(0:k)}_{t-1|t-1} + \mathbf{N}\varepsilon_t.$$

We solve the model iterating over the dynamics of higher-order beliefs. Details are reported in the Appendix F.[90]

[90]There exist other approaches that rely on the fact that average first-order expectations about the endogenous variables can be computed given the guessed laws of motion of

TABLE 2.3 – Expectation Dispersion

	Across all Horizons	Quarter t
Average Dispersion (StD) $\times 10^2$	4.62	2.90
Median Dispersion (StD) $\times 10^2$	4.26	2.48
Average # of Forecasters	47.15	47.61
# of Quarters	337	76

Notes: The table reports the standard deviation in the expectations of future EUR/USD exchange rate across forecasters, averaged across time. Every quarter, forecasters are asked their expectation on the EUR/USD exchange rate one to four quarters ahead. We compute the dispersion across forecaster for every quarter-horizon pair. The first column reports the average dispersion across all quarter-horizon pairs. The second column uses only the one-quarter horizon forecasts. Data are from the ECB Professional Forecasters survey, 2002Q1 to 2020Q4.

We extend the parametrization of the basic framework in Table 2.1 to accommodate for the presence of dispersed information, which requires the calibration of the volatility of private signal, $\sigma_e ta$. We leverage data on exchange rate expectations from the ECB Professional Forecasters survey. The data contain information on the expected euro-dollar exchange rate at different horizons for about 45 professional forecasters at quarterly frequency from 2002 to 2020. We use a Simulated Method of Moments with 150 periods for 100 repetitions and calibrate $\sigma_e ta$ to match the average dispersion (standard deviation) of the same-quarter exchange rate expectation across forecasters reported in Table 2.3.[91] Table 3.17 in Appendix G summarizes the parametrization.

We use the calibrated model to show that dispersed information alone and in conjunction with strategic behavior produces qualitatively similar

the endogenous variables by using the assumption of rational expectations. We find the approach in Nimark (2017) more reliable and fast to implement.

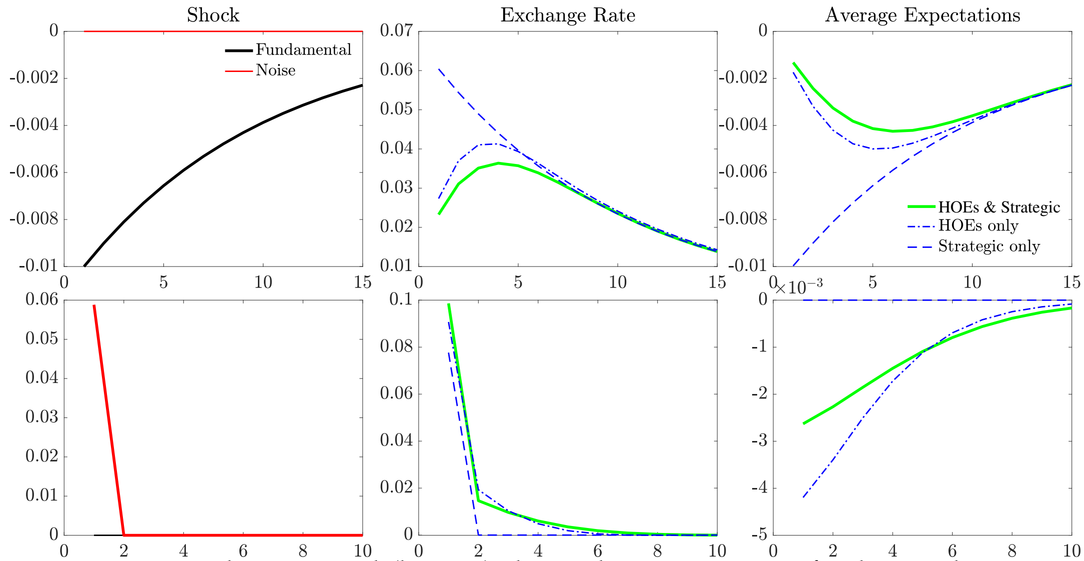
[91] Table 3.16 in Appendix D provides additional moments on expectation dispersion.

implications for exchange rate dynamics (response of exchange rates to exogenous shocks, exchange rate disconnect and volatility), in line with previous literature Bacchetta and Van Wincoop (2006).

Figure 2.6 shows that dispersed information and the resulting rational confusion also amplify (dampen) the effects of noise (fundamental) shocks on the exchange rate, as the presence of strategic behavior in our basic framework. Investors always revise their expectations when the exchange rate moves because they do not whether changes in the exchange rate are driven by noise or fundamental shocks. The top panel shows that, following a negative fundamental shock, investors' expectation do not fully react as part of the response of exchange rates is attributed to the noise component. As a result, the response of exchange rate to a fundamental shock is dampened. Similarly, the response to a positive noise shock (bottom panel) is amplified because the upward movements in the exchange rate is mistakenly confused with a negative change in fundamentals, which puts additional upward pressure on the exchange rate.[92] Figure 3.25 in Appendix G shows that the response of the exchange rate in a model that combines dispersed information and strategic behavior preserves is not the sum of the individual mechanisms but entails a non-linear interaction between the two. The key idea is that strategic behavior makes prices more dispersed for

[92]Notice that we are considering a noise shock without any persistence. Results do not change qualitatively if noise shock exhibits persistence. Moreover, the model generates endogenous persistence as rational confusion takes time to resolve, i.e. average and higher order expectations slowly converge to the full information rational expectation benchmark over time.

FIGURE 2.6 – Impulse Response to Exogenous Shocks



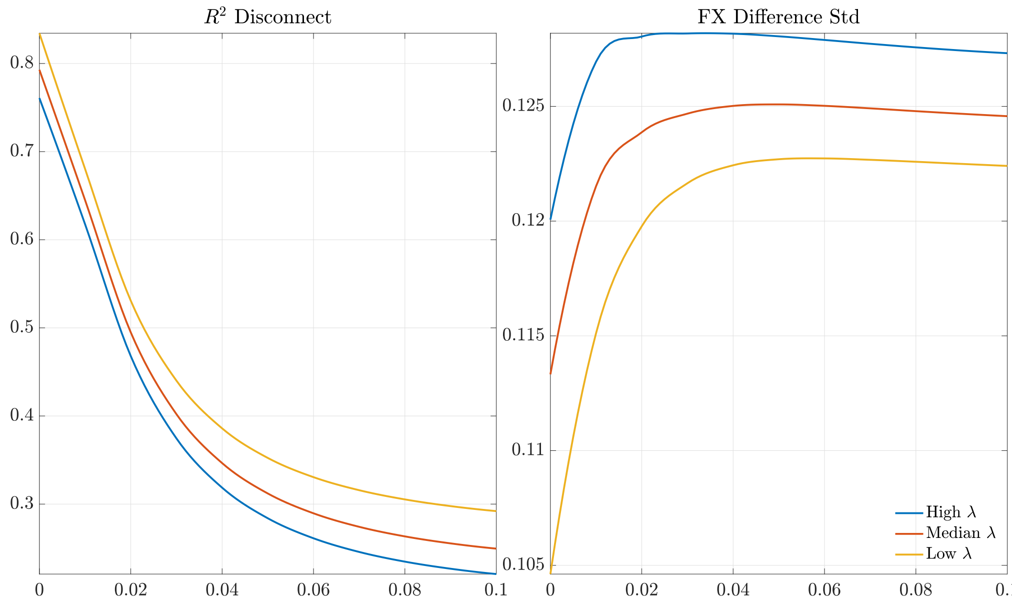
Notes: The top panel (bottom) shows the response to a fundamental (noise) shock. The first column show the dynamics of a one standard deviation shock in fundamental (black line) and noise (red line). The second column shows the dynamics of the exchange rate. Column three shows the response of the average first order ($k = 1$) expectation of future exchange rate defined in Equation (4). The solid green line shows the response in the benchmark parametrization with strategic investors, $\lambda = 0.6$, and dispersed information, $\sigma_\eta > 0$. The blue dashed line shows the response in an economy with strategic investors only, $\lambda = 0.6$ and $\sigma_\eta = 0$. The blue dash-dotted line shows the response in an economy with only dispersed information, $\lambda = 0$ and $\sigma_\eta > 0$. Remaining parameters are common across scenarios, see Table 3.17 in Appendix G.

any level of signal quality σ_η, effectively reducing the weight that investors put on their signal and further amplifying (dampening) noise (fundamental) shocks.[93]

Figure 2.7 shows that dispersed information increases the volatility of the exchange rate, and reduces the connection between fundamentals and exchange rates.The better the quality of the signal about the fundamental (lower σ_η), the more weight is given to the exchange rate in forming future

[93] See Figure 3.26 in Appendix G for the simulated price dispersion for different level of strategic behavior and signal quality.

FIGURE 2.7 – Exchange Rate Volatility and Disconnect



Notes: The left panel shows the estimated R-squared of the one-period disconnect regression in Equation 1 using simulated data with different level of information dispersion and strategic behavior. Similarly, the right panel computes the volatility of exchange rate changes. We run 3000 simulations and, for each iteration, the model runs for 1000 periods with 4000 burn-in. Data are simulated using three levels of strategic behavior: $\lambda = 0.6$ (blue line - "High"), $\lambda = 0.3$ (red line - "Medium"), and $\lambda = 0$ (yellow line - "Low"). For each level of λ, we simulate data for different level of dispersed information: $\sigma_\eta \in [0, 0.1]$ with 0.01 intervals. Remaining parameters are common across scenarios, see Table 3.17 in Appendix G.

expectations. This reduces the amplification of the noise component, increasing the informativeness of the exchange rate and its connection to fundamentals (left panel), and decreasing exchange rate volatility (right panel). As in the basic framework, the presence of strategic behavior increases both the disconnect and the volatility of exchange rate independently of the precision of the signal.[94]

[94] Notice that for a sufficiently low quality of the signal (high σ_η), the volatility of the exchange rate does not increase anymore. As the quality of the signal deteriorates, less weight is given to the fundamental component. This makes the exchange rate less informative, reducing the amplification of the noise component.

2.4.2 *Quantitative Analysis*

We now assume that the model with strategic behavior and dispersed information represents the real data, and decompose the contribution of both elements to the dynamics of exchange rate. Both strategic behavior and dispersed information increase exchange rate volatility and contribute to the low connection between exchange rates and fundamentals. Given our calibration, we use the model to filter the underlying states and perform three different counterfactuals:[95] i) a benchmark economy without strategic investors and dispersed information ($\lambda = \sigma_\eta = 0$); ii) an economy where investors have dispersed information but are not strategic ($\lambda = 0$ and $\sigma_\eta > 0$); iii) an economy where investors are strategic and share the same information set ($\lambda > 0$ and $\sigma_\eta = 0$). By comparing these counterfactual economies to the fully specified model (real data), we can decompose the contribution of each mechanism to the dynamics of exchange rate.

Table 2.4 show that exchange rates are, on average, 30% more volatile and 40% more disconnected relative to a competitive-full information benchmark, respectively.[96] Approximately 80% of the extra disconnect and 35% of the extra volatility can be attributed to dispersed information. Strategic behavior accounts instead for 25% of the extra disconnect and

[95]See Appendix F for additional details on the filtering algorithm.

[96]Even though the increase in predictive power is small in absolute terms, small increments in predictive power are still quantitatively relevant for carry trade.

TABLE 2.4 – Disconnect and Volatility Decomposition

	RMSE Full Model (Actual Data)	Excess Disconnect (%)	% Share Dispersed Information	% Share Strategic Behavior	Non linearity
Average	0.085	39.8	81.2	24.3	-5.5
	Var(s_t) Full Model (Actual Data)	Excess Volatility (%)	% Share Dispersed Information	% Share Strategic Behavior	Non linearity
Average	0.647	28.9	35.1	60.0	4.9

Notes: The table reports the decomposition of the exchange rate disconnect (top panel) and the exchange rate volatility (bottom panel) from the quantitative model in Section 2.4. Exchange rate disconnect is measured using the Root Mean-Square Error of a standard, one-period disconnect regression (Equation (1)). Exchange rate volatility is measured using the standard deviation of the exchange rate. The first column reports the estimated disconnect and volatility from the full model including dispersed information and strategic behavior. We use the full model to match the data and estimate the underlying states. The second column reports excess disconnect and volatility of the full model relative to a benchmark economy that abstract away from both dispersed information and strategic behavior ($\lambda = 0$ and $\sigma_\eta = 0$). The third and fourth columns report the share of the extra disconnect and volatility due to dispersed information and strategic behavior, respectively. The former (latter) is computed comparing RMSE/volatility in the benchmark economy to the RMSE/volatility from an economy without strategic behavior, $\lambda = 0$ and $\sigma_\eta > 0$ (without dispersed information, $\lambda > 0$ and $\sigma_\eta = 0$). The last column reports the discrepancy due to the non-linear interaction between dispersed information and strategic behavior. Appendix D provides additional information on the data. Appendix F provides additional information on the estimation procedure.

60% of the extra volatility.[97] These results suggests that both mechanisms are quantitatively relevant for the dynamics of exchange rate, supporting the importance of combining the two together.

2.5 Conclusion

The high concentration in the foreign exchange market suggests that investors' price impact may be a key element in understanding exchange rate dynamics. In this paper, we have explored the implication of strategic behavior in a simple monetary model of exchange rate determination.

[97]The non-linear interaction between dispersed information and strategic behavior accounts for about -5% as the two mechanisms reinforce each other.

We have shown that strategic behavior reduces the informativeness of the exchange rate, by amplifying the response to non-fundamental shocks and dampening the response to fundamental ones. Thus, it help rationalizing the weak empirical link between fundamentals and exchange rates, and the excess volatility of exchange rates.

The model is stylized in order to derive basic insights and analytic results. Nevertheless, we provide empirical evidence in support of the theoretical predictions using data from a panel of 18 currencies. We also show the quantitative relevance of strategic behavior in a model that includes another dimension of heterogeneity such as information dispersion.

This paper takes a step forward in introducing microstructure institutions in the analysis of exchange rate dynamics. Our framework is tractable and can be integrated into macro models of exchange rate determination. As previous shown in the literature, conclusions about optimal monetary and exchange rate policies qualitatively and/or quantitatively change when introducing investor heterogeneity.

Chapter 3.

The Quality of US Imports and the Consumption Gains from Globalization

with Danial Lashkari

Lack of detailed data on the characteristics and quality of imported goods poses a challenge for measuring consumption gains from rising imports. To tackle this problem, we propose a method that allows us to identify demand and to infer unobserved quality change using data only on prices and market shares. Our method applies to a wide class of homothetic demand systems that allow for heterogeneity in the degree of substitutability across products. For this class, we also characterize the contribution of changes in quality, price, and variety entry/exit to the aggregate price index. To validate our approach, we show that it estimates price elasticities and quality changes similar to those found by the standard BLP strategy in data on the US auto market, without relying on the information on product characteristics and price instruments used by BLP. Applying our strategy to the US customs data (1989–2006), we find the average contributions of quality, price, and variety to the annual fall in the price of US imports relative to CPI to be 0.95%, 0.60%, and 0.25%, respectively. Using a demand system that ignores the heterogeneity in product substitutability leads to a substantial overestimation of the extent of quality improvements.

3.1 Introduction

Globalization has offered consumers around the world access to a wider variety of products at cheaper prices. We can measure the value of the resulting gains for consumers in any given country using available customs records on the volumes and unit values of all imported products. The data allows us to construct aggregate indices for the price of imports that transform the observed changes in the volume and variety of imported products into measures of real consumption gain (Feenstra, 1994, Broda and Weinstein, 2006). However, these indices often leave out yet another potential margin for consumption gains through improvements in the quality of each imported product over time. Part of the challenge for evaluating the extent of quality change lies in the fact that customs records, despite their richness, typically lack comprehensive information on product characteristics.

As an example for the potential magnitude of the quality margin, consider the rapid growth of US imports from China, where the latter's share in the total volume of US imports grew from around 2% in 1989 to around 15% in 2006. We may be tempted to attribute this rise mostly to the gradual availability of cheaper Chinese varieties to US consumers. Surprisingly, however, we find that roughly half of the overall rise in the Chinese share of US imports over this period comes from those products where the prices of

Chinese varieties increased relative to advanced countries.[98] This fact suggests that quality upgrading in Chinese imports may have played a crucial role in their rising appeal for US consumers and importers, while being left out of standard measures that evaluate their value for consumers (e.g., the BLS import price index).

More broadly, the problem of accounting for unobserved quality change applies to many macro settings where we aggregate observed changes in quantities and prices across a wide range of products with limited data on their characteristics.[99] In this paper, we develop and implement a novel strategy to address this problem. Our approach builds on the prior insight that product quality may be inferred as the residual demand after accounting for the contribution of prices (Khandelwal, 2010, Hallak and Schott, 2011), and thus requires the estimation of the demand function. We show how to estimate flexible demand functions and to infer product quality if the data only contains information on prices and market shares. We use our method to quantify the contribution of quality change to the aggregate price index of US imports. We show that access to better quality products is the primary source of consumption gains from the rise in import openness in the US over the period 1989-2006, accounting for about 60% of the total

[98]See Figure 3.43 and further details and discussions about this fact in Appendix I.

[99]This problem is sometimes referred to as the *quality change bias* in the measures of inflation in the cost-of-living (Boskin et al., 1998, Gordon and Griliches, 1997). A related problem is one of changing consumer tastes for products and how it should affect our aggregate indices for price or real economic outcomes (Redding and Weinstein, 2020, Baqaee and Burstein, 2022).

decline in relative import prices. Since these quality improvements have remained unmeasured in the standard imports price index, we substantially raise the estimates of consumption gains from US imports over this period.

The estimates of price elasticities play a key role in determining both the inferred changes in quality and the value of new varieties to consumers. Our approach to the identification of demand allows us to consider a wide class of homothetic demand systems featuring heterogeneous price elasticities, and to allow for correlated shocks to marginal costs and demand. We thus improve upon the standard approach for estimating price elasticities in trade data (Feenstra, 1994), which assumes constant elasticities (CES demand) and imposes uncorrelated product-level supply and demand shocks. The latter assumption is untenable if we associate demand shocks with quality.

The idea of our approach is to apply the dynamic panel (DP) methods to the joint evolution of product-level prices and demand (quality) shocks. More specifically, we assume shocks to the current demand (quality) of each product that, conditional on lagged product demand (quality), are uncorrelated with lagged product prices.[100] This assumption is trivially satisfied if we rule out dynamic pricing (e.g., when prices are flexible and demand

[100] This strategy has been combined with complementary instrumental variables in estimating rich demand systems in several IO applications (e.g., Grennan, 2013, Lee, 2013, Sweeting, 2013), and in estimating firm-level production functions (Caliendo et al., 2020). We note that our assumptions about the dynamics of demand shocks are also in line with Redding and Weinstein (2020), who find a strong persistence in demand shocks in the Nielsen barcode data.

does not directly depend on past prices). Alternatively, even under dynamic pricing, this assumption is still satisfied if current demand shocks are outside the information set of the firms when they choose their prices in past periods. Under either scenario, we can derive moment conditions that identify flexible demand systems in the presence of correlated supply and demand shocks. The only additional requirement is that product prices exhibit strong autocorrelation over time.

Since our identification allows us to estimate flexible demand systems, we also provide a theoretical characterization of the changes in the aggregate price index for a broad family of homothetic demand systems. Our results further decompose these changes into the contributions of changes in price, quality, and the available set of products (product entry/exit). The family of demand systems considered here is characterized by up to two distinct aggregate indices and nests the three homothetic demand systems presented by Matsuyama and Ushchev (2017) (see, also Matsuyama, 2022). Thus, our results generalize the widely used Feenstra (1994) variety correction and the unified price index (CUPI) of Redding and Weinstein (2020) from CES to a wide class of homothetic demand systems.[101]

Before applying our method to the customs data, we validate it in the well studied context of data on the US automobile market (1980-2018). In this setting, we have detailed product characteristics, including horse-

[101] For another generalization of the Feenstra variety correction to alternative family of demand systems, and its application to the cereal market in the US, see Foley (2021).

power, miles-per-dollar, and space, that we can use as proxies for product quality. We show that, controlling for lagged product characteristics, current product characteristics are not correlated with lagged prices. This result provides direct evidence in favor of our main identification assumption.

Using the auto data, we further compare our identification strategy against a standard cost shock instrument based on the real exchange rate (RER) between each car's country of assembly and the US. We find similar demand estimates using the two identification strategies, both for CES demand and for Kimball (1995) demand, which is a homothetic demand system with variable elasticities nested within the family considered in our theory. Moreover, we show that our estimated Kimball demand system leads to own-price elasticities that are higher than those of CES, but closely in line with those found based on a random coefficients logit model (BLP) (Berry, 1994, Berry et al., 1995). The latter is the benchmark demand model commonly used in settings with available data on product characteristics. Lastly, we examine our inferred measures of quality and show that they are correlated with characteristics valued by consumers.

To use our strategy for measuring consumption gains from rising imports in the US, we assume a nested demand structure in which consumers evaluate the varieties of goods supplied by different countries using a CES or Kimball aggregator. We express import prices relative to the US consumer price index (CPI). We then create a basket of OECD countries as our benchmark for quality, assuming that the quality of the varieties produced by these countries on average evolves similarly to that of products covered by the CPI in the US. This allows us to express the quality of the varieties supplied by all other countries relative to this baseline set of products.

In the case of Kimball demand (featuring variable price elasticities), we find that our aggregate index of import prices fell by 32% relative to the US CPI from 1989 to 2006 (1.80% annually), and that quality improvement is responsible for a cumulative decline of about 17% (0.95% annually). The remaining part is mostly due to the decline in the relative unit value (unadjusted price) of imported goods, which accounts for an additional 11% cumulative reduction in the aggregate index of import prices (0.60% annually). A smaller role is played by the availability of new varieties, which accounts for a 4.5% cumulative drop in the aggregate index of import prices.[102] Using CES preferences instead of Kimball doubles the gains from openness arising from the product quality channel, largely overstating the quality gains. This confirms the quantitative importance of relaxing the constant elasticity assumption in the standard CES demand systems for evaluating the consumption gains from trade.

[102] Relying on the standard identification approach ruling out correlated supply and demand shocks, Berlingieri et al. (2018) also find that quality change accounts for the bulk of the gains from openness accruing from the trade agreements signed by the EU. Using scanner-level data, Redding and Weinstein (2020) show that the quality bias is sizable relative to the variety bias. Accounting for the additional effect of imports on the consumption of the domestic varieties, Hsieh et al. (2020) find that the increase in imported varieties may be offset by a decrease of domestic varieties based on data from US-Canada trade flows.

While relative product quality over the period rose across most non-OCECD countries, we find that quality upgrading among Chinese products is the major driver of the quality gains to consumers in the US. This finding is consistent with the extensive literature on the effects of the economic reforms that China has been undertaking before and since its accession to the WTO.

Prior Work Our paper is related to the literature that attempts to measure the welfare gains from trade liberalization.[103] While our focus on the consumption side provides an incomplete picture of the overall gains or losses, it averts the need for structural assumptions on the nature of production and leverages the richness of the price data (see also Feenstra and Weinstein, 2017, Berlingieri et al., 2018).[104] We contribute to this literature by accounting for the role of quality and by proposing a novel approach to the estimation of price elasticities that allows for correlations between supply and demand shocks.

The role of product quality for the patterns of international trade and specialization, at the aggregate and at the firm level, has been the subject

[103]In a class of trade theories that lead to a gravity structure for trade flows, Arkolakis et al. (2017) show that we can uncover a combined measure of both production and consumption gains based only on the changes in the share of imports in domestic consumption expenditure. This result has inspired much subsequent work within the framework of quantitative trade theories (for a review, see Costinot and Rodríguez-Clare, 2015).

[104]This insight has recently been used to study the distributional aspects of the consumption gains from trade (e.g., Borusyak and Jaravel, 2018, Adao et al., 2022, Jaravel, 2021). We emphasize that our measures of consumption gains do not provide the full consumption-side welfare effects of rising imports, since the gains due to imports may partly be compensated by a substitution away from domestic consumption (see, e.g., Hsieh et al., 2020).

of a vast body of theoretical and empirical work (e.g., Linder, 1961, Flam and Helpman, 1987; Hummels and Skiba, 2004; Hallak, 2006; Verhoogen, 2008; Fajgelbaum et al., 2011; Baldwin and Harrigan, 2011; Kugler and Verhoogen, 2012; Manova and Zhang, 2012; Martin and Mejean, 2014; Dingel, 2017; Eaton and Fieler, 2022). Early empirical work on the importance of quality proxied product quality with unit values (e.g., Schott, 2004, Hummels and Klenow, 2005).[105] As already mentioned, we follow the approach pioneered by Khandelwal (2010) and Hallak and Schott (2011) in attributing higher quality to products with higher demand, conditional on price.

Our paper is closely related to **?**, who offer a comprehensive attempt as measuring quality in trade flows across many different countries. Unlike our approach, they impose parametric restrictions on the relationship between quality and income elasticity, on the production cost of quality, and on the distribution of product quality in order to construct their quality measures. Our paper is also closely related to the recent paper by Redding and Weinstein (2021), who decompose the different margins of change in US imports, using a detailed nested CES structure that additionally accounts for firm heterogeneity. Relative to these studies, our contribution is to offer a novel identification strategy that only requires assumptions on the dynamics of demand shocks and, crucially, generalizes beyond CES demand

[105]Several studies have relied on measures of quality available for specific sets of products (e.g., wine as in Crozet et al., 2012), or as narrower proxies such as the ISO 9000 management scores (e.g., Verhoogen, 2008).

to allow for heterogeneous elasticities.[106]

Our paper also contributes to the recent work on the importance of accounting for demand and taste shocks in cost-of-living indices (e.g., Gábor-Tóth and Vermeulen, 2018, Ueda et al., 2019, Redding and Weinstein, 2020, Baqaee and Burstein, 2022).[107] In particular, using US retail scanner data where quality is arguably constant at the barcode-level, Redding and Weinstein (2020) derive a formula for the price index under CES demand that accounts for additional variations in demand due to taste shocks. Our estimation strategy allows us to apply their approach to settings in which changes in demand partially reflect changes in product quality. We also show that the CES assumption may overstate the contribution of taste shocks to the indices of cost-of-living.

Finally, a growing body of work in trade and macro goes beyond the standard CES assumption and allows for variations in price elasticities through specifications such as Kimball and HSA demand to study variable markups and pass-through (e.g., Amiti et al., 2019, Baqaee and Farhi, 2020, Wang and Werning, 2020, Matsuyama and Ushchev, 2022).[108] While

[106]In a recent study, Head and Mayer (2021) study counterfactual trade policy exercises in a models with CES and with BLP, in the context of the original automobile market dataset of Berry et al. (1995). While they find similar results, they emphasize the importance of incorporating heterogeneity in pass-throughs through oligopolistic competition under the CES model.

[107]In addition to changes in taste, the dependence of demand on income (nonhomotheticity) also matters for the measurement of consumption gains. Here, we abstract from this consideration by focusing on homothetic demand. Jaravel and Lashkari (2021) provide a method for tackling this problem based on cross-sectional consumption data.

[108]For instance, allowing for variable markups, Feenstra and Weinstein (2017) and Edmond et al. (2015), among others, show that pro-competitive effects of trade liberalization

this literature typically resorts to calibration to match specific moments of interest in the data, we provide a methodology to identify the parameters of such demand systems using data on observed prices and market shares.[109]

The paper is organized as follow. Section 3.2 presents the homothetic demand systems we consider, our approach to their identification, and our theoretical results on the change in their aggregate price index. Section 3.3 presents the results of our estimation approach in the benchmark setting of the US automobile market. Section 3.4 reports our empirical results from the trade data and quantifies the gains from quality. We conclude in Section 3.5.

3.2 Theory

We consider data on prices and market shares (or quantities) of different products or varieties (we use the two terms interchangeably) in a given market. We observe the sequence $(\boldsymbol{s}_t)_{t=0}^{T-1}$ where $\boldsymbol{s} \equiv (s_i)_{i \in V}$ stands for the vector of market shares chosen by the consumer(s) in a set V of products. At time t, a set V_t of products has nonzero market shares (so that $s_{it} = 0$, $i \notin V_t$). We additionally observe the sequence $(\boldsymbol{p}_t)_{t=0}^{T-1}$ where $\boldsymbol{p} \equiv (p_i)_{i \in V_t}$ stands for the vector of prices faced by the consumer(s) in the set V_t of available products.

are quantitatively relevant in the US and Taiwan, respectively. Since we use aggregate trade data, we cannot directly speak to this margin. However, when we apply our method at the firm-level, we can provide measures of markups based on our estimated price elasticities. In our application to the US auto market, we show that our estimated markups are in line with those found by Grieco et al. (2021) using BLP demand.

[109] For an alternative approach to the estimation of HSA demand, see Kasahara and Sugita (2021).

With slight abuse of notation, we also use the notation $\boldsymbol{s} \equiv (s_i)_{i \in V_t}$, where $\boldsymbol{s}$ may alternatively refer to the vector of expenditure shares limited to the set of available products V_t.

Our goal is to characterize the welfare changes of consumers due to changes in the set of available products V_t, changes in the prices of these products, or changes in their (unobserved) qualities. Crucially, we assume no additional information on the characteristics of the products or the underlying production costs. We proceed in three steps. We first state our assumptions regarding the structure of the underlying demand system that rationalizes the observed prices and market shares in Section 3.2.1. We then present our Dynamic Panel (DP) approach to the estimation of demand in Section 3.2.2. We finally present our characterization of the change in the aggregate price index in the entire market in Section 3.2.3.

3.2.1 Homothetic Demand with Variable Elasticities of Substitution

We consider homothetic demand systems that are rationalizable by a well-defined underlying utility function, which, without loss of generality, we can characterize as follows.

Definition 1 (Homothetic Demand System)**.** A homothetic demand system parameterized by a vector of parameters $\boldsymbol{\varsigma} \in \mathbb{R}^D$ can be characterized by a collection of expenditure-share functions $\mathcal{S}_i\left(\cdot;\boldsymbol{\varsigma}\right)$, satisfying $\sum_{i \in V} \mathcal{S}_i\left(\widetilde{\boldsymbol{p}};\boldsymbol{\varsigma}\right) \equiv 1$ for all $\widetilde{\boldsymbol{p}}$ and V, and a linear homogeneous aggregator $\mathcal{H}\left(\cdot;\boldsymbol{\varsigma}\right)$, satisfying

$\mathcal{H}(\alpha \boldsymbol{p}; \boldsymbol{\varsigma}) = \alpha \mathcal{H}(\boldsymbol{p}; \boldsymbol{\varsigma})$ for all $\boldsymbol{p}$, V, and $\alpha > 0$, such that the expenditure share of product $i \in V$ under prices $\boldsymbol{p}$ is given by $\mathcal{S}_i\left(\frac{\boldsymbol{p}}{\mathcal{H}(\boldsymbol{p};\boldsymbol{\varsigma})}; \boldsymbol{\varsigma}\right)$.

Note that the only constraint implied by Definition 1 is the homotheticity of the underlying preferences, since the composition of demand only depends on the relative prices across products (and not on the total consumer expenditure and/or the average level of prices). At this point, the introduction of the aggregator index $\mathcal{H}(\cdot; \boldsymbol{\varsigma})$ is not strictly necessary; it explicitly ensures that the composition of demand does not depend on the level of prices, since multiplying all prices by a factor $\alpha > 0$ leaves the composition of demand intact. For this reason, we will rely on this explicit index in the inversion of the demand system in the estimation. Appendix I presents the choices of expenditure-share and aggregator functions that lead to alternative demand systems such as mixed logit demand, homothetic AIDS and Translog, and the HSA demand system of Matsuyama and Ushchev (2017).

To characterize a general homothetic demand system following Definition 1, we need to specify a $|V|$-dimensional vector of expenditure-share functions $\boldsymbol{\mathcal{S}}(\cdot; \boldsymbol{\varsigma}) \equiv (\mathcal{S}_i(\cdot; \boldsymbol{\varsigma}))_{i \in V}$, in the space of $|V|$-dimensional price vectors. The dimensionality of the corresponding space of cross-product elasticities of substitution grows quadratically in the size of the product space $|V|$. Given that the number of observations grows proportionally to $|V| \times T$, it is not feasible to estimate a demand system that is fully parameterized in this space, unless if we have access to a long panel (such that $|V| \ll T$).[110]

A common alternative is to summarize the patterns of cross-product elasticities of substitution as a function of one or two aggregate indices of all available products. The following definition specializes the general homothetic demand system of Definition 1 to a broad family of demand systems that use up to two such indices.

Definition 2 (Homothetic with Aggregator Demand System)**.** Consider the homothetic demand system of Definition 1 for a linear homogenous aggregator function $\mathcal{H}(\cdot; \boldsymbol{\varsigma})$ and the expenditure-share functions $\mathcal{S}_i(\cdot; \boldsymbol{\varsigma})$ that satisfy

$$\mathcal{S}_i(\widetilde{\boldsymbol{p}}; \boldsymbol{\varsigma}) \equiv \frac{\widetilde{p}_i \, \mathcal{D}_i(\widetilde{p}_i; \boldsymbol{\varsigma})}{\sum_{i' \in V} \widetilde{p}_{i'} \, \mathcal{D}_{i'}(\widetilde{p}_{i'}; \boldsymbol{\varsigma})}, \tag{1}$$

for a collection of single-argument demand functions $\mathcal{D}_i(\cdot; \boldsymbol{\varsigma})$ that are positive-valued and decreasing over some interval $\widetilde{p} \in \left(0, \underline{\widetilde{p}}_i\right)$ and satisfy $\lim_{\widetilde{p}_i \to \underline{\widetilde{p}}_i} \mathcal{D}_i(\widetilde{p}_i; \boldsymbol{\varsigma}) = 0$ and $\mathcal{D}_i(\widetilde{p}_i; \boldsymbol{\varsigma}) = 0$ for $\widetilde{p} \geq \underline{\widetilde{p}}_i$ where $\underline{\widetilde{p}}_i \in \mathbb{R}_+ \cup \{\infty\}$.

We refer to the demand system in Definition 2 as homothetic with aggregator (HA) since we can characterize them using two aggregate indices $H \equiv \mathcal{H}(\boldsymbol{p}; \boldsymbol{\varsigma})$ and $A \equiv \sum_{i'} \frac{p_{i'}}{H} \mathcal{D}_{i'}\left(\frac{p_{i'}}{H}; \boldsymbol{\varsigma}\right)$ as $s_i \equiv \frac{p_i}{A\,H} \mathcal{D}_i\left(\frac{p_i}{H}; \boldsymbol{\varsigma}\right)$. The demand function $\mathcal{D}_i(\cdot; \boldsymbol{\varsigma})$ for product i only depends on the price of product i rela-

[110] If we have access to information on a vector of product characteristics $\boldsymbol{x}_i$ for each product i, we can still express rich patterns of cross-product elasticities of substitution in the space of product characteristics, whose dimensionality does not grow with the number of products $|V|$. For instance, the mixed logit demand system (McFadden, 1974, Berry, 1994) relies on product characteristics to define the expenditure-share functions as $\mathcal{S}_i(\boldsymbol{p}; \boldsymbol{\varsigma}) \equiv \int \frac{\exp(-\sigma \log p_i + \boldsymbol{\beta}' \boldsymbol{x}_i)}{\sum_{i' \in V} \exp\left(-\sigma \log p_{i'} + \boldsymbol{\beta}' \boldsymbol{x}_{i'}\right)} dF(\sigma, \boldsymbol{\beta}; \boldsymbol{\varsigma})$

tive to the aggregate index H that summarizes the effects of the prices of all other products on the demand for product i. This restriction substantially reduces the potential dimensionality of the parameter space.

For each product i, Definition 2 also defines a constant relative choke price $\underline{\widetilde{p}}_i$, as the value of quality-adjusted relative price for which the demand falls to zero. For instance, for the CES demand systems, the demand elasticity function is a constant and the relative choke price is infinity ($\underline{\widetilde{p}}_i \equiv \infty$). More generally, however, consumer demand may fall to zero for finite values of prices.

Definition 2 nests many well-known homothetic demand systems commonly used in the literature, but does not ensure that they are rationalized by an underlying utility function. The following definition characterizes the rationalizable homothetic demand systems recently introduced by Matsuyama and Ushchev (2017) (see also Matsuyama, 2022), which are all nested in HA demand.

Definition 3. The following families of demand HA demand are rationalizable.

1. Homothetic with a Single Aggregator (HSA). This system is characterized by an aggregator function $\mathcal{H}(\cdot;\varsigma) \equiv H$ that is implicitly defined by the value of H that satisfies $1 = \sum_{i \in V} \frac{p_i}{H} \mathcal{D}_i \left(\frac{p_i}{H};\varsigma\right)$.

2. Homothetic Implicit Additive (HIA). This system is characterized by

an aggregator function $\mathcal{H}(\cdot;\boldsymbol{\varsigma}) \equiv H$ that is implicitly defined by the value of H that, depending on the type of HIA demand, satisfies one of the two following conditions

$$1 = \begin{cases} \sum_{i \in V} \int_0^{\mathcal{D}_i\left(\frac{p_i}{H};\varsigma\right)} \mathcal{D}_i^{-1}(v;\boldsymbol{\varsigma})\, dv, & \text{directly additive type,} \\ \sum_{i \in V} \int_0^{\frac{p_i}{H}} \mathcal{D}_i(v;\boldsymbol{\varsigma})\, dv, & \text{indirectly additive type,} \end{cases} \tag{2}$$

where each condition corresponds to one of the two types of HIA demand: directly or indirectly additive.

The homothetic implicitly additive (HIA) systems requires two distinct aggregate indies H and A to characterize demand. In contrast, the homothetic single aggregator (HSA) system is completely characterized using the aggregate index H, and we always have $A \equiv 1$. As shown by Matsuyama and Ushchev (2017), the only demand system that belongs to both HIA and HSA families is the CES demand system, which corresponds to the choice of expenditure-share function $\mathcal{S}_i(\widetilde{\boldsymbol{p}};\boldsymbol{\varsigma}) \equiv \widetilde{p}_i^{1-\sigma}$ and aggregator function $\mathcal{H}(\boldsymbol{p};\boldsymbol{\varsigma}) \equiv \sum_i p_i^{1-\sigma}$ where $\boldsymbol{\varsigma} \equiv (\sigma)$.

In our empirical exercise, we will particularly focus on a special case of the HIA demand, the Kimball demand system, which assumes identical demand functions $\mathcal{D}_i(\cdot;\boldsymbol{\varsigma}) \equiv \mathcal{D}(\cdot;\boldsymbol{\varsigma})$ across products and implicitly defines the aggregator function $\mathcal{H}(\cdot;\boldsymbol{\varsigma})$ as the directly additive case in Equation (2). In this specification, the space of parameters remains constant and does not change in the number of available products $|V|$.

Demand/Quality Shocks For a general, parametric family of homothetic demand systems given by Definition 1, we can assume that the observed sequence of prices and expenditure shares satisfy

$$s_{it} = \mathcal{S}_i \left(\frac{\left(e^{-\varphi_{it}} p_{it}\right)_{i \in V_t}}{\mathcal{H}\left(\left(e^{-\varphi_{it}} p_{it}\right)_{i \in V_t}; \varsigma\right)}; \varsigma \right), \tag{3}$$

where $\left(e^{-\varphi_{it}} p_{it}\right)_{i \in V_t}$ denotes the vector of prices for all products i at time t adjusted by the structural error φ_{it}. The specification in Equation (4) implies that the higher the demand shock φ_{it} for a product with a fixed level of price, the higher the consumer demand will be for the product. Generally, we may interpret φ_{it} as an unobserved demand shock to the quality or appeal of product i at time t. In what follows, we consider cases where we may interpret the variations in this residual demand as being driven by changes in unobserved characteristics $\boldsymbol{x}_{it}$ of products over time and thus refer to it as quality.[111] For instance, we may assume that $\varphi_{it} \equiv \boldsymbol{\beta}' \boldsymbol{x}_{it} + \psi_{it}$ where $\boldsymbol{\beta}$ is a vector specifying the value of each characteristic for consumers.

Note that a constant shift in all demand shock parameters φ_{it} keeps the demand unchanged. We therefore normalize the demand shocks by assuming that there exists a set of base products $O \subset V_t$, for all t, whose

[111] For instance, if we define a product at the level of barcodes in standard scanner data, in which product characteristics $\boldsymbol{x}_i$ for product i remain constant over time, it is more reasonable to assume that demand shocks are driven by changes in product appeal (consumer taste) (e.g., Redding and Weinstein, 2020). In the settings considered in this paper, in which we define products at more aggregate levels, e.g., at the level of a given product classification code, product characteristics are likely to vary over time, and quality change may be the most likely driver of demand shocks (e.g., Khandelwal, 2010). In the latter case, quality change further includes unobserved changes in the set of varieties within each product classification code that is available to consumers.

quality remains on average constant throughout the entire period, implying $\sum_{o \in O} \varphi_{ot} = 0$. Therefore, we interpret φ_{it} as the (unobserved) quality of i relative to the average base product.

3.2.2 The Dynamic Panel Approach to Demand Estimation

In this section, we consider identifying a parametrized homothetic demand system, as characterized by Definition 1, where data on expenditure shares and prices are assumed to follow Equation (3). Let us define the quality-adjusted relative price of product i at time t as

$$\widetilde{p}_{it} \equiv \frac{e^{-\varphi_{it}} p_{it}}{H_t}, \qquad H_t \equiv \mathcal{H}\left(\left(e^{-\varphi_{it}} p_{it}\right)_{i \in V_t}; \boldsymbol{\varsigma}\right). \tag{4}$$

Note that the space of the quality-adjusted relative price vectors $\widetilde{\boldsymbol{p}}_{it}$ at time t constitutes a $(|V_t| - 1)$- dimensional manifold in $\mathbb{R}^{|V_t|}$ since all such vectors satisfy $\mathcal{H}\left(\widetilde{\boldsymbol{p}}_{it}\right) = 1$. We now assume that the demand system satisfies the connected substitutes property of Berry et al. (2013), and is thus a bijection from the space of quality-adjusted relative prices to the space of consumption expenditure shares. As a result, there exists an inverted demand function $\boldsymbol{\pi}\left(\cdot; \boldsymbol{\varsigma}\right)$ such that we have $\widetilde{p}_{it} = \pi_i\left(\boldsymbol{s}_t; \boldsymbol{\varsigma}\right)$.

Let $\langle v_{it} \rangle \equiv \frac{1}{|O|} \sum_{i \in O} v_{it}$ denote the unweighted mean of variable v_{it} within the set of base products O, where $|O|$ is the size of this set. Using Equation (4) and the normalization of quality in the set of base products, we can then write quality shocks as a function of observed expenditure shares and prices according to[112]

$$\varphi_{it} = \log \widehat{p}_{it} - \log \widehat{\pi}_i \left(\boldsymbol{s}_t; \boldsymbol{\varsigma} \right), \qquad i \in V_t. \tag{5}$$

where we have defined the notation where $\log \widehat{v}_{it} \equiv \log v_{it} - \langle \log v_{it} \rangle$ denotes the difference between the logarithm of variable $v_{i,t}$ and its unweighted mean within the set of base products.

Equation (5) offers a parametrized demand function that may be estimated in the data. Needless to say, the key challenge for the identification of this demand system is the potential correlation between the demand shock, log price, and the expenditure shares. We now turn to our approach for tackling this problem.

3.2.2.1 *Identification Assumptions*

We begin by imposing the following restrictions on the stochastic dynamics of the quality shocks.

Assumption 1 (Dynamics of Demand Shocks)**.** *The following Markov process governs the dynamics of quality (demand) shocks φ_{it} for product i at time t:*

$$\varphi_{it} = g_i \left(\varphi_{it-1}; \boldsymbol{\varrho} \right) + u_{it}, \tag{6}$$

[112]By definition, we have $\log p_{it} - \log h_t - \varphi_{it} = \log \pi_i \left(\boldsymbol{s}_t; \boldsymbol{\varsigma} \right)$. Using the condition $\frac{1}{|O|} \sum_{o \in O} \varphi_{ot} = 0$ then leads to Equation (5).

where u_{it} is a zero-mean i.i.d innovation to the demand shock and where $\boldsymbol{\varrho}$ is a vector of parameters characterizing the persistence of the demand shock process.[113]

Equation (6) implies that despite potential persistence in the process of quality shocks, these shocks cannot be completely predicted based on past realizations due to the arrival of innovations in each period. In our baseline model, we assume that the demand shock process is a stationary AR(1) process with a product-specific mean:[114]

$$g_i\left(\varphi_{it-1};\boldsymbol{\varrho}\right)\equiv\rho\varphi_{it-1}+\left(1-\rho\right)\phi_i, \tag{7}$$

where $\boldsymbol{\varrho}\equiv\left(\rho,\boldsymbol{\phi}\right)$ is the vector of the parameters of the Markov process, and where ϕ_i constitutes the expected long-run mean quality of product i.

We next make our main identification assumption, which rules out the dependence of past decisions by firms and consumers on the current innovation to the demand shock.

Assumption 2 (Identification Assumptions)**.** *Demand shock innovations are zero mean, conditional on lagged log prices (and potentially the latter's powers):*

$$\mathbb{E}\left[u_{it}\right|\left(\log p_{it-1}\right)^m\right]=0,\qquad 1\leq m\leq D, \tag{8}$$

where $D\geq 1$ denotes the dimensionality of the parameters characterizing consumer

[113]Note that we can generalize this condition to higher order Markov dynamics, for instance, assuming $\varphi_{it}=g_i\left(\varphi_{it-1},\varphi_{it-2},\cdots;\boldsymbol{\varrho}\right)+u_{it}$, where the contemporaneous demand shock further depends on its higher-order lags.

[114]This model can also account for a process with stationary growth, e.g., a model with $g_i\left(\varphi_{it-1}\right)\equiv\varphi_{it-1}+\gamma_i$, such that $\gamma_i\equiv\lim_{\rho\to 1}\left(1-\rho\right)\phi_i$.

demand. Moreover, we assume that the log price process has a nonzero autocorrelation $\mathbb{E}\left[\log p_{it-1} \log p_{it}\right] \neq 0$.

In combination with Equations (5) and (6), we can use Equation (8) to derive a number of orthogonality conditions that allow us to estimate the vectors of parameters $\boldsymbol{\varsigma}$ and $\boldsymbol{\varrho}$, leading to the following moment conditions

$$\mathbb{E}\left[\left(\log \widehat{p}_{i,t} - \log \widehat{\pi}_i\left(\boldsymbol{s}_t; \boldsymbol{\varsigma}\right) - g_i\left(\log \widehat{p}_{i,t-1} - \log \widehat{\pi}_i\left(\boldsymbol{s}_{t-1}; \boldsymbol{\varsigma}\right); \boldsymbol{\varrho}\right)\right) \times z_{it-1}\right] = 0, \tag{9}$$

where z_{it} is an instrument that is orthogonal to the value of the quality innovation u_{it} for product i at time t, given by the expression within the main parentheses. The instruments z_{it} include lagged values of different powers of log prices $(\log p_{it-1})^m$ for $m \leq D$, a combination of lagged value of the quality shock φ_{it-1} (and potentially its powers) given by Equation (5), and product dummies, depending on the structure of the process $g_i\left(\cdot; \boldsymbol{\varrho}\right)$. For instance, in the case of AR(1) process considered in Equation (7), we use the lagged quality shocks φ_{it-1} and product dummies to identify ρ and ϕ_i's. The assumption of nonzero autocorrelation ensures that the lagged values of log prices offer meaningful instruments for the corresponding contemporaneous values of the same variables.

Example: CES Demand As an example, let us consider the case of CES demand where, as already mentioned, we have $\mathcal{S}_i\left(\widetilde{\boldsymbol{p}}; \boldsymbol{\varsigma}\right) \equiv \widetilde{p}_i^{1-\sigma}$, $\mathcal{H}\left(\boldsymbol{p}; \boldsymbol{\varsigma}\right) \equiv \sum_i p_i^{1-\sigma}$, and where $\boldsymbol{\varsigma} \equiv (\sigma)$. Here, we can analytically write the inverse demand function $\pi_i(\boldsymbol{s};\sigma) \equiv s_i^{1/(1-\sigma)}$. From Equation (5), we can write the quality shock as $\varphi_{it} = \log \widehat{p}_{it} + \frac{1}{\sigma-1} \log \widehat{s}_{it}$. Since a single parameter σ fully characterizes demand, we only need to use the case of $D = 1$ in Equation (8), and thus use the orthogonality conditions $\mathbb{E}[u_{it}|\log p_{it-1}] = 0$, $\mathbb{E}[u_{it}|\varphi_{it-1}] = 0$, $\mathbb{E}[u_{it}|\varphi_{it-1}] = 0$, and $\mathbb{E}[u_{it}] = 0$ for each product i and each time t.

If we further consider the AR(1) assumption in Equation (7), we can leverage the log-linearity of the model and write the moment conditions in first-differences as

$$\mathbb{E}\left[\left(\Delta \log \widehat{p}_{it} + \frac{1}{\sigma-1} \log \widehat{s}_{it} - \rho\left(\Delta \log \widehat{p}_{it-1} + \frac{1}{\sigma-1} \log \widehat{s}_{it-1}\right)\right) \times z_{it}\right] = 0, \tag{10}$$

where $\Delta \log v_{it} \equiv \log v_{it} - \log v_{it-1}$ for any variable v_{it}, and the instruments z_{it} include *double* lagged log prices and demand shocks, in addition to the time and product dummies, corresponding to the case of $D = 1$ in Equation (8). In this case, we can identify the demand elasticity parameter σ and the demand shock persistence ρ without the need to estimating the long-run mean of product-level demand shocks $\boldsymbol{\phi}$ in Equation (7).

3.2.2.2 Discussion

The Logic of Identification To gain more intuition about the assumption in Equation (8), we present an explicit model of firm price setting that satisfies this assumption. Consider the standard environment in which firms flexibly set prices and thus choose them to maximize contemporaneous profits. In this case, the price at a given point in time should only depend on the current variables, and should not depend on the firm's information or forecasts about future product demand and quality. More specifically, letting q_{it} denote the quantity of product i purchased by consumers, this scenario leads to the following process for the evolution of log prices:

$$\log p_{it} = \log mc_i \left(q_{it}, \varphi_{it}, w_{it}\right) + \log \mu_i \left(\boldsymbol{p}_t, \boldsymbol{s}_t, \boldsymbol{\varphi}_t\right) + v_{it}, \tag{11}$$

where $mc_i\left(\cdot,\cdot,\cdot\right)$ is the marginal cost function, which may depend on quantity q_{it}, quality φ_{it}, and exogenous cost shifters w_{it}, $\mu_i\left(\cdot,\cdot,\cdot\right)$ is the markup function, which may depend on the vector of current prices $\boldsymbol{p}_t$, market shares $\boldsymbol{s}_t$, and demand shocks $\boldsymbol{\varphi}_t$ of all products in the market, and where v_{it} is the residual error that is uncorrelated with all other variables of interest. The price setting Equation (11) satisfies Equation (8) even if the firm knows its future demand shock innovation.[115]

More generally, we may consider a model of dynamic price setting in which the log price additionally depends on the expected value of future cost and demand shocks, as well as those of the competitors, conditional on the information set $\mathcal{I}_{it}$ of the firm at that moment in time. In this case, it is

[115]Note that under the assumption of flexible pricing, our identification assumption is weaker compared to the typical assumptions in the application of the dynamic panel methods to production function estimation (see Ackerberg, 2016). In particular, we do not require the assumption that the innovation u_{it} does not belong to the information set of the firm at time $t-1$. With flexible pricing, even if the firm knows its future demand shock, it does not have an incentive to reflect that in its current pricing decision.

sufficient to assume that the firm does not know the future demand shock innovation $u_{it} \notin \mathcal{I}_{it}$ to again satisfy the assumption in Equation (8). Regardless of the underlying model of price setting, the orthogonality assumption allows us to rule out a *direct* functional dependence of the price p_{it} on the future demand shocks φ_{it+1}. Thus, all systematic correlations between log price and the future demand shocks φ_{it+1} are driven by the persistence of the demand shock process φ_{it}.

Comparison with Alternative Approaches to Identification The standard approach to the identification is to use exogenous cost shifters w_{it}, which affect prices through marginal cost as in Equation (11), as instruments to estimate Equation (5). As already mentioned, we are interested in settings where we only have access to information on prices and quantities. Our identification assumption allows us to use the lagged values of log price as an instrument for current log price, after controlling for the expectation of the demand shock conditional on lagged prices. However, we also emphasize that most cost shock instruments used in practice affect the price or costs of specific inputs. To the extent that in response to these shocks firms substitute away or toward those inputs, it is likely that such substitution may additionally affect product quality, thereby violating the exogeneity of some cost shock instruments.

Finally, the conventional approach to estimating demand in the absence of cost shock instruments is that of Feenstra (1994), which rules out correlations between demand shocks φ_{it} and any shocks to prices that are not driven by quantity changes. In particular, any dependence of the marginal cost on quality in Equation (11), i.e., $\frac{\partial \log mc}{\partial \varphi} \neq 0$, violates this assumption. Intuitively, we expect improvements in quality to be associated with more costly inputs, making it likely that this assumption is indeed violated in practice. Section IV in Appendix I provides a detailed discussion of how our assumptions on the dynamics of demand shocks allows us to estimate demand without the need of the identification assumption of Feenstra (1994).

3.2.3 Accounting for Consumption Gains

Since we consider homothetic preferences, we can define a price index (unit expenditure function) P_t that summarizes the effect of the set of available products, their prices, and their quality at time t for the welfare of the consumer(s) into a single number. In this section, we provide a characterization of the change in the price index that accounts for the contributions of each of the three channels (set of available products, prices, and quality).

For the results of this section, we limit our attention to the family of homothetic with aggregator (HA) demand systems specified in Definition 2. Under this family of demand systems, we can define a demand elasticity as a function of quality-adjusted relative price for each product i as

$$\sigma_i\left(\widetilde{p}\right) \equiv -\frac{\widetilde{p}\,\mathcal{D}_i'\left(\widetilde{p}\right)}{\mathcal{D}_i\left(\widetilde{p}\right)}, \tag{12}$$

where we have suppressed the dependence on the parameter vector $\boldsymbol{\varsigma}$ to simplify the expression. Assuming that the observed data satisfies Equation (3), we let $\sigma_{it} \equiv \sigma_i\left(e^{-\varphi_{it}} p_{it}/H_t\right)$ denote the demand elasticity for product i at time t, and denote the corresponding love-of-variety parameter as $\mu_{it} \equiv \frac{1}{\sigma_{it}-1}$.

3.2.3.1 *Exact Measurement of Consumption Gains*

Consider the changes in the set of products, prices, and qualities faced by consumers in the market between periods $t-1$ and t. Define the common set $V_t^* \equiv V_{t-1} \cap V_t$ to be the set of products common between the two periods. We now assume some smooth paths of prices and qualities $(\boldsymbol{p}_\tau, \boldsymbol{\varphi}_\tau)$ in the interval $\tau \in (t-1, t)$ that in either end of the interval approach the values of prices and qualities in periods $t-1$ and t. Formally, we assume these paths satisfy $\lim_{\tau\to t-1}\left(p_{i\tau}, \varphi_{i\tau}\right) = \left(p_{it-1}, \varphi_{it-1}\right)$ for $i \in V_{t-1}$, $\lim_{\tau\to t}\left(p_{i\tau}, \varphi_{i\tau}\right) = \left(p_{it}, \varphi_{it}\right)$ for $i \in V_t$, and

$$\lim_{\tau\to t-1} e^{-\varphi_{i\tau}} p_{i\tau} = H_{t-1}\,\underline{\widetilde{p}}_i \quad \text{for } i \in V_t \backslash V_t^*, \qquad \lim_{\tau\to t} e^{-\varphi_{i\tau}} p_{i\tau} = H_t\,\underline{\widetilde{p}}_i \quad \text{for } i \in V_{t-1} \backslash V_t^*. \tag{13}$$

Importantly, Equation (13) implies that the quality-adjusted relative price of the products that are unavailable in each period approach their corresponding relative choke prices.

Along the paths above, we can apply the definition of the demand system in Equation (3) to define the corresponding paths of expenditure shares $s_{i\tau}$, the aggregate indices H_τ and A_τ, and demand elasticities and love-of-variety parameters $\sigma_{i\tau}$ and $\mu_{i\tau}$. We also define the total expenditure share of the common set as $\Lambda_\tau^* \equiv \sum_{i\in V_t^*} s_{i\tau}$ and the expenditure shares within the common set as $s_{i\tau}^* \equiv s_{i\tau}/\Lambda_\tau^*$ for $i \in V_t^*$. Correspondingly, we also define the expenditure-share weighted mean of any product-specific variable $v_{i\tau}$ within the common set as $\overline{v}_\tau^* \equiv \sum_{i\in V_t^*} s_{i\tau}^* v_{i\tau}$.

Our first result characterizes the change in the price index for any well-defined homothetic with aggregator (HA) demand system along the paths of prices, qualities, and expenditure shares constructed above.

Proposition 1. *The relative change in the price index of an HA demand, specified in Definition 2, at any point on the interval $\tau \in (t-1, t)$ satisfies*

$$d\log P_\tau = d\log D_\tau^* - d\log \Phi_\tau^* + \overline{\mu}_\tau^* \, d\log \Lambda_\tau^* + (\overline{\mu}_\tau^* - \overline{\mu}_\tau)\, d\log A_\tau + \sum_{i\in V_t^*} \mu_{i\tau} ds_{i\tau}^* - \sum_{i\in V_{t-1}\cup V_t} \mu_{i\tau} ds_{i\tau}, \tag{14}$$

where we have defined the Divisia price and quality indices $d\log D_\tau^*$ and $d\log \Phi_\tau^*$ within the common set as

$$d\log D_\tau^* \equiv \sum_{i\in V_t^*} s_{i\tau}^* d\log p_{i\tau}, \qquad d\log \Phi_\tau^* \equiv \sum_{i\in V_t^*} s_{i\tau}^* d\varphi_{i\tau}. \tag{15}$$

Moreover, based on the normalization of quality in the set O of base products, we can also write this change in terms of changes in prices, expenditure shares, and the A_τ aggregator index, as well as the demand elasticities of each product as

$$d \log P_\tau = \langle d \log p_{it} \rangle + \langle \mu_{it} d \log s^*_{it} \rangle$$
$$+ \langle \mu_{it} \rangle \, d \log \Lambda^*_\tau + (\langle \mu_{it} \rangle - \overline{\mu}_\tau) \, d \log A_\tau - \sum_{i \in V_{t-1} \cup V_t} \mu_{i\tau} ds_{i\tau}, \quad (16)$$

where, as before, $\langle v_{it} \rangle \equiv \frac{1}{|O|} \sum_{i \in O} v_{it}$ denotes the unweighted mean of variable v_{it} within the set of base products.

Proof. See Appendix I. □

Equation (14) expresses the growth in the price index at any point along the path as the sum of three main contributions: the first and the second terms account for the changes in the prices and qualities of the continuing products within the common set. The remaining terms on the second line account for the changes in the sets of entering and exiting products.

To unpack this result, let us first consider the special case of the CES demand system where, as we saw, we have $A_\tau \equiv 1$, $\sigma_{i\tau} \equiv \sigma$ and $\mu_{i\tau} \equiv \frac{1}{\sigma - 1}$. As a result, Equations (14) and (16) simplify to

$$d \log P_\tau = d \log D^*_\tau - d \log \Phi^*_\tau + \tfrac{1}{\sigma - 1} \, d \log \Lambda^*_\tau, \quad (17)$$

$$= \langle d \log p_{it} \rangle + \tfrac{1}{\sigma - 1} \langle d \log s^*_{it} \rangle + \tfrac{1}{\sigma - 1} \, d \log \Lambda^*_\tau. \quad (18)$$

The three terms in the first equation account for the contributions of the change in price, quality, and product entry/exit. Since the means in the set of base products are unweighted, we can explicitly integrate Equation (18) to find the following exact result for the change in the CES price index:[116]

$$\Delta \log P_t = \langle \Delta \log p_{it} \rangle + \tfrac{1}{\sigma-1} \langle \Delta \log s_{it}^* \rangle + \frac{1}{\sigma - 1} \Delta \log \Lambda_\tau^*. \tag{20}$$

Consider the case where we assume that the set of base products corresponds to the current set, $O \equiv V_t$. In this case, Equation (20) corresponds to the logarithm of the CES unified price index (CUPI) defined by Redding and Weinstein (2020): the first term is the logarithm of the Jevons index within the common set, the second term is the logarithm of the geometric mean of the relative change in the expenditure shares within the common set, and the last term is the standard Feenstra (1994) CES correction for the contributions of product entry/exit.

Once we deviate from the CES assumption, Equations (14) and (16) show how the heterogeneity in the demand elasticities σ_{it} affect the change in the unit expenditure function P_τ. First, comparing Equations (14) and (17), we find that in the presence of heterogeneity in demand elasticities, the contribution of product entry and exit to the change in the price index

[116] Integrating Equation (17), we also find the following exact decomposition of the change in the CES price index to changes in price, quality, and the set of available products (Redding and Weinstein, 2020):

$$\Delta \log P_t = \sum_{i \in V_t^*} \tilde{s}_{it}^* \Delta \log p_{it} + \sum_{i \in V_t^*} \tilde{s}_{it}^* \Delta \varphi_{it} + \tfrac{1}{\sigma-1}\, d \log \Lambda_\tau^*, \tag{19}$$

where $\tilde{s}_{it}^* \propto \Delta s_{it}^* / \Delta \log s_{it}^*$ are the Sato-Vartia weights defined in the common set, satisfying $\sum_{i \in V_t^*} \tilde{s}_{it}^* = 1$.

is given by

$$\underbrace{\overline{\mu}_\tau^* \, d\log \Lambda_\tau^*}_{\text{generalized Feenstra correction}} + \underbrace{(\overline{\mu}_\tau^* - \overline{\mu}_\tau)\, d\log A_\tau}_{\text{love-of-variety gap}} + \underbrace{\left(\sum_{i \in V_t^*} \mu_{i\tau} ds_{i\tau}^* - \sum_{i \in V} \mu_{i\tau} ds_{i\tau} \right)}_{\text{love-of-variety reallocation}}.$$

The first term above generalizes the Feenstra (1994) CES variety correction to the case with heterogeneous demand elasticities. In this case, the relevant love-of-variety index is the weighted mean $\overline{\mu}_\tau^*$ of love-of-variety parameters within the common set. The second term accounts for the gap between the mean love-of-variety index within the common set and across all products. The third term shows that we need to additionally account for the effects of the reallocations of consumer expenditure across products that have different degrees of substitutability for consumers. More specifically, this term corresponds to the gap between these reallocations across the set of all products and those within the common set. If reallocations toward products with higher love of variety are stronger outside relative to inside the common set, this expression predicts a lower change in the price index than what is predicted by the Feenstra (1994) CES variety correction.

Equations (14) and (16) expressed in terms of the change in the aggregate index A_τ. We can further simplify these expressions by removing this term for the HSA and HIA demand systems of Definition 3. In the case of HSA, we have that $A_t \equiv 1$. In the HIA case, we can show that the unit expenditure function is given by the product of the two aggregate indices:

$$P_\tau = H_\tau \, A_\tau. \tag{21}$$

Using these observations, the following lemma characterizes the change in the aggregate index A_τ as a function of the love-of-variety weighted change in the market shares of different products. Using this lemma allows us to expressed the change in the unit expenditure function only as a function of changes in prices and expenditure shares, and the demand elasticities.

Lemma 1. *For the HSA and HIA demand systems of Definition 3, the change in the price index satisfies*

$$d \log A_t = \begin{cases} 0, & HSA, \\ \\ -\frac{1}{1+\bar{\mu}_t} \sum_{i \in V} \mu_{it} ds_{it}, & HIA. \end{cases} \tag{22}$$

Proof. See Appendix I. □

Unlike the CES case, in the presence of heterogeneity in demand elasticities, we cannot exactly integrate the above results to construct the exact measures of change in the price index. We will instead construct second-order approximations for the change in the price index that we can compute in the data.

3.2.3.2 Approximate Measures of Consumption Gains

Since the paths that we constructed in Section 3.2.3.1 between periods $t-1$ and t in the limit approach the outcomes in those two periods, we can approximately integrate Equation (16) to find the change in the unit expenditure function between these two periods.

Define the Trnqvist average $\overline{\overline{v_{it}}} \equiv \frac{1}{2}\left(v_{it-1} + v_{it}\right)$ of variable v_{it} between periods $t-1$ and t. In particular, in computing $\overline{\overline{\mu}}_{it}$ for products that are outside the common set ($i \notin V_t^*$), we use the love of variety for product i at its relative choke price. For instance, for the products that enter between periods $t-1$ and t ($i \in V_t \backslash V_t^*$), we let $\overline{\overline{\mu_{it}}} \equiv \frac{1}{2}\left(\underline{\mu}_i + \mu_{it}\right)$ where $\underline{\mu}_i \equiv \lim_{\widetilde{p} \to \underline{\widetilde{p}}_i} \frac{1}{\sigma_i(\widetilde{p})-1}$. Using these definitions, the following lemma characterizes the change in the homothetic price index of any HA demand system up to the second order of approximation.

Lemma 2. *The relative change in the price index of any HA demand system, specified following Definition 2, between periods* $t-1$ *and* t *satisfies*

$$\Delta \log P_t = \sum_{i \in V_t^*} \overline{\overline{s_{it}^*}} \Delta \log p_{it} - \sum_{i \in V_t^*} \overline{\overline{s_{it}^*}} \Delta \varphi_{it} + \overline{\overline{\overline{\mu}_t^*}} \Delta \log \Lambda_t^* + \overline{\overline{\left(\overline{\mu}_t^* - \overline{\mu}_t\right)}} \Delta \log A_t + \sum_{i \in V_t^*} \overline{\overline{\mu_{it}}} \Delta s_{it}^* - \sum_{i \in V} \overline{\overline{\mu_{it}}} \Delta s_{it}^* + O\left(\delta^3\right), \tag{23}$$

up to the second-order terms in $\delta \equiv \max\{\Delta \log \Lambda_t^*, \max_i |\Delta \log p_{it}|, \max_i \Delta \varphi_{it}\}$, *as δ approaches zero. Moreover, this relative change can also be written as*

$$\begin{aligned}\Delta \log P_t = \langle \Delta \log p_{it} \rangle + \langle \overline{\overline{\mu_{it}}}\, \Delta \log s^*_{it} \rangle + \overline{\overline{\langle \mu_{it} \rangle}}\, \Delta \log \Lambda^*_t \\ + \overline{\overline{(\langle \mu_{it} \rangle - \overline{\mu}_t)}}\, \Delta \log A_t - \sum_{i \in V} \overline{\overline{\mu_{it}}}\, \Delta s_{it} + O\left(\delta^3\right),\end{aligned} \tag{24}$$

Proof. See Appendix I □

Equation (23) constitutes one of our main theoretical results. It provides a decomposition of the changes in the price index in a broad family of homothetic demand systems to the contributions of changes in prices, quality, and the set of available products. For any parameterized family of homothetic demand, applying our estimation scheme in Section 3.2.2 allows us to find the implied values of demand elasticity σ_{it} for each product i in the set of products V_t at time t and compute the index A_t. We can then apply Lemma 2 to compute the change in the price index between the two periods. Using the resulting estimates of quality change, we can also find the second-order approximation provided in Equation (23) for the decomposition of the change in the price index to the contributions of price change, quality change, and product entry/exit. Appendix II uses the results of Lemma 1 to remove the need for computing the change in index A_t under the HSA/HIA demand families.

3.3 Validating the Strategy using US Auto Data

In this section, we apply the Dynamic Panel (DP) approach for demand estimation to detailed data on the US automobile market and compare the resulting estimates with those found using benchmark methods of demand estimation including the random coefficient logit model (Berry, 1994, Berry et al., 1995).

3.3.1 Data

We use data on the US automobile market from 1980 to 2018. The Wards Automotive Yearbooks contain information on specifications, list prices and sales by model for all cars, light trucks, and vans sold in the US.[117] Vehicle characteristics include horsepower, miles-per-dollar, miles-per-gallon, weight, width, height, style (car, truck, SUV, van, sport), and producer. Additional information such as the producer's region, whether the model is an electric vehicle, a luxurious brand, or a new design (redesign), complement the data from the yearbooks.[118] We perform standard cleaning to the data following Grieco et al. (2021) and Berry et al. (1995), and, in addition, we exclude models that have an average price higher than $100k over the en-

[117]The Wards Automotive Yearbooks contain information for all trims (variants) of each model. Following standard practice, we aggregate all information at the model level based on the median across trims (Berry et al., 1995, Grieco et al., 2021).

[118]Table 3.18 in Appendix I provides additional details and displays summary statistics for our sample.

tire time period and drop observations with a change in market share above (below) the 99th (1th) percentile within each year.[119]

We follow Grieco et al. (2021) and Goldberg and Verboven (2001) in the construction of an exogenous instrument for prices based on exchange rates. We use the lagged bilateral real exchange rate between the US and the country of assembly of each model, henceforth RER.[120] RER constitutes an arguably exogenous shifter of production costs capturing, in part, local labor market conditions in the country of assembly. This is because exogenous changes in local wages are reflected on the local price level and, in turn, on the real exchange rate. In addition, exogenous movements in the nominal exchange rate between the US and the country of assembly represents another source of variation for the RER as firms can lower their prices when the local currency depreciates.

Before applying our methodology for demand estimation, we rely on the availability of product characteristics to directly test our identification assumption (Assumption 2). In Appendix II, we show that lagged log prices are uncorrelated with current product characteristics after controlling for lagged product characteristics. In addition, product characteristics exhibit strong autocorrelations, supporting our Markov process assumption for the

[119]As in Berry et al. (1995), we define the new variable "space" as the product between length and width and exclude observations with a value larger than 6. Similarly, we define the ratio of horsepower per 10lbs and exclude observations with a value larger than 3.

[120]The RER is constructed as the ratio of the expenditure price levels between the assembly country and the US. The expenditure price levels are available from the Penn World Tables. See Grieco et al. (2021) for additional details.

dynamics of product-level quality.

3.3.2 *Empirical Demand Specification*

In applying our framework to the auto data, we map each car model to a product/variety i in our data. For the specification of demand, we will rely on a particular parameterized family of HIA demand systems (Definition 3) commonly referred to as Kimball demand (Kimball, 1995). This specification corresponds to the directly additive HIA type, as specified in Equation (2), with identical demand functions $\mathcal{D}_i\left(\widetilde{p};\varsigma\right) \equiv \mathcal{D}\left(\widetilde{p};\varsigma\right)$, which are nonnegative-valued and decreasing for all $\widetilde{p} \leq \underline{\widetilde{p}}$ for a relative choke price $\underline{\widetilde{p}}$. This demand system can rationalized by a homothetic utility function (aggregator) Q_t as a function of a vector of quantities $\boldsymbol{q}_t \equiv (q_{it})_{i\in V_t}$, implicitly defined through

$$\sum_{i\in V_t} \mathcal{K}\left(\frac{q_{it}}{Q_t};\varsigma\right) = \mathcal{K}\left(1\right), \tag{1}$$

where the Kimball function is given by $\mathcal{K}\left(\widetilde{q}\right) \equiv \int_0^{\widetilde{q}} \mathcal{D}^{-1}\left(v;\varsigma\right) dv$ for the corresponding demand function $\mathcal{D}\left(\cdot;\varsigma\right)$.

We consider a number of different parameterizations of the Kimball function, characterized using the *Kimball elasticity* functions:

$$\mathcal{E}\left(\widetilde{q};\varsigma\right) \equiv -\frac{\widetilde{q}\,\mathcal{K}''\left(\widetilde{q};\varsigma\right)}{\mathcal{K}'\left(\widetilde{q};\varsigma\right)} = \frac{1}{\sigma\left(\mathcal{D}^{-1}\left(\widetilde{q};\varsigma\right)\right)}, \tag{2}$$

where in the second equality we have used the definition of the demand elasticity function in Equation (12), and the demand relations $\mathcal{K}'(\widetilde{q};\boldsymbol{\varsigma}) = \mathcal{D}^{-1}(\widetilde{q};\boldsymbol{\varsigma})$ and $\widetilde{q} = \mathcal{D}\left(\widetilde{p};\boldsymbol{\varsigma}\right)$. Given our assumptions on the Kimball function $\mathcal{K}(\cdot)$, the elasticity function $\mathcal{E}\left(\cdot\right)$ is positive-valued for all $\widetilde{p} < \underline{\widetilde{p}}$.[121]

We recover standard CES preferences by choosing Kimball function $\mathcal{K}(\widetilde{q};\boldsymbol{\varsigma}) \equiv \widetilde{q}^{1-1/\sigma}$ in Equation (1) with the corresponding choice of parameterization $\boldsymbol{\varsigma} \equiv (\sigma)$. Below, we consider three additional parametric families of Kimball functions $\mathcal{K}\left(\cdot;\boldsymbol{\varsigma}\right)$, each characterized by a corresponding family of elasticity functions $\mathcal{E}\left(\cdot;\boldsymbol{\varsigma}\right)$.

1. *Klenow and Willis (2006).* This case involves an elasticity function

$$\mathcal{E}\left(\widetilde{q};\boldsymbol{\varsigma}\right) \equiv \frac{\widetilde{q}^{\theta}}{\sigma}, \qquad \boldsymbol{\varsigma} \equiv (\sigma,\theta) \tag{3}$$

   that goes from zero (corresponding to infinite price elasticity) to infinity as the normalized quantity goes from zero to infinity.

2. *Finite–Infinite Limits:* This case involves an elasticity function

$$\mathcal{E}\left(\widetilde{q};\boldsymbol{\varsigma}\right) \equiv \frac{1}{\sigma + \left(\sigma_o - \sigma\right)\widetilde{q}^{-\theta}}, \qquad \sigma < \sigma_o,\ \theta > 0,\ \boldsymbol{\varsigma} \equiv \left(\sigma,\sigma_o,\theta\right), \tag{4}$$

   that goes from zero (corresponding to infinite price elasticity) to a finite value $1/\sigma$ as the normalized quantity goes from zero to infinity.

[121] We may consider additional constraints that imply this function is also nonincreasing and is smaller than unity, implying price elasticities of demand that exceed unity and are nondecreasing in quantity (satisfying Marshall's Second Law of Demand).

3. *Finite–Finite Limits:* This case involves an elasticity function

$$\mathcal{E}\left(\widetilde{q};\varsigma\right) \equiv \frac{1}{\sigma_o} + \left(\frac{1}{\sigma} - \frac{1}{\sigma_o}\right) \frac{e^{\theta_o}\widetilde{q}^{\theta}}{1+e^{\theta_o}\widetilde{q}^{\theta}}, \qquad \sigma < \sigma_o,\ \theta > 0,,\ \varsigma \equiv \left(\sigma,\sigma_o,\theta,\theta_o\right), \tag{5}$$

that goes from a finite value $1/\sigma_0$ to another finite value $1/\sigma$ as the normalized quantity goes from zero to infinity.[122]

Appendix II derives the family of Kimball functions $\mathcal{K}(\cdot;\varsigma)$ corresponding to each of the three cases above.

3.3.3 Benchmark Empirical Models

Our goal is to examine two distinct aspects of the approach we proposed in Section 3.2: the effectiveness of the DP approach as an identification strategy, and the ability of a homothetic with aggregator (HA) demand system, e.g., the Kimball demand system, to provide a satisfactory account of heterogeneity in price elasticities. First, to study the identification aspects, we estimate a standard CES specification using the DP approach and compare it against the standard instrumental variable approach that uses cost shocks (RER). In the latter case, we take advantage of the information on product characteristics to directly proxy for product quality. Second, to study the properties of the Kimball specification, we compare it against the current

[122] In the first and the last cases, the marginal utility of consuming every product at a zero level of consumption ($\widetilde{q}_i = 0$) is infinity. Therefore, the demand takes a finite, nonzero value for every finite value of price. In contrast, in the second case, the marginal utility of consuming every product at a zero level of consumption ($\widetilde{q}_i = 0$) is finite. As a result, there is a finite choke price for any product, above which the consumption falls to zero.

workhorse demand model for differentiated products, i.e., the random coefficient logit model (Berry, 1994, Berry et al., 1995). In this exercise, we also compare the estimates of the Kimball specification using the two alternative identification strategies: the DP approach and the standard cost shock IV approach. Below, we discuss the details of these alternative benchmark models.

To study the properties of the DP identification strategy, we consider the CES specification that leads to a simple log-linear relationship between market shares and prices to estimate the elasticity of substitution σ:

$$\log s_{it} = -(\sigma - 1)\ \log p_{it} + \beta \boldsymbol{x}_{it} + make_i + \delta_t + \epsilon_{it}, \tag{6}$$

where $make_i$ specifies the producer of product i. Here, $\boldsymbol{x}_{it}$ stands for the vector of product characteristics, including space, horsepower, miles-per-dollar, luxury brand, vehicle type (sport, electric, truck, suv, van). As mentioned, we can address the endogeneity of prices using a proxy for the costs of production, the real exchange rate (RER) in the assembly country, as a price instrument and also controlling for product characteristics and time and producer fixed effects. We also estimate the specification in Equation (6) using ordinary least squares, as an additional benchmark for the instrumented regressions.

We also estimate Equation (6) with the DP approach, using the moment conditions in first-differences as in Equation (10) and relying on double-lagged prices and market shares as instruments, together with time fixed effects. In this case, we use the Chevrolet Corvette model as the reference product for the estimation. However, for all welfare calculations, the measure of inferred quality is normalized such that the average change in quality of the set of continuing models that are not redesigned is zero (Grieco et al., 2021).[123]

As mentioned, we next compare our Kimball specification against the empirical discrete choice model of differentiated products presented in Berry (1994) and Berry et al. (1995) (henceforth BLP). The BLP method assumes heterogeneous consumers, whereby the utility u_{nit} of consumer n for a product i with the vector $\boldsymbol{x}_{it}$ of product characteristics is given by $u_{nit} = \alpha\, p_{it} + \beta\, \boldsymbol{x}_{it} + \alpha_n p_{it} + \beta_n \boldsymbol{x}_{it} + \epsilon_{nit}$, where the consumer-specific coefficients α_n and β_{nk} on price and characteristic k, respectively, are zero-mean, gaussian-distributed, *i.i.d.* sources of unobserved heterogeneity in consumer taste. Following standard practice, we normalize to zero the utility of the outside option to not purchase any available model. We estimate the random coefficients model including the same set of product characteristics as in the CES specification, using the RER as a cost-shock instrument, and following the best practices as in Conlon and Gortmaker (2020).

[123] For the set O of continuing models that are not redesigned, $\frac{1}{|O|} \sum_{o \in O} \Delta \varphi_{ot} = 0$.

TABLE 3.1 – Price Elasticity

	CES				Kimball	
	OLS	IV	DP	BLP	IV	DP
Mean	1.979	4.637	4.254	7.618	7.862	8.581
	(0.200)	(1.135)	(1.647)	(0.442)	(1.472)	(1.368)
Median				6.706	6.793	7.419
				(0.389)	(1.008)	(1.010)
Weighted Mean				6.890	5.462	5.839
				(0.364)	(0.641)	(0.890)
IQR				4.063	2.929	3.366
				(0.240)	(0.843)	(0.966)

Note: The table reports the estimated own-price elasticities for the full sample. Each column corresponds to a different econometric model: CES OLS, CES IV, CES DP, BLP, Kimball IV, and Kimball DP. For the CES cases, we report the own-price elasticity while for the VES cases (BLP and Kimball) we report a set of moments from the distribution of the estimated price elasticities. For the BLP and the Kimball specifications, we report the mean and the median elasticity together with the expenditure weighted mean elasticity and the interquartile range. For each coefficient we report the 95% confidence intervals. For the CES specifications, standard errors are clustered at product (model) level. The standard errors of the statistics for the Kimball specifcations are obtained from N=100 bootstrapped samples (using models as resampling unit). Due to computational limitations, we follow Conlon and Gortmaker (2020) in copmuting standard errors for the BLP statistics from a parametric bootstrap procedure (we draw 100 different sets of coefficients from the estimated joint distribution of parameters and compute the median under each of these parametric bootstrap samples).

Finally, we estimate the three parametric families of Kimball functions presented in Equations (5), (4) and (3), using the moment condition in Equation (9). We estimate the Kimball specification using both the DP identification strategy and the RER as a cost-shock instrument. Here, too, we choose the Chevrolet Corvette as reference product for the estimation while quality is normalized with respect to the set of continuing models that are not redesigned. For the DP case, we use lagged prices and their quadratic powers as instruments, as well as time and producer fixed effects. For the standard IV approach, we use RER, log(RER) and their powers as instrument.

3.3.4 *The Comparison of Estimated Own-Price Elasticities*

In Table 3.1, we report the estimated price elasticities found by the different approaches for the whole sample. The first three columns show the estimated price elasticity under the CES specification using OLS estimation, using the RER variable as the cost shock instrument (IV henceforth), and using our DP approach. The remaining four columns display different moments of the distribution of the estimated own-price elasticities under the two models with variable elasticities, the BLP and the Kimball specifications. In the latter case, the table also shows the estimates when using the RER as the cost shock instrument and when using our DP approach.

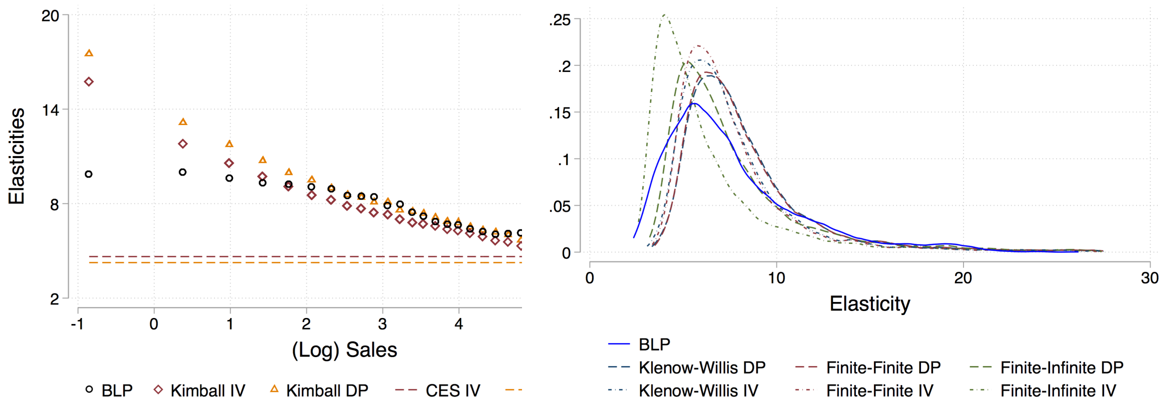
As expected, we find that the OLS estimate of the CES price elasticity displays a bias towards zero due to the positive correlation between demand and price shocks, despite the fact that our specification includes product characteristics to control for quality. When we use the cost shock instrument, the magnitude of the estimated CES elasticity rises relative to its OLS counterpart (1.98 from 4.64). The latter estimates suffer from downward bias due to correlation between prices and demand shocks. This result confirms the need for price instruments to correct for the endogeneity bias in this setting.

Importantly, applying the DP approach to the CES specification delivers a CES elasticity of substitution of 4.25, close to the estimated elasticity obtained with the cost shock instrument. This suggests that our DP approach provides a solution for the endogeneity problem without relying on additional costs shocks, and even without controlling for product characteristics.

How important is accounting for heterogeneity in price elasticities? Comparing the estimates under the CES and the BLP models, we find that ignoring the heterogeneity in price elasticities leads to a bias toward zero under the former. The median, the unweighted, and the weighted means of the estimated elasticities are larger under the BLP specification compared to the CES. Despite its simplicity, the Kimball specification also appears to allow for sufficient heterogeneity to circumvent this problem: all three moments of the distributions of the estimated own-price elasticities under Kimball are closer to those under BLP, when compared to those of CES. Moreover, we again find that the Kimball estimates found using the cost shock instrument and using the DP approach are close, providing additional evidence of the validity of the DP approach.

We next explore the relationship between the volume of sales and the estimated elasticities across products under the BLP and the Kimball models. The left panel of Figure 3.1 shows that this relationship is similar between the BLP specification and the Kimball specification, when estimated under both identification strategies (DP and IV). This result confirms that the Kimball specification can indeed account for the same relationship between sales and price elasticity as that uncovered by the BLP specification,

FIGURE 3.1 – Elasticity Heterogeneity in Kimball and BLP



Note: The left panel plots a binscatter representation of the relationship between (log) sales and the estimated elasticity of substitution. Products with (log) sales less than -1 are dropped. We consider the set of elasticities estimated from: i) the BLP model; ii) the Finite-Finite Kimball model using cost shocks (RER) as instruments (Kimball IV); iii) the Finite-Finite Kimball model using the DP approach (Kimball DP). We also report the CES elasticity estimated using IV and DP. The right panel shows the distribution of elasticities of all Kimball specifications (Finite-Finite, Finite-Infinite and Klenow-Willis) estimated using both the DP and IV instruments. The distribution of BLP elasticities is also reported. Values are truncated at 25.

both qualitatively and quantitatively, and that the DP approach can identify this pattern without the use of any additional information other than prices and market shares.

The right panel of Figure 3.1 shows that the entire distribution of elasticities estimated by the BLP method is similar to those estimated under the different Kimball specifications and using the two different identification strategies.[124] This result, in addition to the evidence on the similarity of the interquartile range values reported in Table 3.1, confirms that the heterogeneity in the price elasticities estimated under the Kimball specification bears a close resemblance to that under the BLP specification.[125] Moreover,

[124] See also Figure 3.27 in Appendix J for additional comparisons across Kimball specifications and identification strategies.

[125] Note that in the Kimball case, the heterogeneity in elasticities is entirely due to the heterogeneity in market shares. In contrast, the heterogeneity in the elasticities estimated by the BLP method may additionally stem from the heterogeneity in product characteristics as well.

it shows that the distribution of elasticities, estimated using both the DP and the IV approaches, is robust to the choice of different families of the Kimball functions (Finite-Finite, Finite-Infinite and Infinite-Infinite).

3.3.5 Inferred Quality and Product Characteristics

Using detailed data on the US automobile market allows us to examine whether our approach retrieves meaningful measures of quality. We examine this question by quantifying the correlation between our inferred measures of quality and the product characteristics valued by consumers available in our dataset. We again compare the results of our DP approach for the CES specification to alternative estimation strategies such as OLS and the standard IV approach using RER. We also explore the implications of accounting for heterogeneity in price elasticities for the inferred quality (compared to the standard CES case).

In the CES case, the inferred quality of each product i at time t is computed according Equation (5) in which we use the elasticity estimated using the DP approach and reported in Table 3.1. Similarly, inverting the Kimball demand, we infer the measure of product quality for the Kimball case using Equation (5).[126] We then study the correlation between the quality measure φ_{it} (inferred using either the CES or Kimball estimates) and a subset of product characteristics tightly linked to product quality in this specific

[126]See the discussion in Appendix III for more details on inverting the Kimball demand.

market, e.g., horsepower, space, miles-per-dollar and style:

$$\varphi_{it} = \beta \boldsymbol{x}_{it} + \eta_t + \gamma_i + \epsilon_{it}, \tag{7}$$

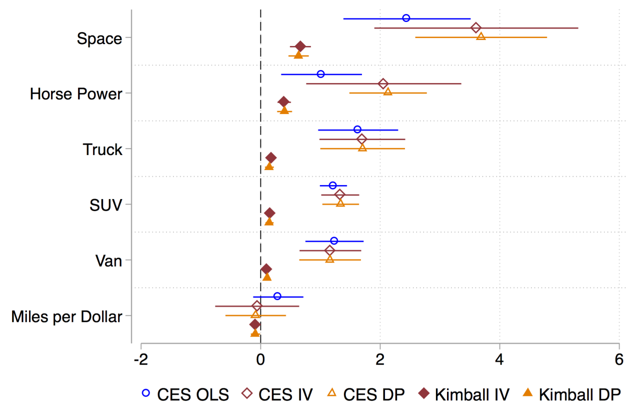
where $\boldsymbol{x}_{it}$ is the set of characteristics listed above. The correlation coefficients estimated from regression (7) are compared against the coefficients estimated from Equation (6) above.[127]

Figure 3.2 shows that the inferred quality estimated using DP and using the cost shock (RER) identification are related to product characteristics almost identically, in both the CES and the Kimball specifications. This is a direct consequence of the ability of the DP approach to correctly estimate price elasticities, as shown in the previous section. Notice that the correlation between inferred quality and product characteristics differs across model specifications. Even though the correlations exhibit the same qualitative patterns, the magnitude is stronger in the CES specification compared to Kimball. The quantitative difference across models suggests that accounting for heterogeneity in price elasticity has a first order role in quantifying the role of quality.

If we assume that the market structure is characterized by monopolistic competition, the markup charged for each vehicle-year is given by $\mu_{it} = \frac{1}{\sigma_{it}-1}$, where σ_{it} is the estimated price elasticity for vehicle i at time t. Given

[127] We re-estimate Equation (6) above using the same set of product characteristics and fixed effects as in regression (7).

FIGURE 3.2 – Correlation between Inferred Quality and Product Characteristics

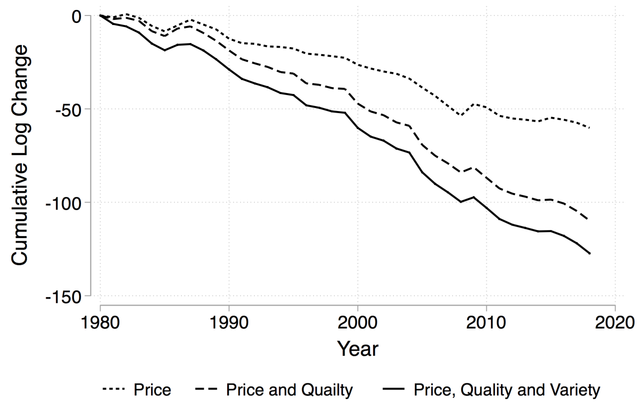


Note: The figure reports the relationship between product characteristics and inferred quality. In the CES DP case, the inferred quality measure follows from Equation (5). For the Kimball specification, inferred quality is obtained inverting demand as in Appendix *III*. The coefficients referring to the DP approach (CES and Kimball) and the Kimball IV case are obtained from regression in Equation (7). We consider the following product characteristics: horse power, space, miles-per-dollar and style (suv, truck, van). The coefficients referring to the OLS and IV estimates of the CES specification are obtained from Equation (6), where product characteristics are used to proxy for quality. All regressions use the entire sample and includes time and product fixed effects. Standard errors are clustered at the producer level, the bands around the estimates show the 95% confidence intervals.

this measure of markups, we infer the marginal cost of each vehicle to be $mc_{it} = \frac{p_{it}}{1+\mu_{it}}$. The right panel of Figure 3.29 in Appendix III shows that there is a strong positive relationship between a proxy of input cost, the weight of the vehicle multiplied by the price of steel, and our measure of inferred marginal cost, supporting the relevance of the latter. The left panel of Figure 3.29 shows that higher quality models have lower price elasticities and, thus, higher markups. The right panel of Figure 3.29 displays a positive relationship between inferred quality and the cost of production, in line with the findings of the prior literature on product quality (e.g., Verhoogen, 2008).[128]

[128]Consistent with this evidence, Figure 3.30 in Appendix III shows that our measure of marginal costs is strongly correlated with the product characteristics consumers value (e.g. horsepower, space and miles-per-dollar). Moreover, these results are also consistent with

FIGURE 3.3 – The Price Index for the US Auto Market



Note: The figure plots the price index for the auto market and its decomposition into the unadjusted price, quality improvement and variety components. We use the estimates from the Finite-Finite Kimball specification estimated using the DP approach. The solid line represents the price index including all three components. The dashed and dotted lines represent the price and quality components together and the price component only, respectively. Prices are deflated using the CPI index from BLS. The measure of inferred quality is normalized such that the average change in quality of the set of continuing models that are not redesigned is zero.

3.3.6 *Consumption Gains in the Auto Market*

We construct the price index for the entire US auto market following Section 3.2.3 and analyze its evolution, quantifying the contribution of changes in unit price, quality, and the set of available models for consumers. We express the price changes relative to the CPI index constructed by the BLS. As before, quality is normalized such that the average quality change in the set of continuing models that are not redesigned between each two consecutive years is zero.[129]

In Figure 3.3 we plot the Kimball price index for the US auto market over the 1980-2018 period, highlighting the role of the price, quality, and

Atkin et al. (2015), who show direct evidence for the relationship between markups and costs

[129]See footnote 123 for details on the normalization of quality.

variety channels. The price index on average declines by around 3.3% annually relative to the CPI over this period. Almost half of the annual decline (1.58%) can be attributed to the decline in unadjusted unit price. Quality improvement contributes substantially to the overall fall in the price index, accounting for an additional 1.3% average annual decline. Figure 3.3 shows that the contribution of the availability of new models is marginal compared to the other two channels, accounting for a 0.46% annual drop in the aggregate price index.[130] Table 3.20 in Appendix III compares the price index for Kimball to the price index for the CES case. The annual decline in the price index is 4% larger in the CES case because the contribution of quality improvements is largely overestimated (4.6% in the CES case compared to 1.3% in the Kimball case). We find that our conclusions about the quantitative role of quality improvement for welfare changes strongly depend on our assumptions about the underlying structure of demand.[131]

3.4 Consumer Gains from Imports in the US

We now turn to the task of evaluating the impact of the changes in the size, content, and composition of US imports for the welfare of consumers

[130] Grieco et al. (2021) also attributes the bulk of the increase in consumer surplus in the auto industry to quality improvements, while a marginal role is played by the entry of new varieties.

[131] We can use our estimation results to explore the evolution of markups and marginal cost in the US auto market. Figure 3.31 in Appendix III shows that markups (marginal cost) are increasing (decreasing) over the period 1980-2018, in line with previous work on this industry, Grieco et al. (2021).

in the United States from 1989 to 2006, as captured by the price index of US import. We first briefly outline a model of consumer demand for imports and define the corresponding price index building on the results of Section 3.2.3. We then present the results of estimating the US import demand with the DP approach and discuss the resulting measures of the change in the price index of US import.

3.4.1 Import Demand and The Import Price Index

We assume that the preferences of the representative US consumer can be characterized by a nested utility function that aggregates imported varieties into a composite import good that is consumed together with a composite domestic good. The first tier of the nested structure is given by $Q_t = \mathcal{F}_1(Q_{D,t}, Q_{M,t})$ where $Q_{D,t}$ is the composite domestically produced good, $Q_{M,t}$ is the composite imported good defined below, and where $\mathcal{F}_1(\cdot,\cdot)$ is an homothetic aggregator function that defines the consumption aggregate Q_t. In the second tier, the composite imported good $Q_{M,t}$ aggregates a vector of K sectoral imported goods $\boldsymbol{Q}_{M,t} \equiv (Q_{kt}) \in \mathbb{R}^K$ according to another homothetic aggregator $Q_{M,t} = \mathcal{F}_2(\boldsymbol{Q}_{M,t})$.

Finally, in the third tier, the composite imported good for each sector k is defined by aggregating all varieties i within that sector:

$$\sum_{i \in V_{kt}} \mathcal{K}\left(e^{\varphi_{kit}} \frac{q_{kit}}{Q_{kt}}; \varsigma_k\right) = \mathcal{K}(1), \tag{1}$$

where $\mathcal{K}(\cdot; \varsigma_k)$ is the Kimball aggregator for the varieties in sector k, q_{kit} and φ_{kit} stand for the consumption level and quality of variety i in sector k, and V_{kt} is the set of all imported varieties consumed in sector k. We follow the standard approach to identify varieties with the country of origin (Armington assumption). As for the Kimball function, we consider the standard CES aggregator and our Finite-Finite specification of the Kimball preferences in Equation (5)

Our goal is to measure the change in the relative price of imports, given by $\Delta \log P_{M,t} \equiv \log(P_{M,t}/P_{M,t-1})$. We take the price of the consumption composite Q_t to be the numeraire, and express the prices of imported goods relative to the price index of the representative US consumer. Assuming that the number of sectors remains constant over time, we can approximate the change in the unit cost of the bundle of imported goods for any homothetic aggregator $\mathcal{F}_2(\cdot)$, up to the second order, using the Trnqvist price index (Diewert, 1976, 1978, Jaravel and Lashkari, 2021):

$$\Delta \log P_{M,t} \approx \sum_k \overline{\overline{s_{kt}}} \, \Delta \log P_{kt}, \tag{2}$$

where the Trnqvist sectoral weight $\overline{\overline{s_{kt}}}$ is the average share of sector k is the total volume of import between periods $t-1$ and t.

To compute the aggregate import price index from Equation (2), we need to compute change $\Delta \log P_{kt}$ in the logarithm of the unit cost for each sector k, relying on the results of Section 3.2.3. As we discuss below, we first estimate the Kimball demand system, separately for each sector, using the technique presented in Section 3.2.2, and then use Equation (23) to approximately decompose the change in the ideal price index for each sector into the change in unadjusted unit value, quality, and variety. We also estimate the CES demand to examine the difference between the contribution of quality as inferred by the Kimball and the CES demand systems.

3.4.2 Data and Estimation

We use product-level data on US imports from 1989 to 2006 compiled by Feenstra et al. (2002). These data record US imports at the 10-digit level of the Harmonized System (henceforth HS10), reporting also the corresponding SITC classification. We define a good to be an HS10 category and we follow the standard approach to identify varieties with the country of origin, e.g., an exporter-HS10 pair. A variety's unit value is defined as the sum of the value, total duties, and transportation costs divided by the import quantity. To correctly evaluate the role of prices, we deflate import prices and expenditure using the official measure of CPI from the Bureau of Labor Statistics.[132] To minimize the effects of noise in the data, we trim the data as follows: we exclude all varieties that report a quantity of one unit or less

[132] In Appendix III we report the welfare calculations using the US producer's price index (PPI) as the price deflator. The main qualitative conclusions of our welfare analysis do not change.

than the 5th percentile within each HS10 product category; we remove varieties with an annual unit value increase that fall below the 5th percentile or above the 95th percentile within each HS10 product category.

We estimate the CES elasticity of substitution across product varieties at the HS10 level, together with the 5, 4 and 3-digit SITC levels of aggregation (SITC5, SITC4 and SITC3, respectively).[133] We use our Dynamic Panel (DP) approach using the moment condition in Equation (10) with double lagged (log) prices and market shares as instruments. We compare our estimates against those found using the conventional Feenstra (1994) and Broda and Weinstein (2006) estimator (henceforth FBW) and as well as the more recent Limited Information Maximum Likelihood estimation approach (Soderbery, 2015, henceforth LIML). We next apply the DP approach to the Finite-Finite specification of the Kimball preferences at the SITC3 level.[134] We use the moment condition in Equation (9) with lagged log prices and quantities and their quadratic power as instruments.[135]

For the purpose of estimation, we use any continuously imported variety over the period from 1989 to 2006 within each product classification

[133]The SITC4 level allows us to map our data to the Rauch product classification (Rauch, 1999).

[134]Note that the contribution of changes in the set of available varieties at more disaggregated levels, e.g., HS10, appears as quality gains at the SITC3 level. As we will discuss below, our measures of variety gains at SITC3 and HS10 levels are similar in the CES case, since the larger changes in the share of common varieties set in the more disaggregated case are mostly counteracted by the correspondingly lower love of variety (higher elasticities of substitution). We can extend our analysis to the more disaggregated levels, e.g., HS10, by focusing on shorter intervals of time over which we can define a continuously imported variety as a reference product in our estimation.

[135]In cases where the estimated values were not feasible with this set of instrument, we added the third power of both lagged log prices and quantities.

TABLE 3.2 – Comparison between DP, FBW and LIML

	HS 10			SITC 5			SITC 3		
	DP	BW	LIML	DP	BW	LIML	DP	BW	LIML
Mean	5.70	4.64	4.50	5.09	3.44	3.21	4.49	2.97	1.70
(SE)	(0.15)	(0.09)	(0.11)	(0.23)	(0.13)	(0.15)	(0.45)	(0.39)	(0.11)
Median	3.35	2.74	2.10	3.08	2.43	1.65	2.79	2.29	1.23
(SE)	(0.05)	(0.02)	(0.02)	(0.10)	(0.04)	(0.04)	(0.25)	(0.08)	(0.03)
T-statistics		7.89	8.08		6.40	6.91		2.56	6.06
Pearson χ^2 p-value		0.00	0.00		0.00	0.00		0.03	0.00
N	7283	7283	7283	1140	1140	1140	127	127	127

Note: Mean and median of the elasticities of substitution estimated with the DP, FBW and LIML methods for the HS10, SITC5 and SITC3 levels of aggregation. Only feasible estimates for common products are reported. Values above 130 are censored. Standard errors for each statistics are bootstrapped. For each level of aggregation, T-statistics refer to a t-test for differences in mean with respect to DP; p-values for Pearson difference in median tests with respect to DP.

as the baseline product to infer quality in Equation (5).[136] For computing the price index, we create a basket of OECD countries as our set of baseline products O_k for quality ($\frac{1}{|O_k|}\sum_{o\in O_k}\varphi_{ot} = 0$) within each product classification, assuming that the average quality of varieties imported from these countries are on average the same as those reflected in the US CPI. This allows us to express the quality of the varieties supplied by all other countries relative to this baseline.

3.4.3 Estimates of the Elasticitiy of Substitution

Elasticities under the CES Model Table 3.2 compares the price elasticities estimated by the different strategies across different product classifications. First, note that the magnitude of the estimated price elasticities falls as we estimate them across more aggregated varieties, as varieties become

[136] In practice, this restricts the possibility to the major advanced economies and few other exporters.

less substitutable at these more aggregated levels.[137] Comparing the magnitudes across different methods, we find that the elasticities estimated using DP are larger compared to those obtained using the FBW or LIML methods, in both mean and median terms, at all levels of aggregation. For instance, at the three-digit level, the mean elasticity for DP is 4.5, 50% greater than the number for FBW and more than twice that for LIML. Similarly, the median elasticity for DP is 2.8, while the value is 2.3 and 1.2 for the conventional methods FBW and LIML, respectively. We can easily reject the hypothesis that the means and the medians are the same.[138]

As we discussed in Section 3.2.2.2, the the FBW and LIML methods assume uncorrelated demand and supply shocks, which is likely to be violated when marginal cost depends on quality. The resulting positive correlation between demand and supply shocks should lead to a downward bias in the price elasticities estimated by the two conventional methods, consistent with the results in Table 3.2. As we will see in the following subsection, the bias in the estimates of the elasticity of substitution plays an important role in the predictions of these methods for the inferred quality gains.

Intuitively, we expect the magnitude of the price elasticities to be higher among more homogenous goods compared to more differentiated ones, since these homogenous goods should be more substitutable (Broda and

[137] Appendix I provides a more extensive discussion of this result for the DP estimates.

[138] Figure 3.37 in Appendix III shows the strong correlation among the estimates found by the three methods.

TABLE 3.3 – Kimball Elasticities

	Kimball	CES
Mean	8.82	5.17
Median	4.66	3.87
Weighted Mean	6.62	7.18
p5	1.85	1.63
p95	27.0	10.2

Note: The table reports the mean, median, and both the 5th and 95th percentiles of the distribution of price elasticities for both the Kimball and CES specifications. For the Kimball specification, we can compute the elasticity for each variety at each moment in time while, in the CES case, each variety-time pair is associated with the corresponding sectoral CES elasticity.

Weinstein, 2006). In Appendix I, we use the standard Rauch (1999) classification to distinguish products at the SITC4 level into three categories: commodities, referenced priced, and differentiated goods, and show that our estimated price elasticities are lower for more differentiated products. More interestingly, we also show that the downward bias in the FBW and LIML methods is stronger for more differentiated product categories, since quality should be more relevant for this type of products compared to more homogenous ones.

Elasticities under the Kimball Model We now turn our attention to the estimated price elasticities for the Kimball model and compare them to the corresponding CES estimates.[139] Table 3.3 compares different moments of the distribution of elasticities across varieties between Kimball and CES estimates.[140] We find larger estimates under the Kimball demand system,

[139]Table 3.24 in Appendix III reports summary statistics of the distribution of the estimated Finite-Finite Kimball parameters.

[140]Recall that for the Kimball specification, we can compute the elasticity for each variety at each moment in time while in the CES case we only compute a common value across time and varieties, within each SITC3. The moments for CES are computed assuming that

in terms of mean, median, and both lower and upper tails of the distribution. This result suggests that ignoring the heterogeneity in price elasticities across varieties leads to a bias in the estimated price elasticity at the variety level. Figure 3.4 orders all sectors from left to right based on the share-weighted mean elasticity under Kimball, reporting the estimated lower and upper limits of the Kimball specification, the expenditure share weighted Kimball elasticitiy, and the estimated CES elasticity for each SITC3. The solid black line shows that there is a strong positive correlation between the expenditure-share weighted mean Kimball elasticity and the corresponding CES elasticity.[141] However, the estimated lower and upper limits of the Finite-Finite specification show the existence of an extensive heterogeneity in the price elasticities across varieties within each sector, suggesting that the CES assumption can be a poor approximation for the degree of own-price elasticity for many individual varieties.[142]

In line with the results from the US auto market, Figure 3.5 shows that across all product codes, varieties with higher inferred quality have higher expenditure shares and lower price elasticities.

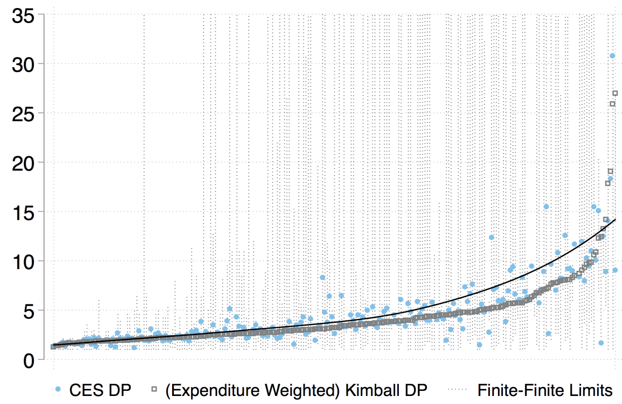
---

each variety-time pair within the same sector has the same elasticity.

[141]The CES elasticities reported in Figure 3.4 are estimated using CES as the limiting case of the Kimball specification ($\sigma_0 \equiv \sigma$). Figure 3.38 in Appendix III shows that there is almost a perfect match between the estimates obtained using the limiting Kimball moment and the moment conditions in first-differences used for elasticities reported in Table 3.21.

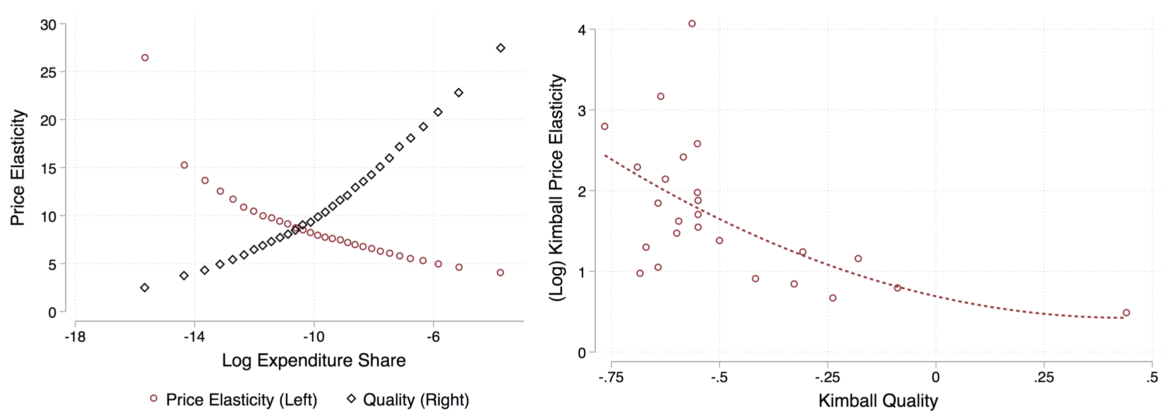
[142]Figure 3.39 in Appendix III illustrates the extent of the heterogeneity in elasticities for the Watches and Clocks sector (SITC3 number 884). The figure reports the entire set of Kimball elasticities, their expenditure-share weighted mean, and the CES estimate. Even if the expenditure-weighted mean Kimball elasticity is very close to the CES estimate (4.02 compared to 4.69), the Kimball prices elasticities range from 2 to 15 and decrease with market share.

FIGURE 3.4 – Comparison with CES Elasticities



Note: In the figure we rank each SITC3 sector by the expenditure-share weighted mean Kimball price elasticity. For each sector, it display the estimated lower and upper limits of the Finite-Finite Kimball specification (dotted line), the expenditure-share weighted mean Kimball price elasticity (gray squares) and the corresponding CES estimate (blue circles). The upper limits are truncated at 35. The solid black line shows a fitted curve through the CES estimates.

FIGURE 3.5 – Kimball Price Elasticities and Implied Quality



Note: The left panel plots the binscattered relationship between (log) expenditure share of each variety-time observation and the Kimball price elasticity (left axis) and product quality (right axis). The right panel directly plots the relationship between product quality and price elasticity.

3.4.4 The Evolution of the US Import Price Index

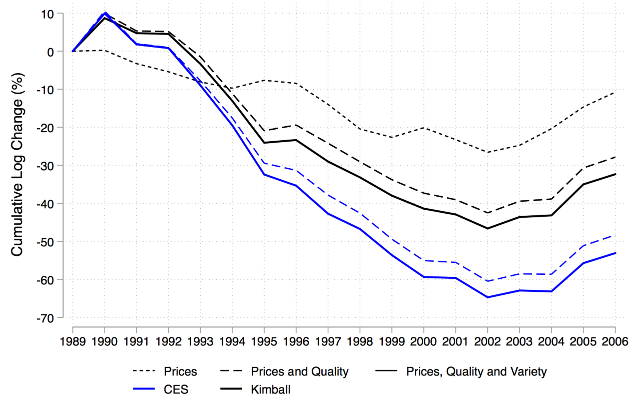
Figure 3.6 reports the cumulative change in the aggregate price of US imports relative to the CPI from Equation (2), where the changes in the sector-level Kimball price indices are approximated using the expression in Equation (23). The figure also provides a decomposition of the change in the aggregate index to the three sources of interest. Improved product quality constitutes the primary source of consumption gains from openness in the US, accounting for more than half of the total decline in relative import prices. The import price index declined by around 32% (1.80% annually) relative to the CPI over the 1989-2006 period. A price index including only changes in unadjusted prices would find the cumulative decline in the aggregate import price index over the period to be around 11%.[143] Figure 3.6 and Table 3.4 also show that the impact of new varieties is marginal compared to the role of quality improvement, accounting for a 4.5% cumulative (0.25% annually) drop in the aggregate import price index. Standard price indices would therefore largely underestimate the overall decline in import prices.[144]

Using CES preferences instead of Kimball doubles the consumption gains arising from the product quality channel, leading to a sizable overestimation of the overall gains. The CES aggregate price index for imports shows a decline of around 53% (2.95% annually), 30% more than the Kimball case. The stark difference with respect to the Kimball aggregate price index arises mainly from the different estimates of the role of quality upgrading. Whereas quality improvement reduces the CES aggregate

[143] Figure 3.40 in Appendix III shows that the year-to-year change in the price component of our aggregate import price index strongly resembles the Import Price Index constructed by the BLS.

[144] In Appendix III, Figure 3.41 and Table 3.25 show the change in the price index of imports and its decomposition when the prices are stated relative to the US PPI. In this case, the unadjusted import prices in fact slightly rise over time and almost all of the fall in the import price index is explained by quality improvements.

FIGURE 3.6 – Dynamics of US Import Price Index



Note: The figure plots the aggregate import price indices for both the CES and Kimball case and their decomposition into the price, quality and variety components, according Equations (19) and (23). Prices are deflated using the CPI index from BLS. The measure of inferred quality is normalized such that the average quality of the set of OECD varieties is zero. The solid lines represent the aggregate import price index including all three components. The dashed and dotted lines represent the price and quality components together and the price component only, respectively. Black (Blue) lines refer to the Kimball (CES) specification.

import price by 37.5%, the corresponding contribution using Kimball is only 17%. Table 3.4 shows that under the CES model the impact of new varieties is still marginal but larger than that suggested by the Kimball specification. This confirms the quantitative importance of departing from the constant elasticity assumption in the standard CES demand systems for evaluating the consumption gains from trade, and in particular the role of product quality.

To better understand the drivers of the gap in the contribution of quality implied by CES and Kimball, Proposition 2 in Appendix III provides a decomposition of this gap to a number of different components. Appendix II.1 uses this decomposition to show that the key reason for the overestimation of the contribution of quality under the CES specification is simply that the corresponding estimated elasticities suffer from a downward bias.

TABLE 3.4 – Change in the Import Price Index in the US (Relative to CPI, 1989–2006)

	Total		Decomposition				
			Price	Quality		Variety	
	Kimball	CES		Kimball	CES	Kimball	CES
Cumulative Change (%)	-32.3	-53.1	-10.8	-17.1	-37.5	-4.48	-4.76
Annual Change (%)	-1.80	-2.95	-0.60	-0.95	-2.09	-0.25	-0.26

Note: The table reports the cumulative and average annual change in the aggregate import price indices defined in Equations (19) and (23) and reported in Figure 3.6, and their decomposition. Prices are deflated using the CPI index from BLS. The measure of inferred quality is normalized such that the average quality of the set of OECD varieties is zero.

The above results show that, although quantitatively less relevant than the role of quality upgrading, the contribution of variety in Lemma 2 also depends on the demand system used to evaluate it. The gains from varieties in the presence of heterogenous demand elasticities are smaller mainly because the index of love of variety, when adjusted for contribution of heterogeneity in demand elasticities, $\overline{\overline{\mu}}^{*}_{k,t}$, is typically smaller than in the CES case, $\mu_k \equiv \frac{1}{\sigma_k - 1}$. Once again, this result is driven by the lower estimates of the price elasticities under the CES case, which leads to an overestimation of the contribution of variety.

3.4.5 *Decomposing Quality Change across Exporters*

We now focus our attention on the main source of consumption gains, quality upgrading, and decomponse the aggregate quality change to the contributions of major exporters to the US, distinguishing China, the OECD economies, and all other exporters.

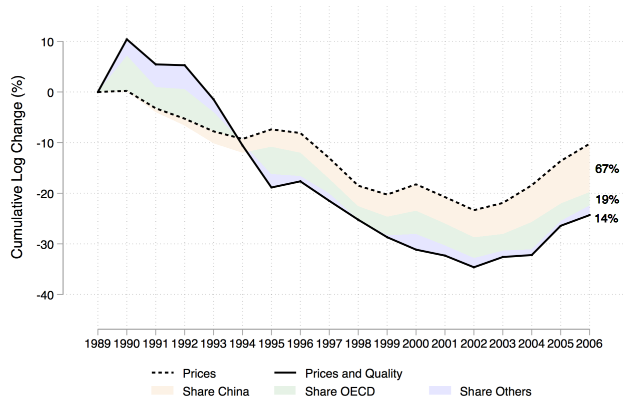
Figure 3.7 shows that about 70% of the total cumulative gains from quality can be attributed to quality improvements of Chinese varieties relative to the baseline , i.e. the average quality across OECD varieties.[145] The contribution of the OECD countries and all the other exporters to the overall quality improvement is about 20% and 14%, respectively.[146] Chinese products represent the largest source of quality improvements and, ultimately, gains from trade experienced by the US. This result is in line with the prior work documenting that the expansion of Chinese exports is not limited to the low-skill labor intensive and low-quality goods (Hsieh and Ossa, 2016). Figure 3.7 further shows that the quality upgrading the quality upgrading already in progress in the 90s but accelerates after China's accession to the WTO. This result is consistent with the fact that the path of economic reforms in China goes further back in time to the late 70s (Brandt et al., 2017, Fan et al., 2015, 2017), and with recent evidence for the substantial effect of the China's entry into the WTO on US prices (Amiti et al., 2020).[147]

[145]Notice that the normalization used to evaluate quality does not imply that the contribution of quality changes of the OECD countries is zero. The contribution of quality change among OECD varieties is the Tornqvist weighted mean of variety-level quality change, while our baseline sets the unweighted mean quality among the OECD varieties to zero.

[146]Figure 3.42 in Appendix III shows the same decomposition for the CES case. Chinese varieties still represent the major source of quality improvements, accounting for 46% of the aggregate quality improvement. OECD and other exporters' varieties account for the 28% and the 26% of the aggregate quality improvement, respectively. Departing from the constant elasticity assumption is important not only in evaluating the aggregate role of quality for the gains from trade, but also in decomposing its sources.

[147]This result is also consistent with the evidence of the effects of trade liberalization on firm performance. Prior work has documented that a reduction in (input and output) tariffs spurs innovation, productivity and product quality (see Shu and Steinwender (2019) for a survey, and see, among others, Brandt et al., 2017, Fan et al., 2015, Hsieh and Ossa, 2016 for discussions of the specific Chinese case). Schott (2008) show that, even if unit values in product-level US import data are higher for advanced economies, Chinese products undertook a rapid process of sophistication. See Appendix I for further discussion.

FIGURE 3.7 – Decomposition of Quality across Countries



Note: The dashed line shows the price component of the aggregate import price index. The solid line shows the price component together with the quality component of the aggregate import price index. The quality contribution is computed using the inferred quality from the Kimball specification. The difference between these two lines quantifies the role of quality changes and is decomposed into the role of Chinese varieties (orange area), OECD varieties (green area) and all other varieties pooled together (purple area).

3.5 Conclusion

In this paper, we examined the role of quality improvements for the consumption gains from globalization in the context of the changes in the size and composition of US imports over the 1989-2006 period. We implemented a novel methodology to infer quality changes in a flexible demand model using only data on prices and market shares, and derived an approximate decomposition of the changes in the relative price of imports into the contribtuions of changes in prices, quality, and the variety in the set of available products. Moreover, we independently validated our approach in the context of the US auto market in which additional information on product chacteristics is available. Our baseline results suggest that, over the period from 1989 to 2006, quality improvements accounted for more than half of gains from trade in the US and 70% of these gains arise from the improvement in the quality of Chinese products. By ignoring the heterogeneity in price elasticities, the gains from quality are largely overestimated, indicating the importance of departing from the standard CES assumption in our accounting of the role of quality. Applying our novel methodology to other economies, as well as to firm-level data to include pro-competitive effects and their interaction with quality, are promising venues for future research.

Bibliography

Acemoglu, D., Akcigit, U. and Kerr, W. (2016), 'Networks and the macroeconomy: An empirical exploration', *Nber macroeconomics annual* **30**(1), 273–335.

Ackerberg, D. A. (2016), Timing Assumptions and Efficiency: Empirical Evidence in a Production Function Context.

Ackerberg, D. A., Caves, K. and Frazer, G. (2015), 'Identification properties of recent production function estimators', *Econometrica* **83**(6), 2411–2451.

Adao, R., Pomeranz, D., Donaldson, D., Adão, R., Carrillo, P., Arnaud, C., Donaldson, D. and Pomeranz, D. (2022), 'Imports, Exports, and Earnings Inequality: Measures of Exposure and Estimates of Incidence', *Quarterly Journal of Economics* .

Admati, A. R. (1985), 'A noisy rational expectations equilibrium for multi-asset securities markets', *Econometrica: Journal of the Econometric Society* pp. 629–657.

Alessandria, G. (2009), 'Consumer search, price dispersion, and international relative price fluctuations', *International Economic Review* **50**(3), 803–829.

Alviarez, V., Fioretti, M., Kikkawa, K. and Morlacco, M. (2021), Two-sided market power in firm-to-firm trade, Technical report, Working Paper No. IDB-DP-884, Department of Research and Chief Economist.

Amiti, M., Dai, M., Feenstra, R. C. and Romalis, J. (2020), 'How did China's WTO entry affect U.S. prices?', *Journal of International Economics* **126**(71973013), 103339.
URL: *https://doi.org/10.1016/j.jinteco.2020.103339*

Amiti, M., Itskhoki, O. and Konings, J. (2014), 'Importers, exporters, and exchange rate disconnect', *American Economic Review* **104**(7), 1942–78.

Amiti, M., Itskhoki, O. and Konings, J. (2019), 'International shocks, variable markups, and domestic prices', *The Review of Economic Studies* **86**(6), 2356–2402.

Antràs, P. (2020), De-globalisation? global value chains in the post-covid-19 age, Technical report, National Bureau of Economic Research.

Arkolakis, C., Costinot, A., Donaldson, D. and Rodriguez-clare, A. (2017), The elusive of pro-competitive effect of trade.

Aruoba, S. B., Fernández, A., Guzmán, D., Pastén, E. and Saffie, F. (2022), 'Pricing under distress'.

Atkeson, A. and Burstein, A. (2008), 'Pricing-to-market, trade costs, and international relative prices', *American Economic Review* **98**(5), 1998–2031.

Atkin, D., Chaudhry, A., Chaudry, S., Khandelwal, A. K. and Verhoogen, E. (2015), 'Markup and Cost Dispersion across Firms: Direct Evidence from Producer Surveys in Pakistan', *American Economic Review: Papers & Proceedings* **105**(5), 537–544.
URL: *http://pubs.aeaweb.org/doi/10.1257/aer.p20151050*

Auer, R. A., Levchenko, A. A. and Sauré, P. (2019), 'International inflation spillovers through input linkages', *Review of Economics and Statistics* **101**(3), 507–521.

Autor, D., Dorn, D., Katz, L. F., Patterson, C. and Van Reenen, J. (2020), 'The fall of the labor share and the rise of superstar firms', *The Quarterly Journal of Economics* **135**(2), 645–709.

Bacchetta, P. and Van Wincoop, E. (2006), 'Can information heterogeneity explain the exchange rate determination puzzle?', *American Economic Review* **96**(3), 552–576.

Bacchetta, P. and Van Wincoop, E. (2010), 'Infrequent portfolio decisions: A solution to the forward discount puzzle', *American Economic Review* **100**(3), 870–904.

Bacchetta, P. and Van Wincoop, E. (2019), Puzzling exchange rate dynamics and delayed portfolio adjustment, Technical report, National Bureau of Economic Research.

Backus, D. K. and Smith, G. W. (1993), 'Consumption and real exchange rates in dynamic economies with non-traded goods', *Journal of International Economics* **35**(3-4), 297–316.

Baldwin, R. and Harrigan, J. (2011), 'Zeros, quality, and space: Trade theory and trade evidence', *American Economic Journal: Microeconomics* **3**(May), 60–88.

Baqaee, D. and Burstein, A. (2022), 'Welfare and Output with Income Effects

and Taste Shocks', *NBER Working Paper* (1947611), No. 28754.
URL: *http://www.nber.org/papers/w28754*

Baqaee, D. R. and Farhi, E. (2020), 'Productivity and misallocation in general equilibrium', *The Quarterly Journal of Economics* **135**(1), 105–163.

Basu, S. (1994), 'Intermediate goods and business cycles: Implications for productivity and welfare'.

Benigno, G. and Benigno, P. (2003), 'Price stability in open economies', *The Review of Economic Studies* **70**(4), 743–764.

Berlingieri, G., Breinlich, H. and Dhingra, S. (2018), 'The impact of trade agreements on consumer welfare—evidence from the eu common external trade policy', *Journal of the European Economic Association* **16**(6), 1881–1928.

Berman, N., Martin, P. and Mayer, T. (2012), 'How do different exporters react to exchange rate changes?', *The Quarterly Journal of Economics* **127**(1), 437–492.

Bernanke, B. S. and Woodford, M. (2005), *The inflation-targeting debate*, Vol. 32, University of Chicago Press.

Berry, S., Levinsohn, J. and Pakes, A. (1995), 'Automobile prices in market equilibrium', *Econometrica: Journal of the Econometric Society* pp. 841–890.

Berry, S. T. (1994), 'Estimating Discrete-Choice Models of Product Differentiation Authors', *The RAND Journal of Economics* **25**(2), 242–262.

Berry, S. T., Gandhi, A. and Haile, P. A. (2013), 'Connected Substitutes and Invertibility of Demand', *Econometrica* **81**(5), 2087–2111.

Borusyak, K. and Jaravel, X. (2018), The Distributional Effects of Trade: Theory and Evidence from the United States.

Boskin, M. J., Dulberger, E. R., Gordon, R. J., Griliches, Z. and Jorgenson, D. W. (1998), 'Consumer Prices, the Consumer Price Index, and the Cost of Living', *Journal of Economic Perspectives* **12**(1), 3–26.

Brandt, L., Van Biesebroeck, J., Wang, L. and Zhang, Y. (2017), 'Wto accession and performance of chinese manufacturing firms', *American Economic Review* **107**(9), 2784–2820.

Broda, C. and Weinstein, D. E. (2006), 'Globalization and the gains from variety', *The Quarterly journal of economics* **121**(2), 541–585.

Burstein, A. and Gopinath, G. (2014), International prices and exchange rates, *in* 'Handbook of international economics', Vol. 4, Elsevier, pp. 391–451.

Burstein, A. T., Neves, J. C. and Rebelo, S. (2003), 'Distribution costs and real exchange rate dynamics during exchange-rate-based stabilizations', *Journal of monetary Economics* **50**(6), 1189–1214.

Caliendo, L., Mion, G., Opromolla, L. D. and Rossi-Hansberg, E. (2020), 'Productivity and organization in portuguese firms', *Journal of Political Economy* **128**(11), 4211–4257.

Calvo, G. A. and Reinhart, C. M. (2002), 'Fear of floating', *The Quarterly journal of economics* **117**(2), 379–408.

Camatte, H., Daudin, G., Faubert, V., Lalliard, A. and Rifflart, C. (2021), 'Global value chains and the transmission of exchange rate shocks to consumer prices'.

Campa, J. M. and Goldberg, L. S. (2005), 'Exchange rate pass-through into import prices', *Review of Economics and Statistics* **87**(4), 679–690.

Carvalho, V. M. and Tahbaz-Salehi, A. (2019), 'Production networks: A primer', *Annual Review of Economics* **11**, 635–663.

Chen, N., Chung, W. and Novy, D. (2022), 'Vehicle currency pricing and exchange rate pass-through', *Journal of the European Economic Association* **20**(1), 312–351.

Chen, N. and Juvenal, L. (2016), 'Quality, trade, and exchange rate pass-through', *Journal of International Economics* **100**, 61–80.

Chen, S., Devereux, M. B., Shi, K. and Xu, J. (2022), Consumption heterogeneity and monetary policy in an open economy, Technical report, National Bureau of Economic Research.

Conlon, C. and Gortmaker, J. (2020), 'Best practices for differentiated products demand estimation with pyblp', *The RAND Journal of Economics* **51**(4), 1108–1161.

Corsetti, G. and Dedola, L. (2005), 'A macroeconomic model of international price discrimination', *Journal of International Economics* **67**(1), 129–155.

Corsetti, G., Dedola, L. and Leduc, S. (2008), 'International risk sharing and the transmission of productivity shocks', *The Review of Economic Studies* **75**(2), 443–473.

Corsetti, G., Dedola, L. and Leduc, S. (2010), Optimal monetary policy in open economies, *in* 'Handbook of monetary economics', Vol. 3, Elsevier, pp. 861–933.

Costinot, A. and Rodríguez-Clare, A. (2015), Trade Theory with Numbers: Quantifying the Consequences of Globalization, *in* 'Handbook of International Economics', Vol. 4, Elsevier B.V., pp. 197–261.
URL: *http://dx.doi.org/10.1016/B978-0-444-54314-1.00004-5*

Cravino, J. and Levchenko, A. A. (2017*a*), 'The distributional consequences of large devaluations', *American Economic Review* **107**(11), 3477–3509.

Cravino, J. and Levchenko, A. A. (2017*b*), 'Multinational firms and international business cycle transmission', *The Quarterly Journal of Economics* **132**(2), 921–962.

Crozet, M., Head, K. and Mayer, T. (2012), 'Quality sorting and trade: Firm-level evidence for french wine', *The Review of Economic Studies* **79**(2), 609–644.

Dasgupta, P. and Stiglitz, J. (1988), 'Learning-by-doing, market structure and industrial and trade policies', *Oxford Economic Papers* **40**(2), 246–268.

De Loecker, J., Eeckhout, J. and Unger, G. (2020), 'The rise of market power and the macroeconomic implications', *The Quarterly Journal of Economics* **135**(2), 561–644.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K. and Pavcnik, N. (2016), 'Prices, markups, and trade reform', *Econometrica* **84**(2), 445–510.

De Loecker, J. and Warzynski, F. (2012), 'Markups and firm-level export status', *American economic review* **102**(6), 2437–71.

Deaton, A. and Muellbauer, J. (1980), 'An almost ideal demand system', *American Economic Review* **70**(3), 312–326.
URL: *http://www.jstor.org/stable/1805222*

Dhyne, E., Kikkawa, A. K., Mogstad, M. and Tintelnot, F. (2021), 'Trade and domestic production networks', *The Review of Economic Studies* **88**(2), 643–668.

Di Giovanni, J., Kalemli-Ozcan, Silva, A. and Yildirim, M. A. (2022), Global supply chain pressures, international trade, and inflation, Technical report, National Bureau of Economic Research.

Di Giovanni, J. and Levchenko, A. A. (2010), 'Putting the parts together:

trade, vertical linkages, and business cycle comovement', *American Economic Journal: Macroeconomics* **2**(2), 95–124.

Di Giovanni, J., Levchenko, A. A. and Mejean, I. (2017), 'Large firms and international business cycle comovement', *American Economic Review* **107**(5), 598–602.

Diewert, W. E. (1976), 'Exact and superlative index numbers', *Journal of Econometrics* **4**(2), 115–145.

Diewert, W. E. (1978), 'Superlative Index Numbers and Consistency in Aggregation', *Econometrica* **46**(4), 883–900.

Dingel, J. I. (2017), 'The determinants of quality specialization', *The Review of Economic Studies* **84**(4), 1551–1582.

Drozd, L. A. and Nosal, J. B. (2012), 'Understanding international prices: Customers as capital', *American Economic Review* **102**(1), 364–95.

Eaton, J., Eslava, M., Jinkins, D., Krizan, C. J. and Tybout, J. R. (2021), A search and learning model of export dynamics, Technical report, National Bureau of Economic Research.

Eaton, J. and Fieler, A. C. (2022), The Margins of Trade.
URL: *http://www.sas.upenn.edu/ afieler/EF_gravity_prices.pdf*

Edmond, C., Midrigan, V. and Xu, D. Y. (2015), 'Competition, markups, and the gains from international trade', *American Economic Review* **105**(10), 3183–3221.

Edmond, C., Midrigan, V. and Xu, D. Y. (2018), How costly are markups?, Technical report, National Bureau of Economic Research.

Engel, C. (2016), 'Exchange rates, interest rates, and the risk premium', *American Economic Review* **106**(2), 436–74.

Engel, C. and West, K. D. (2005), 'Exchange rates and fundamentals', *Journal of political Economy* **113**(3), 485–517.

Errico, M. (2022), 'Importer's dynamics and exchange rate pass-through'.

Errico, M. and Lashkari, D. (2022), 'The quality of us imports and the consumption gains from globalization'.

Evans, M. D. and Lyons, R. K. (2002), 'Order flow and exchange rate dynamics', *Journal of political economy* **110**(1), 170–180.

Fajgelbaum, P., Grossman, G. M. and Helpman, E. (2011), 'Income distribution, product quality, and international trade', *Journal of political Economy* **119**(4), 721–765.

Fama, E. F. (1984), 'Forward and spot exchange rates', *Journal of monetary economics* **14**(3), 319–338.

Fan, H., Li, Y. A., Xu, S. and Yeaple, S. R. (2017), Quality, variable markups, and welfare: a quantitative general equilibrium analysis of export prices.

Fan, H., Li, Y. A. and Yeaple, S. R. (2015), 'Trade liberalization, quality, and export prices', *Review of Economics and Statistics* **97**(5), 1033–1051.

Feenstra, R. C. (1994), 'New product varieties and the measurement of international prices', *The American Economic Review* pp. 157–177.

Feenstra, R. C. (2003), 'A homothetic utility function for monopolistic competition models, without constant price elasticity', *Economics Letters* **78**(1), 79–86.

Feenstra, R. C. (2010), *Product variety and the gains from international trade*, MIT Press, Cambridge, MA.

Feenstra, R. C., Romalis, J. and Schott, P. K. (2002), 'Us imports, exports, and tariff data, 1989-2001'.

Feenstra, R. C. and Weinstein, D. E. (2017), 'Globalization, markups, and us welfare', *Journal of Political Economy* **125**(4), 1040–1074.

Flam, H. and Helpman, E. (1987), 'Vertical product differentiation and advertising', *American Economic Review* **77**(5), 810–822.
URL: *http://eprints.lancs.ac.uk/19876/1/IJEBrevpap.doc*

Foley, C. (2021), Flexible Entry / Exit Adjustment for Price Indices.

Foucault, T., Pagano, M., Roell, A. and Röell, A. (2013), *Market liquidity: theory, evidence, and policy*, Oxford University Press.

Gabaix, X. and Maggiori, M. (2015), 'International liquidity and exchange rate dynamics', *The Quarterly Journal of Economics* **130**(3), 1369–1420.

Gábor-Tóth, E. and Vermeulen, P. (2018), The Relative Importance of Taste Shocks and Price Movements in the Variation of Cost-of-Living: Evidence From Scanner Data.

Garcia-Marin, A., Justel, S. and Schmidt-Eisenlohr, T. (2019), 'Trade credit, markups, and relationships'.

Georgiadis, G., Gräb, J. and Khalil, M. (2020), 'Global value chain participation and exchange rate pass-through', *Available at SSRN 3767294* .

Goldberg, L. S. and Campa, J. M. (2010), 'The sensitivity of the cpi to exchange rates: Distribution margins, imported inputs, and trade exposure', *The Review of Economics and Statistics* **92**(2), 392–407.

Goldberg, P. K. and Hellerstein, R. (2008), 'A structural approach to explaining incomplete exchange-rate pass-through and pricing-to-market', *American Economic Review* **98**(2), 423–29.

Goldberg, P. K. and Verboven, F. (2001), 'The evolution of price dispersion in the european car market', *The Review of Economic Studies* **68**(4), 811–848.

Gopinath, G. (2015), The international price system, Technical report, National Bureau of Economic Research.

Gopinath, G., Boz, E., Casas, C., Díez, F. J., Gourinchas, P.-O. and Plagborg-Møller, M. (2020), 'Dominant currency paradigm', *American Economic Review* **110**(3), 677–719.

Gopinath, G. and Itskhoki, O. (2011), 'In search of real rigidities', *NBER Macroeconomics Annual* **25**(1), 261–310.

Gopinath, G., Itskhoki, O. and Rigobon, R. (2010), 'Currency choice and exchange rate pass-through', *American Economic Review* **100**(1), 304–36.

Gordon, R. J. and Griliches, Z. (1997), 'Quality change and new products', *The American Economic Review* **87**(2), 84–88.

Grennan, M. (2013), 'Price discrimination and bargaining: Empirical evidence from medical devices', *American Economic Review* **103**(1), 145–177.

Grieco, P. L., Murry, C. and Yurukoglu, A. (2021), The evolution of market power in the us auto industry, Technical report, National Bureau of Economic Research.

Hallak, J. C. (2006), 'Product quality and the direction of trade', *Journal of International Economics* **68**(1), 238–265.

Hallak, J. C. and Schott, P. K. (2011), 'Estimating cross-country differences in product quality', *The Quarterly journal of economics* **126**(1), 417–474.

Head, K. and Mayer, T. (2021), Poor Substitutes? Counterfactual Methods in IO and Trade Compared.
URL: *https://drive.google.com/file/d/1FBJS_Y9kds6WzhF9GlNGKglh5ouH5j-j/view*

Heise, S. (2019), 'Firm-to-firm relationships and the pass-through of shocks: Theory and evidence', *FRB of New York Staff Report* (896).

Heise, S., Karahan, F. and Şahin, A. (2022), 'The missing inflation puzzle: The role of the wage-price pass-through', *Journal of Money, Credit and Banking* **54**(S1), 7–51.

Hillberry, R. and Hummels, D. (2013), Trade elasticity parameters for a computable general equilibrium model, *in* 'Handbook of computable general equilibrium modeling', Vol. 1, Elsevier, pp. 1213–1269.

Hsieh, C.-T., Li, N., Ossa, R. and Yang, M.-J. (2020), 'Accounting for the new gains from trade liberalization', *Journal of International Economics* **127**, 103370.

Hsieh, C.-T. and Ossa, R. (2016), 'A global view of productivity growth in china', *Journal of international Economics* **102**, 209–224.

Hulten, C. R. (1978), 'Growth accounting with intermediate inputs', *The Review of Economic Studies* **45**(3), 511–518.

Hummels, D. and Klenow, P. J. (2005), 'The variety and quality of a nation's exports', *American economic review* **95**(3), 704–723.

Hummels, D. and Skiba, A. (2004), 'Shipping the good apples out? An empirical confirmation of the Alchian-Allen conjecture', *Journal of Political Economy* **112**(6), 1384–1402.

Huneeus, F. (2018), 'Production network dynamics and the propagation of shocks', *Graduate thesis, Princeton University, Princeton, NJ* .

Jaravel, X. (2021), 'Inflation inequality: Measurement, causes, and policy implications', *Annual Review of Economics* **13**, 599–629.

Jaravel, X. and Lashkari, D. (2021), Nonparametric Measurement of Long-Run Growth in Consumer Welfare.

Juarez, L. (2022), 'Buyer market power and exchange rate pass-through'.

Kacperczyk, M. T., Nosal, J. B. and Sundaresan, S. (2018), 'Market power and price informativeness', *Available at SSRN 3137803* .

Kaplan, S. B. A. H. G. and Zoch, P. (2020), 'Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data'.

Kasahara, H. and Sugita, Y. (2021), Nonparametric Identification of Production Function, Total Factor Productivity, and Markup from Revenue Data.

Khandelwal, A. (2010), 'The long and short (of) quality ladders', *The Review of Economic Studies* **77**(4), 1450–1476.

Kimball, M. S. (1995), 'The quantitative analytics of the basic neomonetarist model'.

Klenow, P. J. and Willis, J. L. (2006), Real Rigidities and Nominal Price Changes.

Klenow, P. J. and Willis, J. L. (2016), 'Real rigidities and nominal price changes', *Economica* **83**(331), 443–472.

Kugler, M. and Verhoogen, E. (2012), 'Prices, plant size, and product quality', *Review of Economic Studies* **79**(1), 307–339.

Kyle, A. S. (1985), 'Continuous auctions and insider trading', *Econometrica* pp. 1315–1335.

Kyle, A. S. (1989), 'Informed speculation with imperfect competition', *The Review of Economic Studies* **56**(3), 317–355.

La'O, J. and Tahbaz-Salehi, A. (2022), 'Optimal monetary policy in production networks', *Econometrica* **90**(3), 1295–1336.

Leamer, E. E. (1981), 'Is it a Demand Curve, Or is It A Supply Curve? Partial Identification through Inequality Constraints', *Review of Economics and Statistics* **63**(3), 319–327.

Lee, R. S. (2013), 'Vertical integration and exclusivity in platform and two-sided markets', *American Economic Review* **103**(7), 2960–3000.

Levinsohn, J. and Petrin, A. (2003), 'Estimating production functions using inputs to control for unobservables', *The review of economic studies* **70**(2), 317–341.

Linder, S. B. (1961), *An Essay on Trade and Transformation*, Almqvist & Wiksells, Uppsala.

Lyons, R. K. et al. (2001), *The microstructure approach to exchange rates*, Vol. 333, Citeseer.

Mahajan, S. (2018), *Handbook on Supply and Use Tables and Input Output-Tables with Extensions and Applications*, United Nations Publications, Department of Economic and Social Affairs, Statistics Division.

Manova, K. and Zhang, Z. (2012), 'Export prices across firms and destinations', *Quarterly Journal of Economics* **127**(1), 379–436.

Martin, J. and Mejean, I. (2014), 'Low-wage country competition and the quality content of high-wage country exports', *Journal of International Eco-*

nomics **93**(1), 140–152.
URL: *http://dx.doi.org/10.1016/j.jinteco.2014.02.002*

Matsuyama, K. (2022), Non-CES aggregators: a guided tour.

Matsuyama, K. and Ushchev, P. (2017), Beyond CES: Three Alternative Classes of Flexible Homothetic Demand Systems.

Matsuyama, K. and Ushchev, P. (2022), Selection and Sorting of Heterogeneous Firms through Competitive Pressures.

McFadden, D. (1974), Conditional logit analysis of qualitative choice behavior, *in* P. Zarembka, ed., 'Frontiers in Econometrics', Academic Press, pp. 105—-142.

Meese, R. A. and Rogoff, K. (1983), 'Empirical exchange rate models of the seventies: Do they fit out of sample?', *Journal of international economics* **14**(1-2), 3–24.

Miller, R. E. and Blair, P. D. (2009), *Input-output analysis: foundations and extensions*, Cambridge university press.

Mishkin, F. S. (2008), Exchange rate pass-through and monetary policy, Technical report, National Bureau of Economic Research.

Mussa, M. (1982), 'A model of exchange rate dynamics', *Journal of political economy* **90**(1), 74–104.

Nakamura, E. and Steinsson, J. (2008), 'Five facts about prices: A reevaluation of menu cost models', *The Quarterly Journal of Economics* **123**(4), 1415–1464.

Neiman, B. (2010), 'Stickiness, synchronization, and passthrough in intrafirm trade prices', *Journal of Monetary Economics* **57**(3), 295–308.

Nimark, K. (2008), 'Dynamic pricing and imperfect common knowledge', *Journal of monetary Economics* **55**(2), 365–382.

Nimark, K. (2017), 'Dynamic higher order expectations'.

Obstfeld, M. and Rogoff, K. (2000), 'The six major puzzles in international macroeconomics: is there a common cause?', *NBER macroeconomics annual* **15**, 339–390.

Pasten, E., Schoenle, R. and Weber, M. (2020), 'The propagation of monetary policy shocks in a heterogeneous production economy', *Journal of Monetary Economics* **116**, 1–22.

Piveteau, P. (2021), 'An empirical dynamic model of trade with consumer accumulation', *American Economic Journal: Microeconomics* **13**(4), 23–63.

Rauch, J. E. (1999), 'Networks versus markets in international trade', *Journal of international Economics* **48**(1), 7–35.

Redding, S. J. and Weinstein, D. E. (2020), 'Measuring aggregate price indices with taste shocks: Theory and evidence for ces preferences', *The Quarterly Journal of Economics* **135**(1), 503–560.

Redding, S. and Weinstein, D. (2021), Accounting for trade patterns.

Reinert, K. A. and Roland-Holst, D. W. (1992), 'Armington elasticities for united states manufacturing sectors', *Journal of Policy Modeling* **14**(5), 631–639.

Rubbo, E. (2020), 'Networks, phillips curves, and monetary policy', *manuscript, Harvard University* .

Ruhl, K. J. et al. (2008), 'The international elasticity puzzle'.

Schott, P. K. (2004), 'Across-product versus within-product specialization in international trade', *The Quarterly Journal of Economics* **119**(2), 647–678.

Schott, P. K. (2008), 'The relative sophistication of chinese exports', *Economic policy* **23**(53), 6–49.

Shu, P. and Steinwender, C. (2019), 'The impact of trade liberalization on firm productivity and innovation. innovation policy and the economy, vol. 19'.

Singleton, K. (1987), Speculation and the volatility of foreign currency exchange rates, *in* 'Carnegie-Rochester Conference Series on Public Policy', Vol. 26, Elsevier, pp. 9–56.

Soderbery, A. (2015), 'Estimating import supply and demand elasticities: Analysis and implications', *Journal of International Economics* **96**(1), 1–17.

Sweeting, A. (2013), 'Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry', *Econometrica* **81**(5), 1763–1803.

Ueda, K., Watanabe, K. and Watanabe, T. (2019), 'Product Turnover and the Cost-of-Living Index: Quality versus Fashion Effects', *American Economic Journal: Macroeconomics* **11**(2), 310–347.

Verhoogen, E. A. (2008), 'Trade, quality upgrading, and wage inequality in the mexican manufacturing sector', *The Quarterly Journal of Economics* **123**(2), 489–530.

Wang, O. and Werning, I. (2020), Dynamic oligopoly and price stickiness, Technical report, National Bureau of Economic Research.

Appendices

Appendix A Data and Calibration

I Chilean Customs Data and Importing Tenure

For each import transaction, the Chilean Customs dataset includes standard information such as the importer's unique identifier (*importer*), the 8-digit HS product code (*product*), the date of the transaction, the country of origin (*origin*), the FOB and CIF values, the quantity shipped, etc. Data are available from 2009 to 2020. I compute prices as unit values by dividing the shipment value by the quantity shipped. To improve the reliability of the data, I trim the dataset by dropping observations whose price changes are above (below) the 99th (1st) percentile. Additional data cleaning entails the removal of all transactions with missing information, e.g. quantity, value, etc. I aggregate all transactions at the importer-origin-product-quarterly level, by summing over values and quantities. Table 3.6 provides summary statistics of the main variables and Table 3.5 reports information on industry and origin composition of the data.

TABLE 3.5 – Summary Statistics - Breakdown by Industry and Origin

	Numbers of Transactions (%)	Import Value (%)
Industry (SITC):		
Food & Animals	3.871	8.238
Beverages, Tobacco	0.291	0.613
Crude Materials	1.589	2.392
Mineral fuels	0.503	24.34
Animal & Vegetable Oils	0.192	0.524
Chemicals	11.55	13.23
Manufactured Goods	18.57	9.466
Machinery	36.52	33.02
Mix Manufacturing	26.91	8.160
Country:		
China	14.02	6.208
USA	25.01	30.00
EU15	25.41	17.41
Other Americas	18.42	25.03
Others	17.14	21.35

The table reports the breakdown by industry (2-digit SITC level) and country of origin of the cleaned universe of import transactions from the Chilean Customs, 2009-2019. The breakdown is computed in terms of i) number of transactions and ii) import values.

II Construction of IO Matrix and Distribution Costs

I construct the input-output matrix for the Chilean economy combining the 2013 "make" and "use" tables provided by the the Central Bank of Chile (*Banco Central de Chile*).[148] I combine the make and use tables to construct a product-by-product input-output matrix that quantifies how much of each product is used in the production of other products. I choose to construct a product-by-product matrix, rather than an industry-by-industry, to leverage the larger product dimension of the make and use tables.

I follow standard best practice in Mahajan (2018) and Miller and Blair

[148] The most recent version of the tables provided by the Central Bank of Chile is from 2013. Data are available at the following website: `https://si3.bcentral.cl/estadisticas/Principal1/Excel/CCNN/cdr/excel.html`.

TABLE 3.6 – Description of the Data - Customs

	Whole Sample				
	Mean	Median	StD	p5	p95
Importers	41,186	.	.	.	.
Products	7,518	.	.	.	.
Origin Countries	168	.	.	.	.
Products per importer	10.66	3	27.28	1	43
Origins per importer	2.227	1	2.931	1	7
Unit value (USD/quantity)	1,732.7	21.35	76,930.6	0.934	1,569.2
% Δ log unit value	0.446	0.417	0.690	-116.6	118.1
Transaction value (USD)	130,817.5	7,214.3	2,659,917.9	239.5	286,991.7
Observations (N)	3,044,931	.	.	.	.

The table reports summary statistics of the cleaned universe of import transactions from the Chilean Customs, 2009-2019. Transaction values and unit values are defined in USD.

(2009) in constructing the input-output table under the industry technology assumption. Consider the product-by-industry make matrix, V^T, and the product-by-industry use matrices of domestic and imported products U_d and U_m, respectively. Define g^T the row vector of industry output, i.e. the column sum of V^T. I construct the product-mix matrix C,

$$C = V^T \left[\texttt{diag}(g^T)\right]^{-1},$$

that collects the share of each product in the output of an industry. Under the industry technology assumption, each industry has its own specific way of production, irrespective of its product mix.[149] I obtain the domestic and international Leontief matrices by multiplying the product-mix matrix C to the use matrices U_d and U_m:

$$S_d = U_d C^T \qquad S_m = U_m C^T,$$

[149]Compared to the most common alternative assumption (product technology assumption), the key advantage of the industry technology assumption is that negative elements in the input-output table cannot arise.

TABLE 3.7 – Summary Statistics - Importing Tenure

	p5	p25	Median	p75	p95	Mean	N
Observation:							
Importer X Time	1	1	4	11	44	11.8	965,043
Importer X Product X Time	1	1	1	1	2	1.19	9,524,237
Importer X Country X Time	1	1	2	5	18	4.94	2,299,882
Tenure:							
Main	1	1	1	2	6	1.99	2,391,689
Alternative	1	1	1	3	13	3.26	2,391,689

The table reports summary statistics on the distribution of the number of observations along different dimension (importer, time, product and country) from the cleaned universe of import transactions from the Chilean Customs, 2009-2019. The table reports summary statistics on importing tenure, defined as: i) the number of quarters the importer has been consecutively importing a Product X Origin pair (main); ii) the number of quarters the importer has been importing a Product X Origin pair (alternative).

where S_d and S_m represent the domestic and international product-by-product Leontief matrices, respectively. The left (right) panel of Figure 1.2 plots the domestic (international) Leontief matrix. As expected, the matrices are highly sparse given the granularity of the product classification used.

III Markup Elasticity

In this section, I provide additional information on how markup elasticities are estimated and calibrated. In the main text, I assume that the Kimball aggregator in Equation (7) takes the form of a Klenow and Willis (2016) aggregator. In this case, the firm-level markup elasticity depends on two parameters, the industry-specific elasticity of demand, σ_i, and the

TABLE 3.8 – Distribution Margins - Summary Statistics

	Intermediate Goods		Final Goods	
	Domestic	Imported	Domestic	Imported
Farms	0.0701	0.0778	0.258	0.183
Fishing and Forestry	0.0135	0.000166	0.113	0.0224
Oil, Coal and Gas Extraction	0.0000500	0.0236	0	0
Mining	0.000593	0.0216	0	0
Food, Beverages and Tobacco	0.0896	0.207	0.265	0.366
Textile and Apparel	0.128	0.248	0.342	0.529
Wood, Paper and Printing	0.103	0.142	0.181	0.257
Petroleum and Chemical Products	0.150	0.172	0.307	0.386
Plastic Rubber and Construction	0.0580	0.146	0.146	0.401
Fabricated Metal Products	0.0577	0.133	0.0309	0.0809
Machinery and Equipment	0.0918	0.194	0.134	0.336
Motor Vehicles	0.0335	0.0988	0.0744	0.333
Furniture	0.112	0.225	0.312	0.369
Utilities	0.0310	0.000800	0.106	0
Construction	0.00269	0	0	0
Wholesale and Retail Trade	0.00384	0.00180	0.0229	0
Transportation	0.0107	0.00803	0.0183	0
Health Care and Education	0.00190	0	0.0250	0
Accomodation and Recreation	0.0381	0.0216	0.0894	0
Professional Services	0.0208	0.0157	0.0525	0.0226
Communication	0.0451	0.0153	0.149	0
Other Products or Services	0.0908	0.0701	0.0391	0.118

The table reports the average distribution margin for each (2-digit) industry. I distinguish across products depending on their use, final vs intermediate use, and on their origin, imported vs domestically produced.

super-elasticity of demand, ϵ_i, as follows:

$$\Gamma_{ik} = \frac{\epsilon_i}{\sigma_i - 1 + \epsilon_i \log\left(\frac{\widetilde{p_{ik}}}{\widetilde{p}_i}\right)}. \tag{1}$$

I follow Gopinath et al. (2010) and Amiti et al. (2019) and calibrate the steady-state elasticity of markups, assuming $\widetilde{p_{ik}} = \widetilde{p}_i$.[150] I calibrate the demand elasticity parameter σ to match the average, steady-state markup. I then follow Edmond et al. (2018) in estimating the superelasticity parameter ϵ using the firm-level relationship between markups and market shares

[150]Under the condition $\widetilde{p_{ik}} = \widetilde{p}_i$, Equation (1) can be interpreted as the markup elasticity for an average firm (Amiti et al., 2019) or at the steady-state markup elasticity (Gopinath et al., 2010).

implied by the Klenow and Willis (2016) function form of Equation (7).

I now provide details on how I estimate markups and markup elasticities to calibrate the model in Section 1.2.

ENIA Data: I use the Annual National Industrial Survey (ENIA) from 2000 to 2007, that provides information for approximately 5000 plants per year with more than 10 employees. It reports detailed information on sales, inputs expenditures, employment and wage bill, investment, industry code (ISIC rev 3). I consider the following variables: `REMPAG` as wage bill; `EMPTOT` and `THHANO` as total number of employees and total hours worked, respectively; `VSTK` as capital stock; `FABVAL` as production value; `VBPB` as gross production value and `VA` as value added; the sum of `TCOVAL` and `MTMPVAL` as total material expenditure; `ELECONS` as electricity consumption in MW. Table 3.9 presents a few basic summary statistics for the leading variables used in the analysis.

TABLE 3.9 – Description of the Data - ENIA

	Mean	p25	Median	p75
Sales	5,666,147	151,802	407,989	1,607,334
Wage Bill	438,828.1	37,268	88,067	279,700
Material Expenditure	3,067,797	74,545	209,090	866,560
Capital Stock	3,001,394	31,636	130,379	620,612
Electricity Used (MW)	3,520.978	27	77	357
Observations	31,027			

The table reports summary statistics of the cleaned ENIA dataset from 2000 to 2007. All variables but electricity consumption are in millions of Chilean pesos.

I drop firms that have zero or negative employees, wage bill, production, material expenditure or electricity usage, and capital stock. I also drop observations for which i) the gross value of production is lower than the total value added; ii) the wage bill is larger than the total value added. To obtain a real measure of the main nominal variables, I use deflators provided by the Central Bank of Chile or the National Statistical Agency (*INE*). Production value is deflated using industry-specific deflators; the value of capital stock is deflated by the investment good deflator; wage bill is deflated by the domestic CPI and material expenditure by industry-specific producer price indices.

Markup estimation & σ_i: I use production function estimation to estimate markups at the three-digit ISIC industry level following state-of-the-art techniques and best practices, Levinsohn and Petrin (2003), Ackerberg et al. (2015) and De Loecker and Warzynski (2012).

As specified in the theoretical model in Section 1.2, I estimate a Cobb-Douglas production function of the form:

$$\log y_{ik} = \beta_i^k \log k_{ik} + \beta_i^l \log l_{ik} + \beta_i^x \log x_{ik} + \omega_{ik} + \xi_{ik} \tag{2}$$

where y_{ik}, k_{ik}, l_{ik}, x_{ik}, ω_{ik} and ξ_{ik} represent quantity sold, capital stock, labor, materials, log productivity and the error term, respectively. I follow the control function literature, Levinsohn and Petrin (2003) and Ackerberg et al. (2015), to tackle the endogeneity challenge due to unobserved time-varying firm-level productivity ω_{ik} and consistently estimate the production function in Equation (2).

I treat capital as a dynamic input that faces adjustment costs. I use the consumption of electricity in megawatts as proxy variable. I favor a composite variable of the cost of goods sold as benchmark measure for variable input. I construct this variable summing the total cost of labor (wage bill) to the total expenditure in materials.[151]

Given the estimated output elasticities, markups are constructed following De Loecker and Warzynski (2012); hence, firm-level markups are given by:

$$\mu_{ik} = \widehat{\beta_i^{\texttt{Cost}}} \frac{\texttt{Sales}_{ik}}{\texttt{Cost}_{ik}} \tag{3}$$

where $\texttt{Cost}_{ik}$ is the sum of wage bill and material expenditure and $\beta_i^{\texttt{Cost}}$ is the associated output elasticity estimated from Equation (2). For each industry, I calibrate the industry-specific demand elasticity using the estimated revenue-weighted average markup $\bar{\mu}_i$, $\sigma_i = \frac{\bar{\mu}_i}{\bar{\mu}_i - 1}$.

Estimating Kimball Super-elasticity ϵ_i: The Klenow and Willis (2016) functional form of the Kimball aggregator implies the following within-industry relationship between markups and market shares, up to a

[151]Using this measure as variable input implicitly imposes an additional assumption in the estimation, as it assumes that labor and materials are perfectly substitutable, De Loecker et al. (2020). As robustness, in the section below, I relax this assumption, treating labor costs and materials separately and using the former to estimate markups. Markups and markup elasticities are highly correlated to the one I obtain from my preferred specification.

constant:

$$\frac{1}{\mu_{ik}} + \log\left(1 - \frac{1}{\mu_{ik}}\right) = a_i + b_i \log \texttt{share}_{ik}, \qquad b_i = \frac{\epsilon_i}{\sigma_i},$$

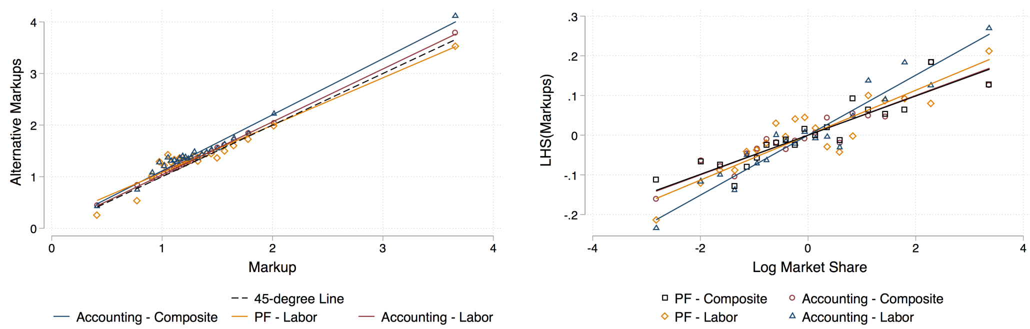
where $\texttt{share}_{ik}$ is the market share of firm k in industry i. I estimate the slope coefficient b_i for each industry introducing firm and year fixed effects. I can then retrieve the sectoral super-elasticities ϵ_i given the estimated demand elasticity σ_i.

Robustness: It is well known that standard production data, as those used here, report revenues and expenditures rather than physical units. The standard practise of deflating using sectoral indices can introduce an additional bias due to unobserved firm-specific input price variation, De Loecker et al. (2016).[152] Moreover, recent work by Kaplan and Zoch (2020) shows that it is not possible to consistently estimate output elasticities when only revenue data is available in the presence of variable markups.

To assess the robustness of the estimates from my preferred specification, I compute markups using the simple alternative cost share approach (Autor et al., 2020, De Loecker et al., 2016). Under constant return to scale, the output elasticity of each input is equal to the share of that input in total costs. I assume that the output elasticity is common to all firms within each industry and I calibrate it to the median input share in each industry. I also

[152]Without more detailed data on output prices and quantities, it is not possible to implement the control function approach proposed by De Loecker et al. (2016) to tackle the input price bias.

FIGURE 3.8 – Comparison with Alternative Markup Estimates



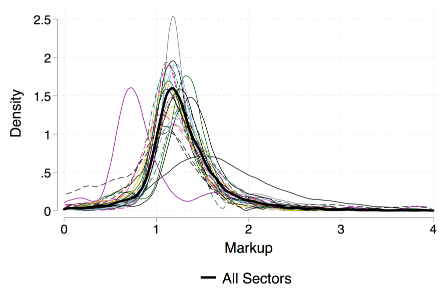
The left panel plots the relationship between the preferred measure of markups (x-axis) and the alternative measures of markups estimated as robustness (y-axis). The preferred measure of markups is estimated using production function estimation and a composite measure of cost of goods sold as variable input. Alternative measures of markups include: i) estimates using production function estimation and labor as variable input ("PF - Labor"); ii) estimates using the cost share approach and a composite measure of cost of goods sold as variable input ("Accounting - Composite"); iii) estimates using the cost share approach and labor as variable input ("Accounting - Labor"). The right panel shows the relationship between the log market share of a firm and the left-hand-side of Equation (3), $\frac{1}{\mu_{ik}} + \log\left(1 - \frac{1}{\mu_{ik}}\right)$, where μ_{ik} is the firm-level markup. I consider both the preferred measure of markups ("PF - Composite") and the alternative measures estimated as robustness ("PF - Labor", "Accounting - Composite" and "Accounting - Labor"). I use the whole sample and include both year and industry fixed effects.

relax the assumption of a composite variable input used in my preferred specification. I re-estimate markups and markup elasticities treating labor and materials separately using both the production function and the cost share approaches.

The left panel of Figure 3.8 plots the alternative estimates of markups against the markups obtained from the preferred specification. The right panel of Figure 3.8 shows the relationship in Equation (3) between (log) market share and markups, using the whole sample and controlling for year and industry fixed effects. Overall, these additional estimates show qualitative and quantitative patterns that are similar to the benchmark specification. Markup distributions are very similar, independently of the approach or variable input used. Similarly, the estimated super-elasticities (the slope coefficient on the right panel of Figure 3.8) are very close.

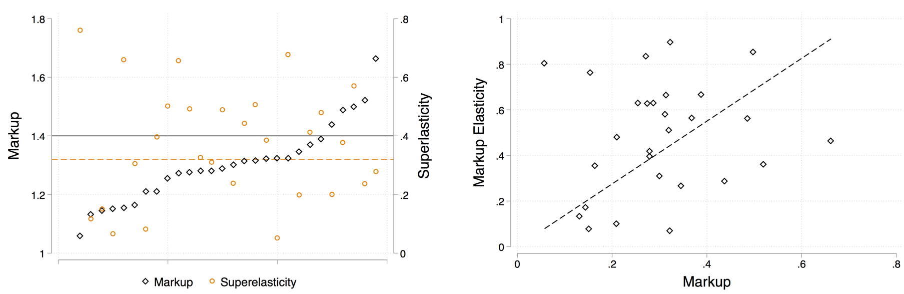
III.1 Additional tables and figures

FIGURE 3.9 – Markup Distributions



The figure plots the distribution of estimated markups for each 3-digit ISIC industry. The thick solid black line represents the aggregate distribution pooling all industries together. Markups are estimated using the preferred specification, i.e. production function estimation and a composite measure of cost of goods sold as variable input.

FIGURE 3.10 – Markup Elasticity and Super-elasticity



In the left panel I rank each 3-digit ISIC industry by the estimated revenue-weighted average markup. For each industry I plot the estimated revenue-weighted average markup and the corresponding estimated demand super-elasticity, ϵ_i. The solid horizontal line shows the aggregate revenue-weighted average markup in the whole sample. The right panel shows the relationship between the estimated revenue-weighted average markup (x-axis) and the implied markup elasticity at the 3-digit ISIC industry level. The dashed line shows a linear fit through the implied markup elasticities. Markups are estimated using the preferred specification, i.e. production function estimation and a composite measure of cost of goods sold as variable input. Markup elasticity is defined according to Equation (2), where σ_i is calibrated using the revenue-weighted average markup and ϵ_i is estimated using Equation (3).

IV Pass-through $\Psi(T)$

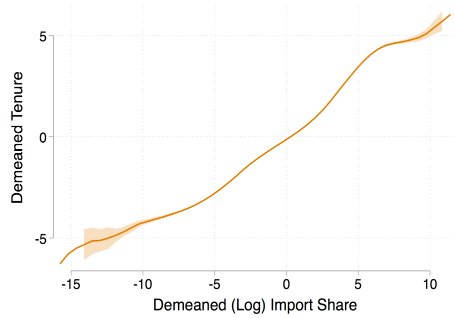
TABLE 3.10 – Estimated Average Pass-through

	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \log e$	0.7149	0.8324	0.7759	0.7092	0.8118	0.7641
	(0.107)	(0.105)	(0.103)	(0.111)	(0.107)	(0.105)
Log Tenure X $\Delta \log e$		-0.0816			-0.0727	
		(0.014)			(0.015)	
Tenure X $\Delta \log e$			-0.0109			-0.0100
			(0.002)			(0.002)
Importer X Product X Country	Yes	Yes	Yes			
Importer X Product				Yes	Yes	Yes
Product X Country				Yes	Yes	Yes
Observations	2,368,422	2,368,422	2,368,422	2,413,107	2,413,107	2,413,107

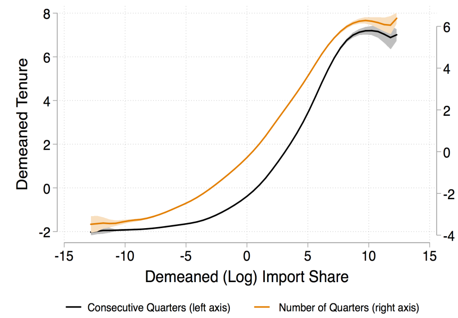
The table reports the estimated coefficients from the specification in Equation (4) without the set of controls included, X, and time fixed effects, ν_t. Columns (1) and (4) do not control for the effect of importing tenure. Columns (2) and (5) ((3) and (6)) control for the interaction between exchange rate change and the log (level) of importing tenure. Columns (1), (2) and (3) ((4), (5) and (6)) include Import X Product X Country (Importer X Product and Product X Country) fixed effects. Coefficients for variables in level (log importing tenure, importing tenure and inflation of origin country) and left and right censorship dummies are omitted. Standard errors clustered at country level. Importing tenure is defined as the number of quarters the importer has been consecutively importing a Product X Origin pair.

Appendix B Importing Tenure: Robustness

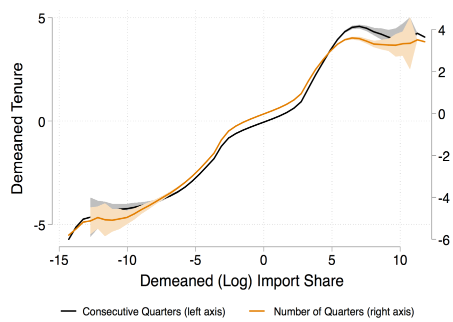
FIGURE 3.11 – Heterogeneity in Tenure - Robustness



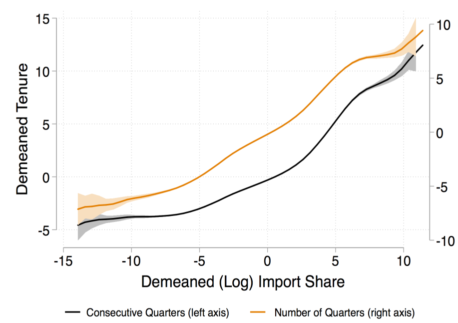
(A) Alternative measure of tenure



(B) Demeaning at quarter level



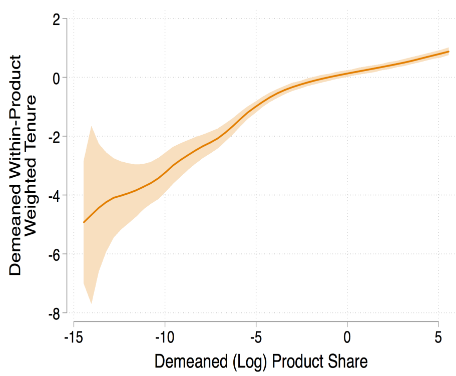
(C) Demeaning at quarter-firm-sector level



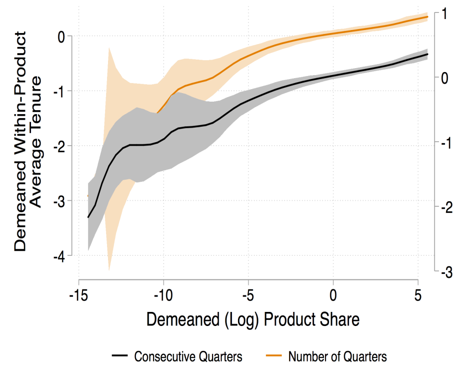
(D) Second half of the sample

All figures plot the non-parametric relationship between the (log) import share and importing tenure in the whole sample. Share and tenure are computed at the quarterly level. Import shares and tenure are defined at the firm-product-origin-quarter level, with product defined at the 8-digit level. Variables are demeaned to avoid mechanical increase in tenure due to time passing and make it comparable over time. Panel a) uses the alternative definition of tenure, the number of quarters a firm is importing the same product-origin pair (dropping the consecutive requirement of the main definition). Panel b) uses both definitions of tenure but demeans variables at the quarterly level only. Similarly, panel c) plots the variables demeaned at the quarterly-firm-sector level, where sector is defined at the 3-digit level. Finally panel d) shows the relationship between the (log) import share and tenure in the second half of the sample, using both definitions of tenure. In all panels, I report the 99% confidence intervals.

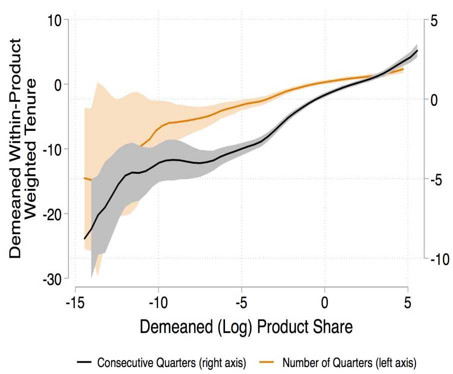
FIGURE 3.12 – Heterogeneity in Tenure at Product Level - Robustness



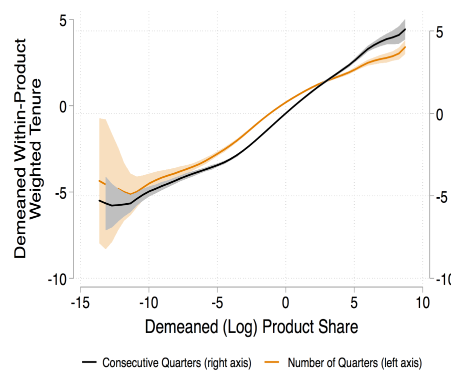
(A) Alternative measure of tenure



(B) Average tenure



(C) Second half of the sample



(D) 5-digit classification

All figures plot the non-parametric, cross-sectional relationship between the (log) import share of a product and the average tenure across all firms importing that product. Share and average tenure are computed at the quarterly level. Variables are demeaned to avoid mechanical increase in tenure due to time passing and make it comparable over time. Panel a) computes the expenditure-weighted tenure using the alternative definition of tenure, the number of quarters a firm is importing the same product-origin pair (dropping the consecutive requirement of the main definition). Panel b) computes the average tenure, considering both the main (left) and the alternative (right) definition of tenure. Panel c) plots the relationship between the (log) import share of a product and the expenditure-weighted average tenure across all firms importing that product using only the second half of the sample. Finally panel d) defines products at the 5-digit level. In all panels, I report the 99% confidence intervals.

TABLE 3.11 – Pass-through Robustness

	Level	Alternative Tenure		Alternative FEs	Alternative Own Size		Alternative Strategic	
		Cum Quarters	Cum Sales		Trans Value	Importer Size		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta \log e$	0.3591	0.4100	0.3359	0.3734	0.3383	0.3906	0.3904	0.4168
	(0.110)	(0.127)	(0.106)	(0.122)	(0.115)	(0.125)	(0.115)	(0.125)
Log Tenure X $\Delta \log e$		-0.0334	-0.0154	-0.0305	-0.0409	-0.0391	-0.0348	-0.0357
		(0.020)	(0.007)	(0.014)	(0.014)	(0.015)	(0.012)	(0.013)
Tenure X $\Delta \log e$	- 0.00375							
	(0.0017)							
Size X $\Delta \log e$	-0.0097	-0.0104	-0.0102	-0.0093	-0.5117	-0.0032	-0.0097	-0.0109
	(0.003)	(0.004)	(0.004)	(0.003)	(0.146)	(0.008)	(0.003)	(0.004)
Strategic Δp_i	0.2664	0.2524	0.2871	0.3001	0.2950	0.3019	-0.0980	-0.3127
	(0.312)	(0.312)	(0.317)	(0.271)	(0.303)	(0.295)	(0.112)	(0.130)
Time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Importer X Product X Country	Yes	Yes	Yes		Yes	Yes	Yes	Yes
Importer X Product				Yes				
Product X Country				Yes				
Observations	2,365,619	2,365,619	2,365,619	2,410,260	2,365,619	2,365,619	2,365,619	2,314,387

Coefficients for terms in levels (log tenure, tenure, average size and inflation of origin country) and left and right censorship dummies are omitted. Standard errors clustered at country level. All columns re-runs the baseline specification in Equation (4) using different controls. Column (1) reports the main specification from column (5) in Table 1.2 using tenure in levels, instead of log. Column (2) is estimated using an alternative definition of tenure, the number of quarters a firm is importing the same product-origin pair (dropping the consecutive requirement of the main definition). Column (3) defines tenure as the cumulative sum of sales at the product-origin pair. Column (4) uses Product X Origin and Product X Importer fixed effects. Column (5) controls for the actual value of the transaction in the quarter, as alternative measure of own-size. Similarly column (6) uses the size of the importer defined as the sum of the all imports across products and origins. Column (7) computes the index of competitor price change using expenditure weights. Finally, column (8) specifies the index of competitor price change at the Product X Origin level.

Appendix C Additional Figures and Tables

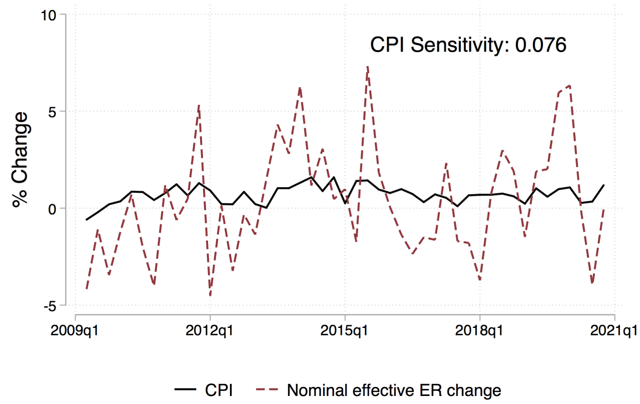
Estimating average CPI sensitivity: I estimate the aggregate CPI sensitivity for the period 2009-2020 at the quarterly level using the following specification:

$$\Delta \log CPI_t = \sum_{\tau=0}^{6} \beta_\tau \Delta \log e_{t-\tau} + \sum_{\tau=0}^{6} \gamma_\tau \pi_{t-\tau} + \varepsilon_t, \tag{1}$$

where CPI is the Chilean consumer price index at the quarterly level; e is the trade-weighted nominal exchange rate between the Chilean peso and the exporting country's currency; π is the trade-weighted inflation rate in the exporting country as proxy for trading partners' costs (Campa and Goldberg, 2005, Burstein and Gopinath, 2014). I include up to 6 lags to control for gradual adjustments and auto-correlation in inflation and exchange rates. Inflation and exchange rate data are sourced from IMF and Datastream, respectively. Figure 3.13 shows the relationship between the change in domestic prices (CPI) and the trade-weighted measure of nominal exchange rate. The estimated contemporaneous, short-run CPI sensitivity from Equation (1) is 7.6%, in line with estimates from the literature (Goldberg and Campa, 2010). The coefficient is robust to the number of lags included and to the inclusion of lagged domestic CPI as additional control.

Estimating CPI trends: I estimate the trend in aggregate short-run CPI sensitivity over the period 2009-2020 using the regression in Equation (1)

FIGURE 3.13 – Estimated CPI Sensitivity

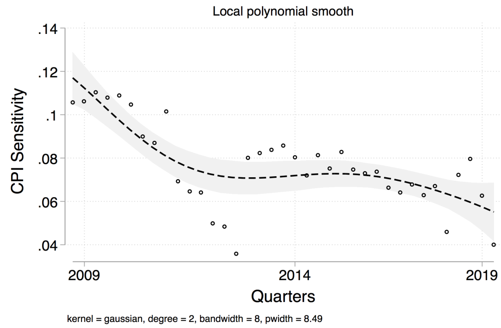


The figure plots the relationship between the change in domestic prices (CPI) and the trade-weighted measure of nominal exchange rate. Inflation and exchange rate data are sourced from IMF and Datastream, respectively. Trade shares are computed from the universe of import transactions from 2009 to 2020. The coefficient reported is the contemporaneous CPI sensitivity estimated from Equation (1) in Appendix C.

with a rolling time window of five years (20 quarters). I extend the sample to the beginning of 2007 so that the mid-point of the initial window is approximately 2009. Differently from Equation (1), I include lags up to one year as the number of data points in each window is reduced. I then estimate the trend using a polynomial approximation given that the CPI sensitivity is moderately noisy at quarterly level. Figure 3.14 plots the estimated CPI sensitivities and the corresponding downward trend.

Figure 3.15 plots the trend in short-run CPI sensitivity over the period from the late 1970s to 2020 using a rolling time window of ten years (40 quarters). Given the longer horizon considered, I augment the regression in Equation (1) to also control for the growth rate in real GDP of the importing

FIGURE 3.14 – Trend in CPI Sensitivity



The figure plots the estimated trend in short-run CPI sensitivity for Chile over the period from the late 2007 to 2020s. I use a 20-quarter rolling time window and plot the estimated trend at the midpoint of the window. CPI sensitivity is estimated at the quarterly level using regression in Equation (1). Appendix C provides additional details on the data and estimation. The trend is computed using a Gaussian polynomial approximation with bandwidth 8 and degree two. Shaded area plot the 95% confidence intervals.

country, Chile, and its lagged values (Campa and Goldberg, 2005):

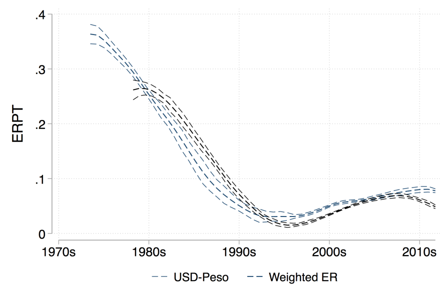
$$\Delta \log CPI_t = \sum_{\tau=0}^{6} \beta_\tau \Delta \log e_{t-\tau} + \sum_{\tau=0}^{6} \gamma_\tau \pi_{t-\tau} + \sum_{\tau=0}^{6} \eta_\tau \Delta \mathrm{GDP}_{t-\tau} + \varepsilon_t. \qquad (2)$$

In this case, the trade-weighted nominal exchange rate is downloaded directly from the IMF (series "NEU" from International Financial Statistics). I use the real effective exchange rate in combination to the nominal effective exchange rate from the IMF ("REU" and "NEU" respectively) to compute a trade-weighted measure of exporters' costs.[153] As robustness, I consider the bilateral USD-CLP exchange rate and the US inflation rate as proxy for exporters' costs.[154] I again estimate the trend using a polynomial approximation given that the CPI sensitivity is moderately noisy at quarterly level.

[153] I follow Campa and Goldberg (2005) and construct the proxy for exporters' cost, π, taking advantage of both the real and nominal exchange rate series. I compute $\pi = NEER \times CPI/REER$, where CPI is the measure of domestic prices in Chile.

[154] Using these alternative series allows to extend the period of analysis back to 1975.

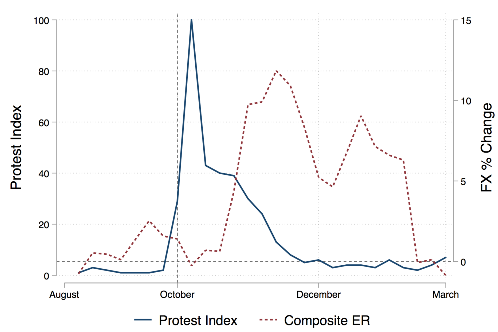
FIGURE 3.15 – Long-Run Trend in CPI Sensitivity



The figure plots the estimated long-run trend in short-run CPI sensitivity for Chile over the period from the late 1970s to 2020s. I use a 40-quarter rolling time window and plot the estimated trend at the midpoint of the window. CPI sensitivity is estimated at the quarterly level using regression in Equation (2). I use a trade-weighted exchange rate and exporters' costs series from the IMF International Financial Statistics ("Weighted ER"). As robustness, I also consider the bilateral USD-CLP exchange rate and the US inflation rate as cost proxy ("USD-Peso"). The trend is computed using an Epanechnikov polynomial approximation with bandwidth 15 and degree one. Dashed lines plot the 95% confidence intervals.

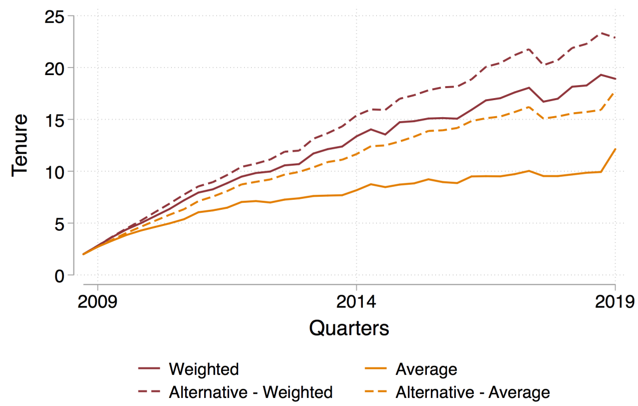
Figure 3.15 shows that sensitivity decreased since the late 1970s and the pattern is robust to the exchange rate series considered.

FIGURE 3.16 – Riot Index and Exchange Rate Dynamics



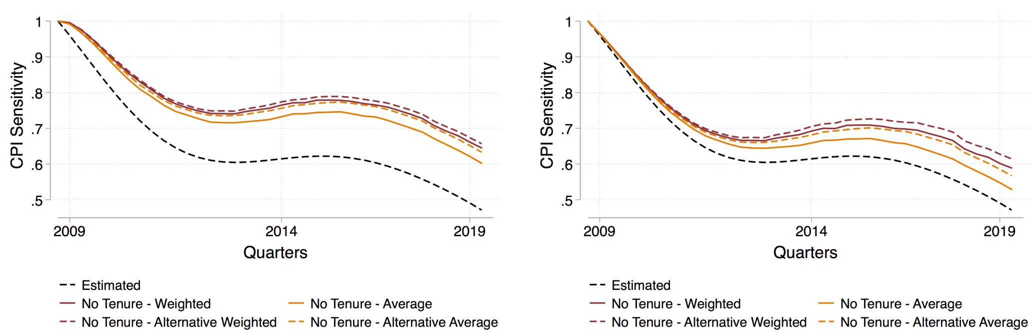
The figure plots, on the left axis, the daily Google search index for protests ("protestas" in Spanish) in Chile. The value is normalized so that the maximum over the time period considered is set equal to 100. On the right axis, I plot the weekly 3-month depreciation rate of the Chilean peso against a composite index of foreign currencies. The composite index of foreign currency is sourced from the Central Bank of Chile and is constructed as a trade-weighted average of bilateral exchange rates.

FIGURE 3.17 – Trends in Tenure



The figure plots the trend in average tenure from 2009 to 2019 from the universe of import transactions. Solid lines use the main definition of tenure, i.e. the number of consecutive quarters a firm is importing the same product-country pair. Dashed lines ("Alternative") use the less conservative measure of tenure, the number of quarters a firm is importing the same product-country pair (dropping the consecutive requirement of the main definition). Red lines compute average tenure as the expenditure-weighted average tenure across all importer-product-origin triples. Orange lines compute average tenure as a simple average across importer-product-origin triples.

FIGURE 3.18 – Trends in Tenure & CPI Sensitivity



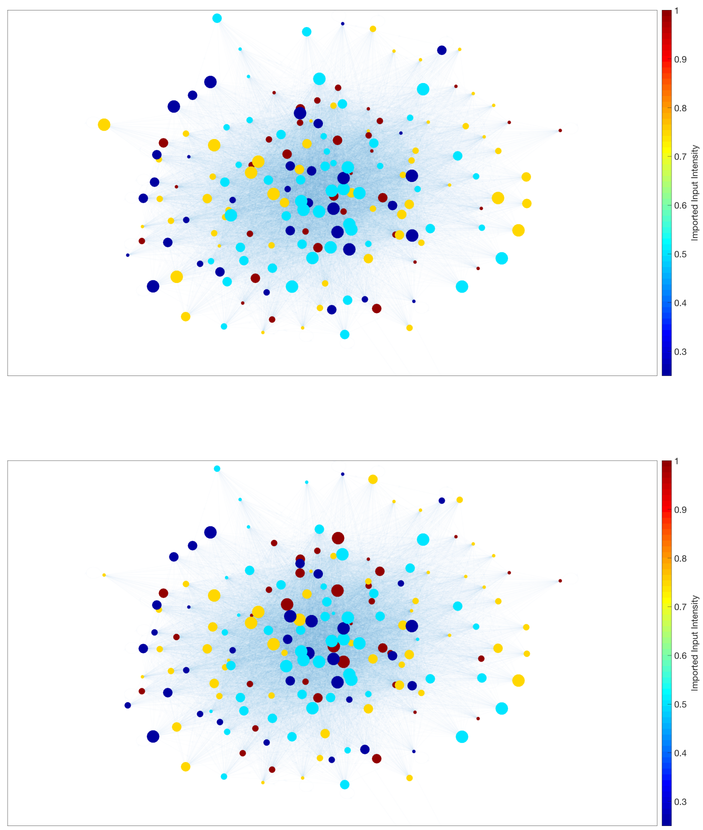
The left panel plots the counterfactual trends in CPI sensitivity using different definitions of average tenure. The measure of average tenure are described in Figure 3.17. The trends are computed using the estimated effect of importing tenure from Table 1.2 (i.e. in logs). The right panel plots the counterfactual trends in CPI sensitivity using the same definitions of average tenure used in the left panel. Differently from the left panel, the trends are computed using the estimated effect of tenure from column (1) in Table 3.11 (i.e. in levels). In both panels the black, dash line plots the trend in CPI sensitivity to exchange rate estimated as explained in Appendix C.

TABLE 3.12 – CPI Sensitivity w/out IO linkages

	Tenure Heterogeneity		No Tenure Heterogeneity	
	IO	w/out IO	IO	w/out IO
	(1)	(2)	(3)	(4)
Frictionless:				
Domestic	8.65	4.69	10.5	5.67
Imported	9.55	9.55	11.5	11.5
Total	18.2	14.2	22.0	17.2
Distribution Only:				
Domestic	6.27	3.59	7.63	4.35
Imported	5.62	5.62	6.76	6.76
Total	11.9	9.21	14.4	11.1
Distribution & Markups:				
Domestic	3.50	2.43	4.26	2.95
Imported	5.62	5.62	6.76	6.76
Total	9.12	8.05	11.0	9.71
All Frictions:				
Domestic	1.98	1.58	2.40	1.92
Imported	5.62	5.62	6.76	6.76
Total	7.60	7.20	9.17	8.68

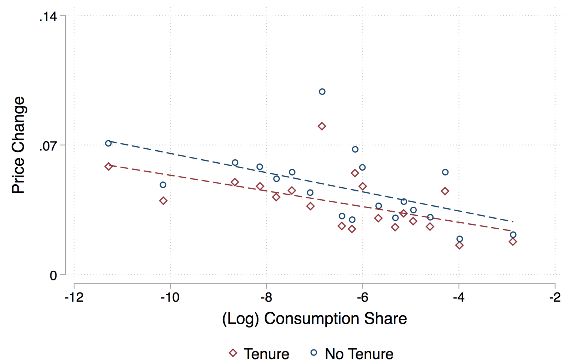
The table reports the implied aggregate CPI sensitivity to exchange rates ("Total") and its decomposition into imported final consumption ("Imported"), i.e. direct exposure, and domestic final consumption ("Domestic"), i.e. indirect exposure. I consider four different scenarios in terms of domestic frictions (distribution margin, markup elasticity, and Calvo rigidity). From top to bottom, I consider a domestic economy with: no frictions; distribution costs only; distribution costs and markup elasticity; all frictions together. In all scenarios, pass-through into import prices is incomplete. Columns (1) and (2) (columns (3) and (4)) include (omit) heterogeneous pass-through rate due to importing tenure. Columns (1) and (3) (columns (2) and (4)) include (omit) input-output linkages.

FIGURE 3.19 – Network Centrality and Import Intensity



The figure shows the relationship between import intensity in production and network centrality across domestically produced goods. I plot the domestic production network of the Chilean economy in 2013 as described by the input-output matrix. Each node represents one of the 180 products making part of the economy. The size of each node is proportional to the centrality of the product in the domestic network: the more central the product is, the larger the node. The top panel uses the PageRank centrality measure while the bottom panel uses the average between the in-degree and out-degree centrality measures. Both measures are computed weighting the edges according to the input-output linkages. The coloring of the nodes depends on the import intensity in the production of that good. Import intensity of a product is computed as the share of imported intermediate inputs over total costs. Warmer colors refer to higher import intensity. Appendix II provides additional details on the construction of the domestic input-output matrix.

FIGURE 3.20 – Consumption Share and Price Change



The figure plots the relationship between the share of each domestic good in the final consumption basket and the change in price due to a depreciation of the exchange rate. The change in price is computed in the fully calibrated model. The dashed line plots a linear fit. Table 3.15 in Appendix C reports the corresponding coefficient. Section 1.3 and Appendix A provide additional details on how consumption shares are computed.

TABLE 3.13 – On the Role of Importing Tenure

	Tenure Heterogeneity	No Tenure Heterogeneity
Frictionless	18.2	22.0
Distribution only	11.9	14.4
Distribution & Markups	9.12	11.0
All Frictions	7.60	9.17

The table compares the CPI sensitivity computed in the presence of importing tenure or omitting it across different scenarios. In the presence of importing tenure, the pass-through rate into import price is incomplete and heterogeneous. When abstracting away from importing tenure, the pass-through rate into import price is incomplete but homogeneous. I consider the following scenarios: "Frictionless", referring to a domestic economy with no frictions (i.e. no distribution costs, markup elasticity or Calvo rigidities); "Distribution only" consider a domestic economy with only distribution costs; "Distribution and Markups" refers to an economy including both distribution costs and markup elasticity; "All frictions" considers all domestic frictions together. I consider input-output linkages in all scenarios.

TABLE 3.14 – Identify Effect and Correlation across Rankings

	Tenure	Calvo	Markups	Distribution	IO
Tenure	1				
Calvo	-0.062	1			
Markups	0.12	0.029	1		
Distribution	0.16	-0.074	-0.17	1	
IO	0.13	$-0.78^{\star}$	-0.0078	0.15	1

The table reports the correlation coefficients between the change in the ranking of the products contributing the most to the overall CPI sensitivity with respect to the fully calibrated model across different scenarios. I consider the change in ranking of the products contributing the most to the overall CPI sensitivity between the fully calibrated model and an alternative scenario. I consider the following alternative scenarios: a fully calibrated economy that omits, one at the time, the role of importing tenure, nominal rigidities, distribution costs, real rigidities, and input-output linkages. I then compute the correlation between changes in ranking across scenarios. All values are not significant except for the correlation between the Calvo and input-output linkages scenarios.

TABLE 3.15 – Import Exposure and Friction Heterogeneity

	Imported Input Share								Δ Domestic Price
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PageRank Centrality	-3.680				-3.071	-3.253	-0.147	-0.156	
	(1.54)				(1.54)	(1.56)	(0.074)	(0.075)	
Distribution Margin - Intermediate		0.363			0.323		0.150		
		(0.16)			(0.16)		(0.074)		
Distribution Margin - Weighted			0.122			0.0971		0.0846	
			(0.086)			(0.085)		(0.074)	
Markup Elasticity				0.0555	0.0519	0.0507	0.110	0.107	
				(0.035)	(0.035)	(0.035)	(0.073)	(0.074)	
Final Consumption Share									-0.475
									(0.20)
Constant	0.270	0.228	0.235	0.222	0.221	0.231	-6.35e-17	-6.38e-17	0.0424
	(0.016)	(0.016)	(0.017)	(0.022)	(0.026)	(0.026)	(0.073)	(0.073)	(0.0028)
N	180	180	180	180	180	180	180	180	180

Columns (1) to (4) report the correlation coefficients between the share of imported intermediate inputs and product level characteristics in the whole sample of domestically produced goods. The share of imported intermediate inputs is computed as the share of imported intermediate inputs used in production over total costs. I consider the following characteristics: the PageRank centrality of the product in the domestic network, column (1); the distribution margin of the product, computed considering only intermediate inputs or as a weighted average between intermediate and final goods (column (2) and (3), respectively); the markup elasticity of the product, column (4). PageRank centrality is computed weighting the edges according to the input-output linkages. Appendix A provides additional information on how distribution margins and markup elasticities are computed. Column (5) regresses the PageRank centrality measure, the markup elasticity and the distribution margin for intermediate goods all together on the share of imported intermediate inputs. Similarly, column (6) uses the weighted measure of distribution costs. Column (7) and (8) run the regressions in column (5) and (6), respectively, after standardizing all the variables. Finally, column (9) reports the correlation coefficient between the change in domestic prices after a depreciation in the exchange rate and the final consumption share in the whole sample of domestically produced goods. The change in domestic prices is computed in the fully calibrated model.

Appendix D Data

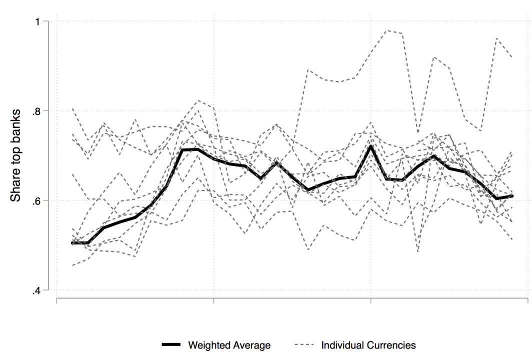
I Additional Figures and Tables

TABLE 3.16 – Expectation Dispersion - Multiple Horizons

	Across all Horizons	Quarter t	Quarter t+1	Quarter t+2	Quarter t+3	Quarter t+4
Average Dispersion (StD)**	4.62	2.90	4.01	4.96	5.76	6.54
Median Dispersion (StD)**	4.26	2.48	3.63	4.44	5.02	5.71
Average # of Forecasters	47.15	47.61	47.39	46.91	46.50	47.64
# of Quarters*	337	76	76	76	76	33

Notes: The table reports the standard deviation in the expectations of future EUR/USD exchange rate across forecasters, averaged across time. Every quarter, forecasters are asked their expectation on the EUR/USD exchange rate one to four quarters ahead. We compute the dispersion across forecaster for every quarter-horizon pair. The first column reports the average dispersion across all quarter-horizon pairs. All the other columns average across forecasts with the same horizon. Data are from the ECB Professional Forecasters survey, 2002Q1 to 2020Q4. Data on four-quarter ahead forecasts are available from 2002Q1 to 2010Q2 only.

FIGURE 3.21 – Market Share of Top Quintile



Notes: The figure shows the market share of the top quintile of investors in the New York OTC foreign exchange market. Market share are computed in terms of total transactions. The thick black shows the weighted average across all currencies, weighted by turnover. All other dotted lines represent individual currencies. Data are from the NY Fed Biannual FXC report, from 2005 to 2020. Appendix D provides additional information on the data used.

II *Mapping Strategic Behavior*

Data sources on the foreign exchange market are hardly available or comprehensive, reflecting the opaque and decentralized structure that characterized the market. Since 1990, the Bank of International Settlements collects and publishes information on turnover, instruments used, market participants etc..., providing one of few sources of data at global level. The BIS Triennial Surveys provide a clear picture of the high concentration in the foreign exchange market, both geographically and within market.

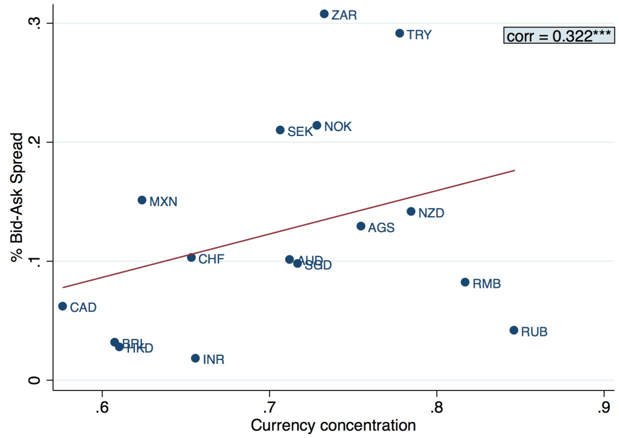
The Triennial Survey complements more frequent regional surveys conducted by national foreign exchange committees like the New York Fed Biannual FXC Report, which provides similar information at higher frequency (biannual) since 2005 for the US market. We choose to calibrate our model focusing on the US market because i) we believe it closely reflects the overall global dynamics in the foreign exchange rate market since it represents the second market worldwide (after London) and ii) data are more granular and allow cross-currency analysis.

In our model we use the market share of the top dealers to calibrate the size of the non-competitive segment, λ. From figure 2.1, a reasonable value is 70%, which is the average share of the first quintile of investors. The number of investors that have strategic behavior in the foreign market, N, is calibrated to the number of players falling into the first quintile of the distribution. The average number of top traders over the time horizon considered is $4-5$. Importantly, the NY Fed also provides information for each currency pairs (USD againt other currencies). In particular, for each currency pair, the report provides information on the market share for each quintile of the distribution. In other words, we have a cross-currency measure of λ. The number of investors in the first quintile does not change across currency (N is constant).

We evaluate the relevance of our proxy for strategic behavior, testing a standard theoretical relationship between strategic behavior and liquidity. An extensive literature in market microstructure associates the presence of non-competitive traders to higher bid-ask spreads.[155] We collect the daily bid-ask spread from Bloomberg for the set of currencies used in the main analysis for the 2019 calendar year.[156] Figure 3.22 shows the existence of a positive relationship between our measure of strategic behavior and the average daily bid-ask spread on the cross section of currencies, supporting our conjecture that concentration in the dealership market can be considered as a (approximate) measure of the size of non-competitive traders in the market.

[155]For instance, standard argument in a market microstructure textbook is that opaqueness and information asymmetry can lessen competition among investors and dealers and thus reduce market liquidity, increasing the bid-ask spread (Foucault et al. (2013), Chapter 8).
[156]See Appendix D for additional information on the data used.

FIGURE 3.22 – Cross-currency relationship between λ and bid-ask spread.



Notes: The figure shows the relationship between the presence of strategic investors and the bid-ask spread across currencies. Data are from 2019. Daily bid-ask spreads are from Bloomberg. Strategic behavior is measured with the market share of the top quintile of investors in the New York FX market in 2019, data are from the 2019 NY FED FX Report. The set of currencies coincides with those used in the main analysis. See Appendix D for additional information on the data used.

Appendix E Derivations

I *Derivation Demand Functions - Rational Expectation Case*

Each investor j solves the following problem:

$$\max_{b_t^j} E_t^j(w_{t+1}^j|\Omega_t^j) - \frac{\rho}{2}Var_t^j(w_{t+1}^j|\Omega_t^j)$$

$$\text{s.t. } w_{t+1}^j = (\omega - b_t^j)i_t + (i_t^\star + s_{t+1} - s_t)b_t^j$$

We assume that investors have symmetric rational expectation information sets, so that all j indexes on expectation and variance are dropped. We take the derivative of the objective function w.r.t. b_t^j. If the investor is strategic ($j = S$), they internalize the effect of their demand on the exchange rate.

Thus, the demand schedule is:

$$b_t^{S,i} = \frac{E_t(s_{t+1}) - s_t + i_t^\star - i_t}{\rho Var_t(s_{t+1}) + \frac{\partial s_t}{\partial b_t^{S,i}}},$$

where the $\frac{\partial s_t}{\partial b_t^j}$ represents the price impact. If the investor is competitive ($j = C$), the demand schedule follows a standard mean-variance specification:

$$b_t^C = \frac{E_t(s_{t+1}) - s_t + i_t^\star - i_t}{\rho Var_t(s_{t+1})}.$$

We can now derive an expression for the price impact of a strategic investor. Assume there are NC strategic investors, each with positive mass λ_i. Then, the market clearing condition for the foreign bond market is:

$$(1-\lambda)b_t^C + \sum_i^N \lambda_i b_t^{S,i} + (x_t + \bar{x})\bar{W} = B(1 + s_t).$$

Substituting the demand schedule and applying the Implicit function theorem, we can write:

$$(1-\lambda)\frac{\partial b_t^C}{\partial s_t}\frac{\partial s_t}{\partial b_t^{S,i}} + \lambda_i = B\frac{\partial s_t}{\partial b_t^{S,i}}$$

Thus:

$$\frac{\partial s_t}{\partial b_t^{S,i}} = \frac{\lambda_i}{B - (1-\lambda)\frac{\partial b_t^C}{\partial s_t}} \qquad \text{with } \frac{\partial b_t^C}{\partial s_t} \equiv -\frac{1}{\rho Var_t(s_{t+1})}$$

Therefore:

$$\frac{\partial s_t}{\partial b_t^{S,i}} = \frac{\lambda_i \rho Var_t(s_{t+1})}{B\rho Var_t(s_{t+1}) + (1-\lambda)} \equiv \frac{1}{N}\frac{\lambda\rho\sigma_t^2}{B\rho\sigma_t^2 + (1-\lambda)} > 0$$

where the last equality holds in case of a symmetric oligopoly (i.e. $\lambda_i = \frac{\lambda}{N}\forall i$). The price impact is positive for $\forall(B, \lambda, N, \lambda_i, \rho, \sigma)$.

Lastly, in international portfolio choice models, the value of the supply of foreign assets in domestic currency (indirectly) depends on the value of the exchange rate when foreign assets are denominated in foreign currency. Differently from standard models of strategic trading (as Kyle (1989)), non-competitive traders internalize not only their price effect on the quantity demanded but also on the quantity (value) supplied. Compared to closed economy models or cases in which foreign assets are denominated in domestic currency, the presence of this "supply effect" implies a weakly lower price impact. Let pi^F and pi^D be the price impact on a foreign and a domestic asset, respectively.

$$pi^F \equiv \frac{\partial s_t}{\partial b_t^{S,i}} = \frac{\lambda_i \rho \sigma_t^2}{B\rho\sigma_t^2 + (1-\lambda)} \qquad pi^D \equiv \frac{\partial p_t}{\partial b_t^{S,i}} = \frac{\lambda_i \rho \sigma_t^2}{(1-\lambda)}$$

where p_t is the price of the domestic asset. It is easy to show that $pi^F \leq pi^D \quad \forall (B, \rho, \sigma_t^2, \lambda_i, \lambda)$. The intuition is fairly simple. The increase in the price of a currency (foreign currency appreciates) increases the nominal value of the supply of foreign assets when denominated in domestic currency. The supply shift dampens the initial rise in price, reducing the magnitude of the price impact. The overall effect of tradings on the exchange rate is lower due to the presence of a revaluation effect. In other words, the residual net demand faced by strategic investors is more elastic than in a case with no valuation effects. The main implication is that strategic investors still reduce their exposure to foreign assets compared to competitive investors but not as much as in the case there was no valuation effect.

II Effect of Strategic Behavior on Noise and Fundamental Shock

The presence of strategic investors amplifies (dampens) the response of the exchange rate to noise (fundamental) shock.

Proof. Consider the law motion of the exchange rate, Equation 8. s_t can be rewritten as a forward looking sum of fundamentals and noises as follow:

$$s_t = -\mu \sum_{k=0}^{\infty} \mu^k \left(\Delta i_{t+k}\right) + \frac{1-\mu}{b} \sum_{k=0}^{\infty} \mu^k \left(x_{t+k}\right),$$

where $\Delta i_{t+k} = i_{t+k} - i^{\star}_{t+k}$. Therefore, the response of the exchange rate to a unit shock in noise and fundamental at impact is:

$$\text{IRF}\left(s_{t+j}, j=0\right) = \begin{cases} \dfrac{\mu}{1-\mu\rho_u}, & \text{for } \varepsilon_u = -1 \\ \\ \dfrac{(1-\mu)}{(1-\mu\rho_x)b}, & \text{for } \varepsilon_x = 1 \end{cases}$$

Taking the derivative w.r.t. μ, we find:

$$\frac{\partial \text{IRF}\left(s_{t+j}, j=0\right)}{\partial \mu} = \begin{cases} \dfrac{1}{(1-\mu\rho_u)^2} > 0 \\ \\ -\dfrac{(1-\rho_x)}{(1-\mu\rho_x)^2 b^2} < 0 \end{cases}$$

Since μ is decreasing (increasing) function of λ (N), the response of the exchange rate to a unit shock in fundamental is dampened while noise shock are amplified as λ increases (N decreases). □

III Monotonicity of Unconditional Variance

The unconditional volatility of the exchange rate is non-monotonic in the presence of strategic investors.

Proof. Consider the law of motion of the exchange rate, Equation 8, and substitute the process for fundamental and noise:

$$s_t = -\mu \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \mu^k \rho^j \varepsilon^u_{t+k-j} + \frac{1-\mu}{b} \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \mu^k \rho_x^j \varepsilon^x_{t+k-j}.$$

After some algebra, s_t can be written as summation of its backward and forward components:

$$s_t = -\frac{\mu}{1-\mu\rho_u} \left[\sum_{k=0}^{\infty} \mu^k \varepsilon^u_{t+k} + \sum_{k=1}^{\infty} \rho_u^k \varepsilon^u_{t-k} \right] + \frac{1-\mu}{b(1-\mu\rho_u)} \left[\sum_{k=0}^{\infty} \mu^k \varepsilon^x_{t+k} + \sum_{k=1}^{\infty} \rho_x^k \varepsilon^x_{t-k} \right].$$

Thus, the unconditional variance of the exchange rate is:

$$\mathrm{Var}(s) = \frac{\mu^2 \sigma_u^2}{(1-\mu\rho_u)^2} \left[\frac{1}{1-\mu^2} + \frac{\rho_u^2}{1-\rho_u^2} \right] + \frac{(1-\mu)^2 \sigma_x^2}{(1-\mu\rho_x)^2 b^2} \left[\frac{1}{1-\mu^2} + \frac{\rho_x^2}{1-\rho_x^2} \right],$$

which is a combination of the variances of fundamental and noise shocks. Taking the derivative of $\mathrm{Var}(s)$ w.r.t. μ, we find:

$$\begin{aligned} \frac{\partial \mathrm{Var}(s)}{\partial \mu} =& \frac{\mu \sigma_u^2}{(1-\mu\rho_u)^3} \left[\frac{1}{1-\mu^2} + \frac{\rho_u^2}{1-\rho_u^2} \right] + \frac{\mu^3 \sigma_u^2}{(1-\mu\rho_u)^2 (1-\mu^2)^2} - \\ & \frac{(1-\mu)(1-\rho_x)\sigma_x^2}{(1-\mu\rho_x)^3 b^2} \left[\frac{1}{1-\mu^2} + \frac{\rho_x^2}{1-\rho_x^2} \right] + \frac{\mu (1-\mu)^2 \sigma_x^2}{(1-\mu\rho_x)^2 (1-\mu^2)^2 b^2}. \end{aligned}$$

The unconditional volatility of the exchange rate is increasing in λ iff:

$$\frac{(1+\mu\rho_x)\sigma_x^2}{(1-\mu\rho_x)^2(1+\mu)(1+\rho_x)b^2} - \frac{\mu\sigma_x^2}{(1-\mu\rho_x)^2(1+\mu)^2b^2} >$$
$$\frac{\mu\sigma_u^2}{(1-\mu\rho_u)^2}\frac{(1+\mu\rho_u)}{(1-\mu^2)(1-\rho_u^2)} + \frac{\mu^3\sigma_u^2}{(1-\mu\rho_u)^2(1-\mu^2)^2},$$

that can be rewritten as follows:

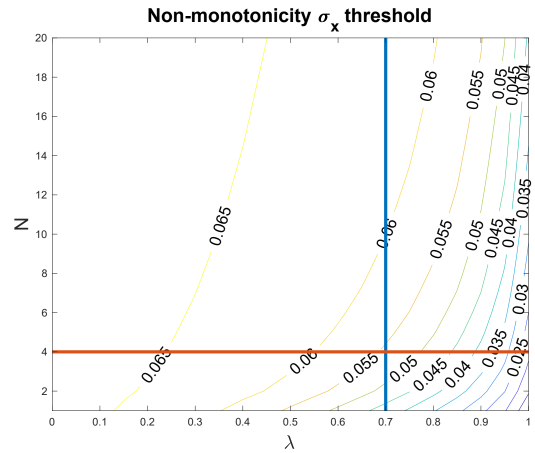
$$\frac{\text{Var}(x)}{\text{Var}(\Delta i)}\frac{1}{b^2} > \left[\frac{(1+\mu^2\rho_x)(1-\rho_x)}{\mu(1+\mu\rho_u)(1-\mu^2)+\mu^3(1-\rho_u^2)}\frac{(1-\mu\rho_u)^2(1-\mu)^2}{(1-\mu\rho_x)^2}\right]^{-1}. \tag{1}$$

□

Equation (1) implies that the unconditional variance of the exchange rate is increasing in λ if the variance of the noise shock is sufficiently high relative to the variance of the fundamental process.

The non monotonic case is not relevant given standard parametrizations, including ours. Let define $\underline{\sigma_x}$ as the minimum value of the volatility of the noise process such that the relationship between the extent of strategic behavior and exchange rate variance is not monotone anymore. Figure 3.23 shows $\underline{\sigma_x}$ for different combinations of N and λ. In our calibration, the noise shock should be at least 45% less volatile in order not to have a monotonic relationship between strategic behavior (λ and/or N) and unconditional variance. For other values of λ or N, $\underline{\sigma_x}$ is at least 50% lower than the value of σ_x implied by Figure 3.24 in Appendix G. In a competitive market, $\sigma_x \approx 0.12$ and monotonicity does not arise if $\underline{\sigma_x} < 0.066$. In highly non-competitive markets, $\sigma_x \approx 0.06$ and monotonicity does not arise

FIGURE 3.23 – Noise threshold for different combinations of N and λ.



Notes: The figure shows the minimum value of the volatility of the noise process, σ_x, that guarantees that the volatility of the exchange rate is monotonically increasing in the presence of strategic behavior (higher λ and/or lower N). The threshold is computed using Equation (1). We compute the minimum value of σ_x for different levels of λ and N. The horizontal and vertical lines pin down the combination of λ and N used in the parametrization of the basic framework. Remaining parameters are constant, see Table 2.1.

if $\underline{\sigma_x} < 0.015$.

Moreover, notice that the threshold value depends on ρ_x, ρ_u and b. The monotonicity in the relationship between strategic behavior and unconditional variance is robust because our calibration is particularly conservative. Only more persistent noise processes or less persistent fundamental processes would be consistent with standard calibrations; similarly, only higher values of home bias (lower b) would be acceptable. Higher values of ρ_x, lower values of ρ_u and lower b all decrease the threshold, relaxing the condition for monotonicity.

Excess return is more predictable as λ increases.

Proof. Consider the law motion of the exchange rate, Equation 8:

$$s_t = \mu \left(E_t \left(s_{t+1}\right) + i_t^{\star} - i_t\right) + (1-\mu)\frac{\bar{x}}{b} + (1-\mu)\frac{1}{b}x_t,$$

where only the first term depends on fundamentals. Manipulating it, we can derive the j-period change in currency price as follows:

$$\Delta s_{t+j} = -\mu \sum_{k=0}^{\infty} \mu^k \left(\Delta i_{t+j+k} - \Delta i_{t+k}\right).$$

With Δs_{t+j} in hand, we can then calculate:

$$\begin{aligned}
\beta_1 = \frac{\text{Cov}\left(\Delta s_{t+1} - \Delta i_t; \Delta i_t\right)}{\text{Var}(\Delta i_t)} &= \left[\text{Cov}\left(-\mu \sum_{k=0}^{\infty} \mu^k \left(\Delta i_{t+k+1} - \Delta i_{t+k}\right); \Delta i_t\right) - \text{Var}\left(\Delta i_t\right)\right] / \text{Var}(\Delta i_t) \\
&= \left[-\mu \sum_{k=0}^{\infty} \mu^k \text{Cov}\left(\Delta i_{t+k+1} - \Delta i_{t+k}; \Delta i_t\right) - \text{Var}\left(\Delta i_t\right)\right] / \text{Var}(\Delta i_t) \\
&= \left[-\mu \sum_{k=0}^{\infty} \mu^k \rho_u^k (\rho_u - 1) \text{Var}(\Delta i_t) - \text{Var}(\Delta i_t)\right] / \text{Var}(\Delta i_t) \\
&= -(1-\mu)\frac{1}{1-\mu\rho_u} < 0,
\end{aligned}$$

which is negative for each value of μ and increasing (decreasing) in μ (in λ). □

Notice that predictability reversal does not arise in our model, differently from Bacchetta and Van Wincoop (2010) and Engel (2016). Formally define the j-period ahead excess return as $q_{t+j} = s_{t+j+1} - s_{t+j} - (i_{t+j} - i^*_{t+j})$, and consider the following regression:

$$q_{t+j} = \alpha + \beta_j(i_t - i_t^*) + \epsilon_{t+j}. \tag{2}$$

The coefficient of interest, β_j, is:

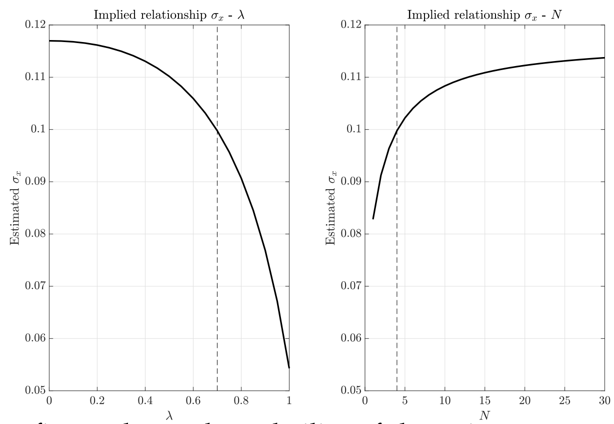
$$\begin{aligned}
\beta_j =& \frac{\text{Cov}(q_{t+j}, \Delta i_t)}{\text{Var}(\Delta i_t)} \\
& \frac{1}{\text{Var}(\Delta i_t 9} \left(\text{Cov}(\Delta s_{t+j}, \Delta i_t) - \text{Cov}(\Delta i_{t+j-1}, \Delta i_t)\right) \\
& \frac{1}{\text{Var}(\Delta i_t)} \left[\text{Cov}\left(-\mu \sum_{k=0}^{\infty} \mu^k \left(\Delta i_{t+k+j} - \Delta i_{t+k+j-1}\right); \Delta i_t\right) - \text{Cov}\left(\Delta i_{t+j-1}, \Delta i_t\right)\right] \\
& \frac{1}{\text{Var}(\Delta i_t)} \left[\left(-\mu \sum_{k=0}^{\infty} \mu^k \text{Cov}\left(\Delta i_{t+k+j} - \Delta i_{t+k+j-1}\right); \Delta i_t\right) - \text{Cov}\left(\Delta i_{t+j-1}, \Delta i_t\right)\right] \\
& -\mu \sum_{k=0}^{\infty} \mu^k (\rho_u^{k+j} - \rho_u^{k+j-1}) - \rho_u^{j-1} \\
& -\mu\rho_u^{j-1}(\rho_u - 1)\frac{1}{1-\mu\rho_u} - \rho^{j-1} = -\rho^{j-1}\frac{1-\mu}{1-\mu\rho_u} \leq 0.
\end{aligned}$$

Lastly, notice that $\frac{\partial \beta_j}{\partial j} = -(j-1)\rho_u^{j-1}\left(\frac{1-\mu}{1-\mu\rho_u}\right) < 0$. Therefore, for $j \to \infty$, the coefficient $\beta_j \to 0$ monotonically, excluding any reversal.

Appendix F Solution Method of Dispersed Information Model

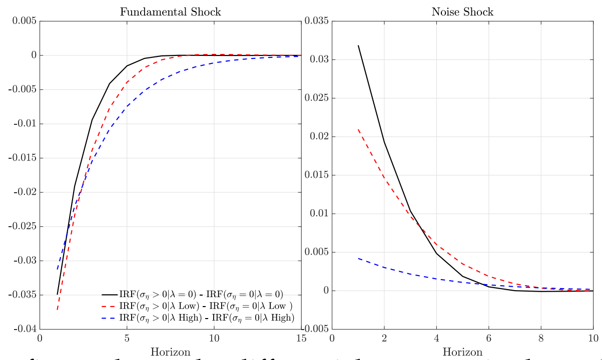
Appendix G Additional Tables and Figures

FIGURE 3.24 – Relationship between Strategic Behavior and Noise Volatility



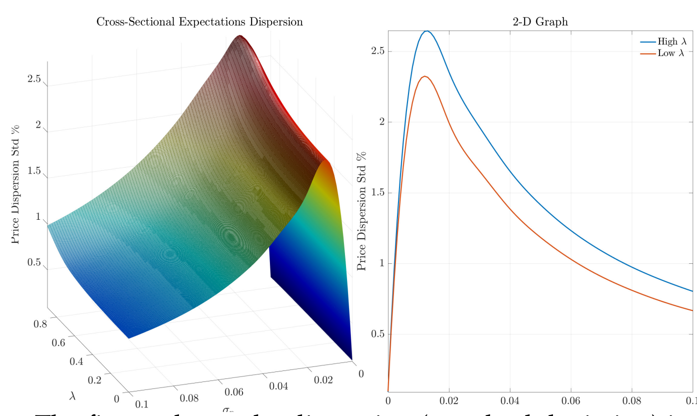
Notes: The figure shows the volatility of the noise component, σ_x, required to match the target volatility of the exchange rate change in the basic framework, for different level of strategic behavior. The left panel considers different levels of strategic behavior in terms of λ for a number of strategic investors equal to $N = 4$. The left panel considers different levels of strategic behavior in terms of N for a total size of strategic investors equal to $\lambda = 0.7$. All other parameters are constant and summarized in Table 2.1.

FIGURE 3.25 – Strategic Behavior and Dispersed Information: Non-linear Interaction



Notes: The figure shows the differential response in the exchange rate to fundamental (left panel) and noise (right panel) shocks across different scenarios in the quantitative model of Section 2.4. The black solid line shows the difference in the response of the exchange rate computed by comparing two competitive economies ($\lambda = 0$), one with dispersed information and one without dispersed information. The dash red line shows the difference in the response of the exchange rate computed by comparing two economies with an intermediate level of strategic behavior ($\lambda = 0.3$), one with dispersed information and one without dispersed information. The dash blue line shows the difference in the response of the exchange rate computed by comparing two economies with the benchmark level of strategic behavior ($\lambda = 0.6$), one with dispersed information and one without dispersed information. All other parameters are constant and summarized in Table 3.17.

FIGURE 3.26 – Exchange Rate Expectation - Dispersion



Notes: The figure shows the dispersion (standard deviation) in the one-quarter exchange rate expectations across investors for different level of strategic behavior (λ) and precision of the signal on fundamentals (σ_η) implied by the model in Section 2.4. The left panel shows the dispersion in expectations for values of $\lambda \in [0,1]$, and $\sigma_\eta \in [0,0.1]$. The right panel shows the dispersion in expectation for two levels of strategic behavior ("Low" with $\lambda = 0$, and "High" with $\lambda = 0.6$) and a precision of the signal σ_η between 0 and 0.1.

TABLE 3.17 – Parametrization Quantitative Model

	Value	Moment - Target	Data	Model
λ	0.6	Share transactions top dealers (1st quintile) in NYFXC		
N	4	Number of top dealers (1st quintile) in NYFXC		
ρ_u	0.85	Average persistence AR(1) Δi_t		
σ_u	0.0088	Average std innovation AR(1) Δi_t		
$\sigma(\Delta s_t)$	0.04	(Average) Std FX change		
σ_η	0.012	Same Quarter FX Dispersion	0.02	0.017
σ_x	0.0352	$\sigma(\Delta s_t)$	0.04	0.041
ρ_x	0.9	FX RW/Average R^2	0.024 (0.006)	0.0055
ρ	20	UIP deviation level	-1.24	-0.84
b	0.33	Home Bias		
$\bar{k}$	10			

Notes: The table summarizes the parametrization used in the main quantitative exercise. For each parameters, we report the value used in the model, the corresponding moment and data used to calibrate, and, if applicable, the target moment used to estimate it. Appendix A provides additional information on the data used.

Appendix H Additional Theoretical Results

I Examples of Homothetic Demand Systems

Here, we show how a few popular choices of demand system can be written as the specification in Definition 1.

1. Mixed Logit (McFadden, 1974, Berry, 1994). This demand system is given through the choice of

$$\mathcal{S}_i\left(\widetilde{\boldsymbol{p}};\boldsymbol{\varsigma}\right) \equiv \int_{\sigma\in\Sigma} \frac{\exp\left(-\sigma\log\widetilde{p}_i\right)}{\sum_{i'\in V_t}\exp\left(-\sigma\log\widetilde{p}_{i'}\right)} dF\left(\sigma;\boldsymbol{\varsigma}\right), \qquad i\in V,$$

   with $\mathcal{H}\left(\boldsymbol{p};\boldsymbol{\varsigma}\right) = \prod_{i\in V} p_i^{1/|V|}$.

2. Homothetic Translog (e.g., Diewert, 1976, Feenstra, 2003). This demand system is given through the choice of

$$\mathcal{S}_i\left(\widetilde{\boldsymbol{p}};\boldsymbol{\varsigma}\right) \equiv \theta_i + \sum_{i'\in V}\sigma_{ii'}\log\widetilde{p}_{i'}, \qquad i\in V,$$

   with $\sum_{i\in V}\theta_i = 1$ and $\sigma_{ii'} = \sigma_{i'i}$ and $\sum_{i'}\sigma_{ii'} = 0$, and $\mathcal{H}\left(\boldsymbol{p};\boldsymbol{\varsigma}\right) = \prod_{i\in V} p_i^{1/|V|}$. Note that this demand system also coincides with the AIDS demand system of Deaton and Muellbauer (1980) when the latter is restricted to be homothetic.

3. Homothetic with a Single Aggregator (HSA) (Matsuyama and

Ushchev, 2017, Matsuyama, 2022). This demand system is given by

$$\mathcal{H}\left(\boldsymbol{p};\boldsymbol{\varsigma}\right)=h,\ \text{such that}\ \sum_{i\in V_t}\mathcal{S}_i\left(\frac{\boldsymbol{p}}{h};\boldsymbol{\varsigma}\right)=1.$$

We may particularly focus on the family $\mathcal{S}_i\left(\cdot;\boldsymbol{\varsigma}\right)$ defined as[157]

$$\log\mathcal{S}_i\left(\boldsymbol{p};\boldsymbol{\varsigma}\right)\equiv\log\mathcal{S}\left(p_i;\boldsymbol{\varsigma}\right)\equiv-\int_o^{\log p_i}\exp\left(\sum_{k=0}^{K}\theta_k x^k\right)dx,\qquad(1)$$

where $\boldsymbol{\varsigma}\equiv\left(\theta_0,\cdots,\theta_K\right)$ are the parameters of the family. Note that Equation (1) nests CES demand for the case of $\theta_0=\log\left(\sigma-1\right)$ and $\theta_k=0$ for $k\geq 1$. In particular, the elasticity of demand for the product i at time t corresponding to the HSA demand in Equation (1) is given by

$$\left.\frac{\partial\log q_i}{\partial\log p_i}\right|_{h\,\text{const.}}=-\left[1+\exp\left(\sum_{k=0}^{K}\theta_k\left(\log p_i\right)^k\right)\right],$$

which varies from $-e^{\theta_o}$ whenever $\theta_k\neq 0$ for some $k\geq 1$.

II Second-Order Approximation of the Change in Price Index for HSA/HIA

We can provide second order approximations for the change in the unit expenditure function for the HSA/HIA demand systems introduced in Definition 3. Using Proposition 1 and Lemma 1, we find the following approximations in the case of each demand system

[157] The expression in Equation (1) for function $\mathcal{S}$ satisfies the conditions in Matsuyama and Ushchev (2017) to ensure that there exits a well-defined homothetic utility function rationalizing this demand function. Since $d\log\mathcal{S}(p)/d\log p<0$, the implication is that all products are gross substitutes everywhere.

$$
\begin{aligned}
\Delta \log P_t \approx & \sum_{i \in V_t^*} \overline{\overline{s_{it}^*}}\, \Delta \log p_{it} - \sum_{i \in V_t} \overline{\overline{s_{it}^*}}\, \Delta \varphi_{it} \\
& + \overline{\overline{\overline{\mu}_t^*}}\, \Delta \log \Lambda_t^* + \sum_{i \in V_t^*} \overline{\overline{\mu_{it}}}\, \Delta s_{it}^* - \begin{cases} \sum_{i \in V} \overline{\overline{\mu}}_{it}\, \Delta s_{it}, & \text{HSA}, \\ \sum_{i \in V} \overline{\overline{\left(\left(1 + \frac{\overline{\mu}_t^* - \overline{\mu}_t}{1 + \overline{\mu}_t} \right) \mu_{it} \right)}}\, \Delta s_{it}, & \text{HIA}. \end{cases}
\end{aligned} \tag{2}
$$

$$
\approx \langle \Delta \log p_{it} \rangle + \langle \overline{\overline{\mu}}_{it}\, \Delta \log s_{it}^* \rangle \tag{3}
$$

$$
+ \overline{\overline{\langle \mu_{it} \rangle}}\, \Delta \log \Lambda_t^* - \begin{cases} \sum_{i \in V} \overline{\overline{\mu}}_{it}\, \Delta s_{it}, & \text{HSA}, \\ \sum_{i \in V} \overline{\overline{\left(\left(1 + \frac{\langle \mu_{it} \rangle - \overline{\mu}_t}{1 + \overline{\mu}_t} \right) \mu_{it} \right)}}\, \Delta s_{it}, & \text{HIA}. \end{cases} \tag{4}
$$

The proof of the result closesly follows that of Lemma 2 to approximate the integrals in Equations (14) and (16), with the additional simplification given by Equation (22), over time τ from $t - 1$ to t.

III The Gap Between the CES- and Kimball-Inferred Quality Contribution

We study the implications of inferring quality if we misspecify the underlying Kimball preferences to be CES. The next proposition compares the contribution of quality changes under CES and Kimball.

Proposition 2. *Consider using a misspecified CES demand system with elasticity of substitution σ^c to infer quality φ_{it}^c based on observed sequences of prices and* *quantities that are rationalized by an underlying Kimball demand system. The gap between the true and the misspecified measures of quality change is approximately given by*

$$\sum_{i \in V_t^*} \overline{\overline{s_{it}^*}} \left(\Delta \varphi_{it} - \Delta \varphi_{it}^c\right) \approx \left\langle \left(\mu^c - \overline{\overline{\mu_{it}}}\right) \Delta \log s_{it}^* \right\rangle + \sum_{i \in V_t^*} \overline{\overline{s_{i\tau}^* \mu_{i\tau}}} \Delta \log s_{i\tau}^*$$
$$+ \left(\overline{\overline{\overline{\mu}_\tau^*}} - \overline{\overline{\langle \mu_{it} \rangle}}\right) \left(\Delta \log \Lambda_t^* + \Delta \log A_t\right), \qquad (5)$$

where, as before, $\overline{\overline{v_{it}}} \equiv \frac{1}{2}(v_{it-1} + v_{it})$ *stands for the Trnqvist mean of variable* v_{it}*, and where* $\mu^c \equiv \frac{1}{\sigma^c - 1}$.

Proof. See Appendix I. □

The first term on the right hand side of Equation (5) (CES-Baseline Gap) depends on the gap in the love-of-variety proxies measured by the CES and Kimball own-price elasticities and the growth in expenditure shares. For instance, if the CES estimate of own-price elasticity is lower than that of Kimball, and thus the measure of love of variety μ^c exceeds the average of the Kimball proxies $\overline{\overline{\mu_{it}}}$, the contribution of this term is negative or positive, depending on whether the shares of base products in the common set falls or rises.

The second term on the right hand side of Equation (5) (Elasticity Heterogeneity) accounts for the contribution of reallocations of expenditure across products and the heterogeneity in own-price elasticities. Under Kimball, we infer higher quality change if expenditure shifts toward products with lower price elasticities. Finally, the last term on the second line of Equation (5) (CVS-Baseline Gap), accounts for the gap between the love-of-variety proxies between the common set V_t^* and the base set O of products. If the underlying demand is indeed CES, then both the second and the third term are always zero since the own-price elasticities are constant at $\sigma_{it} \equiv \sigma$.

IV Comparison with Feenstra (1994)

In this section, we provide a brief comparison of the conceptual distinction between our approach and that of Feenstra (1994), which in turn builds on earlier insights of Leamer (1981). For this purpose, let us consider a CES demand specification presented in Section 3.2.2.1, which leads to the following simple specification of demand

$$\Delta \log \widehat{q}_{it} = -\sigma \Delta \log \widehat{p}_{it} + \Delta \varphi_{it},$$

where we have defined log quantity and price relative to the base product $\widehat{q}_{it} \equiv q_{it}/q_{ot}$ and $\widehat{p}_{it} \equiv p_{it}/p_{ot}$ in a simple setting where the set of base products is a singleton $O \equiv \{o\}$, and where, as before, φ_{it} stands for the demand shock. The Leamer–Feenstra approach to identification begins with positing a supply relationship of the form

$$\Delta \log \widehat{p}_{it} = \zeta \log \Delta \widehat{q}_{it} + \Delta \xi_{it}, \tag{6}$$

where $\zeta > 0$ stands for the supply elasticity. The first key identification assumption is that the supply and demand shocks are uncorrelated $\mathbb{E}\left[\Delta\xi_{it}\Delta\varphi_{it}\right] = 0$. If we know the supply elasticity ζ, then this assumption leads to a synthetic instrument $z_{it}^{F-L}(\zeta) \equiv \Delta \log \widehat{p}_{it} - \zeta \log \Delta \widehat{q}_{it}$ that allows us to identify σ through the moment condition

$$\mathbb{E}\left[\left(\Delta \log \widehat{q}_{it} + \sigma \Delta \log \widehat{p}_{it}\right) \times z_{it}^{F-L}(\zeta)\right] = 0. \tag{7}$$

As shown in Feenstra (2010), the second key identification assumption is that there exists at least two products i and j for which the ratio of the variances of demand schock and supply shocks are not identical $\left(\mathbb{V}\left[\Delta\varphi_{it}\right]/\mathbb{V}\left[\Delta\xi_{it}\right] \neq \mathbb{V}\left[\Delta\varphi_{jt}\right]/\mathbb{V}\left[\Delta\xi_{jt}\right]\right)$.[158] We can think of the role of this additional identification by heteroskedasticity assumption as that of identifying the supply elasticity ζ, which would then enable condition (7) to identify the price elasticity of demand σ. In practice, the estimation strategy combines these identification assumptions to simultaneously estimate both ζ and σ.

Now, let us compare Equation (6) with our pricing Equation (11). Assuming small relative changes in all variables, we can write the change in log price in terms of the change in log quantity and other variables as:

$$\Delta \log p_{it} \approx \underbrace{\frac{\frac{\partial \log mc_{it}}{\partial \log q_{it}} + \frac{\partial \log \mu_{it}}{\partial \log q_{it}}}{1 - \frac{\partial \log \mu_{it}}{\partial \log p_{it}}}}_{\equiv \zeta_{it}} \Delta \log q_{it} + \underbrace{\frac{\frac{\partial \log mc_{it}}{\partial \varphi_{it}} + \frac{\partial \log \mu_{it}}{\partial \varphi_{it}}}{1 - \frac{\partial \log \mu_{it}}{\partial \log p_{it}}} \Delta \varphi_{it} + \frac{\frac{\partial \log mc_{it}}{\partial w_{it}}}{1 - \frac{\partial \log \mu_{it}}{\partial \log p_{it}}} \Delta w_{it} + \Delta v_{it}}_{\equiv \Delta \xi_{it}}.$$

[158] See also Soderbery (2015) for a detailed discussion of how this condition helps identify the elasticities using specific examples from trade data.

We can make two observations. First, in general the supply elasticity may vary over time and across products. Second, and more importantly, there are two potential grounds for the violations of the Leamer–Feenstra identification assumption $\mathbb{E}\left[\Delta\xi_{it}\Delta\varphi_{it}\right] = 0$. First, to the extent that marginal cost depends on quality, i.e., $\frac{\partial \log mc_{it}}{\partial \varphi_{it}} \neq 0$, there is a mechanical correlation between supply shocks $\Delta\xi_{it}$ and demand shocks $\Delta\varphi_{it}$. In addition, to the extent that shocks to production costs Δw_{it} leads to endogenous responses in product quality, we find another potential source of correlation between supply and demand shocks.

In contrast, our approach begins by assuming a simple dynamic process like that of Equation (7) on demand shocks. The same pricing Equation (11) now implies that $\mathbb{E}\left[\Delta u_{it} \log p_{it-2}\right]$, which leads to the following moment condition:

$$\mathbb{E}\left[\left(\Delta \log \widehat{q}_{it} + \sigma\Delta \log \widehat{p}_{it} - \rho\left(\Delta \log \widehat{q}_{it-1} + \sigma\Delta \log \widehat{p}_{it-1}\right)\right) \times \log p_{it-2}\right] = 0.$$

If we know ρ, the term $\rho\left(\Delta \log \widehat{q}_{it-1} + \sigma\Delta \log \widehat{p}_{it-1}\right)$ gives us a control function that accounts for the potential persistence between lagged price and current change in demand shocks, allowing us to identify the price elasticity σ. To recover the persistence parameter ρ, the same Equation (7) also implies that $\mathbb{E}\left[\Delta u_{it}\varphi_{it-2}\right]$ leading to another moment condition

$$\mathbb{E}\left[\left(\Delta \log \widehat{q}_{it} + \sigma\Delta \log \widehat{p}_{it} - \rho\left(\Delta \log \widehat{q}_{it-1} + \sigma\Delta \log \widehat{p}_{it-1}\right)\right) \times \varphi_{it-2}\right] = 0.$$

Just like the Leamer–Feenstra approach, we also combine the moment conditions in a GMM framework to jointly estimate both σ and ρ.

To summarize, our approach averts the need to make the counterfactual assumption that marginal costs do not depend on product quality by relying on the panel structure of the data and imposing restrictions on the dynamics of demand shocks.

Appendix I Proofs and Derivations

I *Proofs*

Proof for Proposition 1. First, for the HA demand, defined by Equations (3) and (1), the change in log expenditure share of product $i \in V$ satisfies

$$d\log s_{i\tau} = -\left(\sigma_{i\tau} - 1\right)\left(d\log p_{i\tau} - d\log \varphi_{i\tau} - d\log H_{\tau}\right) - d\log A_{\tau}, \qquad (1)$$

where $\sigma_{i\tau}$ is defined by Equation (12), and where $H_{\tau} \equiv \mathcal{H}\left(\left(e^{-\varphi_{i\tau}} p_{i\tau}\right)_{i\in V}; \varsigma\right)$ and $A_{\tau} \equiv \sum_{i\in V}\left(e^{-\varphi_{i\tau}}\frac{p_{i\tau}}{H_{\tau}}\right)\mathcal{D}_i\left(e^{-\varphi_{i\tau}}\frac{p_{i\tau}}{H_{\tau}}\right)$. Equation (1), in turn, leads to the following equality for any $i \in V$:

$$d\log p_{i\tau} - d\varphi_{i\tau} = d\log H_{\tau} - \mu_{i\tau}\left(d\log A_{\tau} + d\log s_{i\tau}\right). \qquad (2)$$

Now, we can expand the change in the unit expenditure function of any homothetic preferences as

$$d\log P_\tau = \sum_{i\in V} \frac{\partial \log P_\tau}{\partial \log p_{i\tau}} \left(d\log p_{i\tau} - d\varphi_{i\tau}\right) = \sum_{i\in V_{t-1}\cup V_t} s_{i\tau}\left(d\log p_{i\tau} - d\log \varphi_{i\tau}\right), \tag{3}$$

where we have used the Shephard's lemma in the second equality. Substituting from Equation (2) in Equation (3), we find

$$d\log P_t = d\log H_\tau - \overline{\mu}_\tau d\log A_\tau - \sum_{i\in V_{t-1}\cup V_t} \mu_{i\tau} ds_{i\tau}, \tag{4}$$

where $\overline{\mu}_\tau \equiv \sum_{i\in V_{t-1}\cup V_t} s_{i\tau}\mu_{i\tau}$.

Now, we compute the change in the logarithm of the expenditure share of common set

$$\begin{aligned} d\log \Lambda_t^* &= \frac{\sum_{i\in V_t^*} ds_{i\tau}}{\Lambda_\tau^*} = \sum_{i\in V_t^*} s_{it}^* d\log s_{i\tau}, \\ &= -\sum_{i\in V_t^*} s_{it}^*\left(\sigma_{i\tau}-1\right)\left(d\log p_{i\tau} - d\log\varphi_{i\tau} - d\log H_\tau\right) - d\log A_\tau, \\ &= -\left(\overline{\sigma}_\tau^* - 1\right)\left(d\log\left(\frac{D_\tau^*}{\Phi_\tau^*}\right) - d\log H_\tau\right) - d\log A_\tau - \mathbb{C}^*\left(\sigma_{i\tau}, d\log p_{i\tau} - d\varphi_{i\tau}\right), \end{aligned} \tag{5}$$

where in the second line, we have used Equation (1), and where in the third line we have used the definitions of the Divisia and Quality indices in Equation (15), and have defined the covariance between the demand elasticity σ_{it} and the change in quality-adjusted prices as

$$\mathbb{C}^*\left(\sigma_{it}, d\log p_{it} - d\varphi_{it}\right) \equiv \sum_{i\in V_t^*} s_{it}^*\left(\sigma_{i\tau} - \overline{\sigma}_\tau^*\right)\left(d\log p_{i\tau} - d\log\varphi_{i\tau}\right).$$

We now use Equation (5) to substitute for $d \log H_\tau$ in Equation (4) and find:

$$d \log P_\tau = d \log \left(\frac{D_\tau^*}{\Phi_\tau^*} \right) + \frac{1}{\overline{\sigma}_\tau^* - 1} d \log \Lambda_t^* + \left(\frac{1}{\overline{\sigma}_\tau^* - 1} - \overline{\mu}_\tau \right) d \log A_\tau$$
$$+ \frac{1}{\overline{\sigma}_\tau^* - 1} \mathbb{C}^* \left(\sigma_{i\tau}, d \log p_{i\tau} - d\varphi_{i\tau} \right) - \sum_{i \in V_{t-1} \cup V_t} \mu_{i\tau} ds_{i\tau}. \tag{6}$$

The last step for proving Equation (14) is to compute the covariance term. Using Equation (2), we can rewrite this covariance as $\mathbb{C}^* \left(\sigma_{i\tau}, d \log p_{i\tau} - d\varphi_{i\tau} \right) = -\mathbb{C}^* \left(\sigma_{i\tau}, \mu_{i\tau} \right) d \log A_\tau - \mathbb{C}^* \left(\sigma_{i\tau}, \mu_{i\tau} d \log s_{i\tau} \right)$. The first term can be simplified to:

$$\mathbb{C}^* \left(\sigma_{i\tau}, \mu_{i\tau} \right) = \sum_{i \in V_t^*} s_{i\tau}^* \left(\sigma_{i\tau} - \overline{\sigma}_\tau^* \right) \mu_{i\tau} = 1 - \overline{\mu}_\tau^* \left(\overline{\sigma}_\tau^* - 1 \right),$$

while the second term can be written as

$$\begin{aligned}
\mathbb{C}^* \left(\sigma_{i\tau}, \mu_{i\tau} d \log s_{i\tau} \right) &= \sum_{i \in V_t^*} s_{i\tau}^* \left(\sigma_{i\tau} - \overline{\sigma}_\tau^* \right) \mu_{i\tau} d \log s_{i\tau}, \\
&= \sum_{i \in V_t^*} s_{i\tau}^* \left(\sigma_{i\tau} - \overline{\sigma}_\tau^* \right) \mu_{i\tau} \left(d \log \Lambda_\tau^* + d \log s_{i\tau}^* \right), \\
&= \left[1 - \overline{\mu}_\tau^* \left(\overline{\sigma}_\tau^* - 1 \right) \right] d \log \Lambda_\tau^* \\
&\qquad + \sum_{i \in V_t^*} s_{i\tau}^* \left(1 - \mu_{i\tau} \left(\overline{\sigma}_\tau^* - 1 \right) \right) d \log s_{i\tau}^*, \\
&= \left[1 - \overline{\mu}_\tau^* \left(\overline{\sigma}_\tau^* - 1 \right) \right] d \log \Lambda_\tau^* - \left(\overline{\sigma}_\tau^* - 1 \right) \sum_{i \in V_t^*} \mu_{i\tau} ds_{i\tau}^*.
\end{aligned}$$

Combining the two terms, we find

$$\mathbb{C}^* \left(\sigma_{i\tau}, d \log p_{i\tau} - d\varphi_{i\tau} \right) = \left(\overline{\sigma}_\tau^* - 1 \right) \left[\sum_{i \in V_t^*} \mu_{i\tau} ds_{i\tau}^* - \left(\frac{1}{\overline{\sigma}_\tau^* - 1} - \overline{\mu}_\tau^* \right) \left(d \log \Lambda_\tau^* + d \log A_\tau \right) \right].$$

Substituting the above expression in Equation (6) leads to Equation (14).

To prove Equation (16), we use Equation (2) , and the normalization that $\langle d\varphi_{i\tau}\rangle \equiv 0$, to find

$$d\log H_\tau = \langle d\log p_{i\tau}\rangle + \langle \mu_{i\tau} d\log s_{i\tau}\rangle + \langle \mu_{i\tau}\rangle\, d\log A_t. \tag{7}$$

Using the definitions in Equation (15) and Equation (2), and using the above result leads to

$$d\log\left(\frac{D^*_\tau}{\Phi^*_\tau}\right) = \langle d\log p_{i\tau}\rangle + \langle \mu_{i\tau} d\log s_{i\tau}\rangle + \left(\langle \mu_{i\tau}\rangle - \overline{\mu}^*_\tau\right) d\log A_t - \sum_{i\in V^*_t} s^*_{i\tau}\mu_{i\tau} d\log s_{i\tau},$$

which in turn leads to the desired result if we note that $d\log s_{i\tau} = d\log s^*_{i\tau} + d\log \Lambda^*_\tau$. □

Proof for Lemma 1. The case of HSA trivially follows from the observation that $d\log A_\tau \equiv 0$. In the HIA case, the results of Matsuyama and Ushchev (2017) along with the definitions of the indices H_τ and A_τ imply that $P_\tau = A_\tau H_\tau$. Combining $d\log P_\tau = d\log H_\tau + d\log A_\tau$ and Equation (4) implies

$$d\log A_\tau = -\frac{1}{1+\overline{\mu}_\tau}\sum_{i\in V} s_{i\tau}\mu_{i\tau} d\log s_{i\tau},$$

which, using the leads to the desired result. □

Proof for Lemma 2. We use the following standard result on the error of the trapzoidal integration rule:

$$I \equiv \int_{v_{\tau-1}}^{v_\tau} f(v)\, dv = \sum_j \tfrac{1}{2}\left(f(v_{t-1}) + f(v_{t-1})\right)(v_t - v_{t-1}) - \frac{1}{12} f''\left(v^\dagger\right)(v_t - v_{t-1})^3, \tag{8}$$

for some $v^{\dagger} \in [v_{t-1}, v_t]$. From this, it then follows that $I = \overline{\overline{f(v_t)}}\Delta v_t + O\left(|\Delta v_t|^3\right)$. We apply this result to each of the terms in Equation (14). For instance, doing a change of variable $v \equiv \log p_{i\tau}$, we find for the first term (from Equation 15) that

$$\int_{t-1}^{t} s_{i\tau}^* d\log p_{i\tau} = \int_{\log p_{it-1}}^{\log p_{it}} s_i^*(v)\, dv = \overline{\overline{s_{it}^*}}\, \Delta \log p_{it} + O\left(|\Delta \log p_{it}|^3\right).$$

Applying the same treatment to the other terms leads to the desired result.

□

Proof of Proposition 2. Let σ^c denote the constant CES elasticity and consider the same path as that in Proposition 1 between periods $t-1$ and t. From Equations (2) and (7), the inferred change in quality under the Kimball and the CES demand are given by

$$d\varphi_{i\tau} = d\log p_{i\tau} + \mu_{i\tau} d\log s_{i\tau} - \langle d\log p_{i\tau}\rangle - \langle \mu_{i\tau} d\log s_{i\tau}\rangle + (\mu_{i\tau} - \langle \mu_{i\tau}\rangle)\, d\log A_\tau,$$

$$d\varphi_{i\tau}^c = d\log p_{i\tau} + \mu^c d\log s_{i\tau} - \langle d\log p_{i\tau}\rangle - \mu^c \langle d\log s_{i\tau}\rangle,$$

where we have let $\mu^c \equiv \frac{1}{\sigma^c - 1}$. We can therefore write

$$\begin{aligned}\sum_{i\in V_t^*} s_{i\tau}^* \left(d\varphi_{i\tau} - d\varphi_{i\tau}^c\right) = & \sum_{i\in V_t^*} s_{i\tau}^* \left(\mu_{i\tau} - \mu^c\right) d\log s_{i\tau} + \frac{1}{|O|}\sum_{i\in O} \left(\mu^c - \mu_{i\tau}\right) d\log s_{i\tau} \\ & + \left(\overline{\mu}_\tau^* - \langle \mu_{i\tau}\rangle\right) d\log A_\tau, \\ = & \sum_{i\in V_t^*} s_{i\tau}^* \mu_{i\tau} d\log s_{i\tau}^* + \frac{1}{|O|}\sum_{i\in O} \left(\mu^c - \mu_{i\tau}\right) d\log s_{i\tau}^* \\ & + \left(\overline{\mu}_\tau^* - \langle \mu_{i\tau}\rangle\right)\left(d\log A_\tau + d\log \Lambda_\tau^*\right).\end{aligned}$$

Approximating this integral following the same arguments as in the proof of Corollary 2 leads to

$$\sum_{i \in V_t^*} s_{i\tau}^* \left(\Delta\varphi_{i\tau} - \Delta\varphi_{i\tau}^c\right) \approx \sum_{i \in V_t^*} \overline{\overline{s_{i\tau}^* \mu_{i\tau}}} \Delta \log s_{i\tau}^* + \left\langle \left(\mu^c - \overline{\overline{\mu_{i\tau}}}\right) \Delta \log s_{i\tau}^* \right\rangle$$
$$+ \left(\overline{\overline{\mu}}_t^* - \overline{\overline{\langle \mu_{it} \rangle}}\right) \left(\Delta \log \Lambda_t^* + \Delta \log A_t\right).$$

□

II Derivations for Kimball Specifications

Below, we derive the Kimball functions corresponding to each of the three cases discussed in Section 3.3.2. We have that $\mathcal{E}(\widetilde{q}) \equiv -d\log \mathcal{K}'(\widetilde{q})/d\log \widetilde{q}$. This allows us to integrate the function $\mathcal{E}(\cdot)$ twice to arrive at $\mathcal{K}(\cdot)$.

Klenow-Willis In this case, we have:

$$\psi(\log \widetilde{q}) \equiv \log \mathcal{K}'(\widetilde{q}) = \xi - \frac{1}{\sigma} \int_{-\infty}^{\log \widetilde{q}} e^{\theta v} dv,$$
$$= \xi - \frac{1}{\sigma\theta} \widetilde{q}^{\theta},$$

for any constant ξ. Integrating this expression again, we find:

$$\mathcal{K}(\widetilde{q}) = -e^{\xi} \int_{\log \widetilde{q}}^{\infty} e^{-v^{\theta}/\sigma\theta} dv,$$
$$= e^{\xi} (\sigma\theta)^{\frac{1}{\theta}} \frac{1}{\theta} \Gamma\left(\frac{1}{\theta}, \frac{1}{\sigma\theta} \widetilde{q}^{\theta}\right),$$

where $\Gamma(\cdot, \cdot)$ is the incomplete Gamma function.

Finite-Infinite Limits (FIL) In this case, we have:

$$\psi\left(\log \widetilde{q}\right) \equiv \log \mathcal{K}'\left(\widetilde{q}\right) = \xi - \int_{-\infty}^{\log \widetilde{q}} \frac{dv}{\sigma + \left(\sigma_o - \sigma\right) e^{-\theta v}},$$
$$= -\frac{1}{\sigma} \log \widetilde{q} + \xi - \frac{1}{\sigma\theta} \log\left(\frac{\sigma}{\sigma_o - \sigma} + \widetilde{q}^{-\theta}\right).$$

Next, we integrate to find the expression for $\mathcal{K}\left(\cdot\right)$:

$$\mathcal{K}\left(\widetilde{q}\right) = e^{\xi} \int_{0}^{\log \widetilde{q}} \left(\frac{\sigma v^{\theta} + \sigma_o - \sigma}{\sigma_o - \sigma}\right)^{-\frac{1}{\sigma\theta}} dv,$$
$$= e^{\xi}\, \widetilde{q} \cdot {}_2F_1\left(\frac{1}{\theta}, \frac{1}{\sigma\theta}; 1 + \frac{1}{\theta}; -\frac{\sigma}{\sigma_o - \sigma} \widetilde{q}^{\theta}\right),$$

where ${}_2F_1$ is the hypergeometric function. The functional form above implies the following expression for log demand:

$$d\left(\log \widetilde{p}\right) \equiv \psi^{-1}\left(\log \widetilde{p}\right),$$
$$= \frac{1}{\theta} \log\left[\frac{\sigma_o - \sigma}{\sigma}\left(e^{\theta\sigma\left(\xi - \log \widetilde{p}\right)} - 1\right)\right].$$

In this case, there exists a finite chocke price for any product, above which demand drops to zero.

Finite-Finite Limits (FFL) In this case, we have:

$$\psi\left(\log \widetilde{q}\right) \equiv \log \mathcal{K}'\left(\widetilde{q}\right) = \xi - \int_{-\infty}^{\log \widetilde{q}} \left[\frac{1}{\sigma_o} + \left(\frac{1}{\sigma} - \frac{1}{\sigma_o}\right) \frac{e^{\theta_o} e^{\theta v}}{1 + e^{\theta_o} e^{\theta v}}\right] dv,$$
$$= \xi - \frac{1}{\sigma_o} \log \widetilde{q} - \left(\frac{1}{\sigma} - \frac{1}{\sigma_o}\right) \frac{1}{\theta} \log\left(1 + e^{\theta_o} \widetilde{q}^{\theta}\right).$$

Finally, we integrate to find the expression for $\mathcal{K}(\cdot)$:

$$\mathcal{K}(\widetilde{q}) = e^{\xi} \int_0^{\widetilde{q}} v^{-\frac{1}{\sigma_o}} \left(1 + e^{\theta_o} v^{\theta}\right)^{-\left(\frac{1}{\sigma} - \frac{1}{\sigma_o}\right)\frac{1}{\theta}} dv,$$
$$= e^{\xi} \frac{\sigma_o}{\sigma_o - 1} \widetilde{q}^{1-\frac{1}{\sigma_o}} \cdot {}_2F_1\left(\left(1 - \frac{1}{\sigma_o}\right)\frac{1}{\theta}, \left(\frac{1}{\sigma} - \frac{1}{\sigma_o}\right)\frac{1}{\theta}; 1 + \left(\frac{1}{\sigma_o} + 1\right)\frac{1}{\theta}; -e^{\theta_o}\widetilde{q}^{\theta}\right),$$

where ${}_2F_1$ is the hypergeometric function.

III Inverting Kimball Demand

We implement the demand inversion through the dual problem, meaning that we map the vector of observed expenditure shares $\boldsymbol{s}_t$ to a corresponding vector of normalized quantities $\widetilde{\boldsymbol{q}}_t$. Formally, we solve for the function $\mathcal{D}(\pi_i(\cdot; \boldsymbol{\varsigma}); \boldsymbol{\varsigma})$ corresponding to the definition (1).

To invert the demand, for any collection of $(\boldsymbol{p}_t, \boldsymbol{s}_t)$ at time t, we need to solve for the vector $(\log \widetilde{q}_{it})_i$, such that:

$$\log s_{it} = \log \widetilde{q}_{it} + \psi(\log \widetilde{q}_{it}) - \log\left[\sum_{j \in V_t} \exp\left(\log \widetilde{q}_{jt} + \psi\left(\log \widetilde{q}_{jt}\right)\right)\right], \qquad \forall i \in V_t, \tag{9}$$

$$k(1) = \log\left[\sum_{i \in V_t} \exp\left(k\left(\log \widetilde{q}_{it}\right)\right)\right], \tag{10}$$

where $k(\cdot) \equiv \log \mathcal{K}(\exp(\cdot))$ and $\psi(\cdot) \equiv \log \mathcal{K}'(\exp(\cdot))$. We can rewrite Equation (9) as (assuming $O \equiv \{o\}$):

$$\log\left(\frac{s_{it}}{s_{ot}}\right) = \log\left(\frac{\widetilde{q}_{it}}{\widetilde{q}_{ot}}\right) + \psi(\log \widetilde{q}_{it}) - \psi(\log \widetilde{q}_{ot}), \qquad \forall i \in V_t. \tag{11}$$

Using the identity

$$k'\left(\log \widetilde{q}\right) = \exp\left(\log \widetilde{q} + \psi\left(\log \widetilde{q}\right) - k\left(\log \widetilde{q}\right)\right),$$

we can substitute Equation (11) in Equation (10), we find:

$$\begin{aligned}
k\left(1\right) &= \log\left[\sum_{i\in V_t} \exp\left(k\left(\log \widetilde{q}_{it}\right)\right)\right], \\
&= \log\left[\sum_{i\in V_t} \exp\left(\log \widetilde{q}_{it} + \psi\left(\log \widetilde{q}_{it}\right) - k'\left(\log \widetilde{q}_{it}\right)\right)\right], \\
&= \log\left[\sum_{i\in V_t} \exp\left(\log \widetilde{q}_{ot} + \psi\left(\log \widetilde{q}_{ot}\right) + \log\left(\frac{s_{it}}{s_{ot}}\right) - k'\left(\log \widetilde{q}_{it}\right)\right)\right], \\
&= \log \widetilde{q}_{ot} + \psi\left(\log \widetilde{q}_{ot}\right) + \log\left[\sum_{i\in V_t} \frac{s_{it}}{s_{ot}} \exp\left(-k'\left(\log \widetilde{q}_{it}\right)\right)\right], \\
&= k\left(\log \widetilde{q}_{ot}\right) + \log\left[\sum_{i\in V_t} \frac{s_{it}}{s_{ot}} \exp\left(k'\left(\log q_{ot}\right) - k'\left(\log \widetilde{q}_{it}\right)\right)\right]. \qquad (12)
\end{aligned}$$

We use an iterative approach: starting with some initial guess for $\widetilde{q}_{ot}$, we iterate between updating values of $\widetilde{q}_{it}$ for $i \neq o$ from Equation (11) and updating the value of $\widetilde{q}_{ot}$ from Equation (12).

Appendix J Details on the Auto Data

I Data

TABLE 3.18 – Summary Statistics

	Mean	Std. Dev	Min	Max
Sales	60135.09	87493.58	10	891482
Price ('000 USD)	36.18	17.47	11.14	124.05
Space	1.34	.19	.65	2
Horsepower	.53	.17	.12	1.90
Miles/$	.90	.43	.30	5.84
Luxury	.30	.46	0	1
Sport	.09	.29	0	1
SUV	.23	.42	0	1
Truck	.07	.26	0	1
Van	.06	.24	0	1
Electric	.048	.21	0	1
Observations	9493			

Note: The table displays summary statistics of the main variables of our sample of vehicles. An observation is defined as a model-year pair. Prices are in thousands of current US Dollars. Space is defined as the product between the length and the width of the vehicle in inches divided by one thousand. Horsepower is defined as the horsepower of the vehicle divided by its curbweight. Miles-per-dollar is scaled down by a factor of 10. The Electric dummy refers to EV (electric vehicles), PHEV (plug-in hybrid electric vehicles) and HEV (hybrid electric vehicles).

II Testing the Identification Assumption

We are able to test the identification assumption in Equation (8) leveraging the additional data on product characteristics available for the US auto market. The identification assumption relies on the orthogonality between demand shocks innovations, u_{it}, and lagged log prices and quantities. Under the assumption in Equation (6), the identification assumption between demand shocks innovations and lagged log prices can be rewritten as:

$$\mathbb{E}\left[\varphi_{it}|g_i\left(\varphi_{it-1};\boldsymbol{\varrho}\right),\log p_{it-1}\right] \;=\; g_i\left(\varphi_{it-1};\boldsymbol{\varrho}\right)+\alpha\,\log p_{it-1}.$$

where α is expected to be equal to zero when the orthogonality condition holds. Under the assumption that the demand shock process is a stationary AR(1) process, $g_i\left(\varphi_{it-1};\boldsymbol{\varrho}\right)\equiv\rho\varphi_{it-1}+(1-\rho)\,\phi_i$ as in Equation (7), we use the set of characteristics available in our dataset as a proxy for φ_{it} and test whether the current value of product characteristics are correlated to lagged log prices after controlling for lagged characteristics. In other words, for each characteristic k, we estimate the following specification:

$$x_{kit}=\alpha\log p_{it-1}+\boldsymbol{\rho}_k'\,\boldsymbol{x}_{it-1}+\eta_t+\gamma_i+\epsilon_{it}, \tag{1}$$

where $\boldsymbol{x}_{it-1}$ is the entire set of lagged product characteristics. Table 3.19 reports the set of coefficients estimated using Equation (1). No estimated $\widehat{\alpha}$ coefficients are statistically different from zero, validating our identification assumption. Moreover, all product characteristics exhibit a strong degree of autocorrelation, supporting our choice for the process of demand shocks.[159] We also standardize all variables and re-estimate Equation (1) in order to compare the coefficient of lagged price to the coefficients of lagged characteristics in terms of magnitude. Lagged product characteristics still exhibit strong and significant correlations, while lagged prices are not correlated to

[159] The only exception is Truck, which exhibits a weak autocorrelation.

current product characteristics.

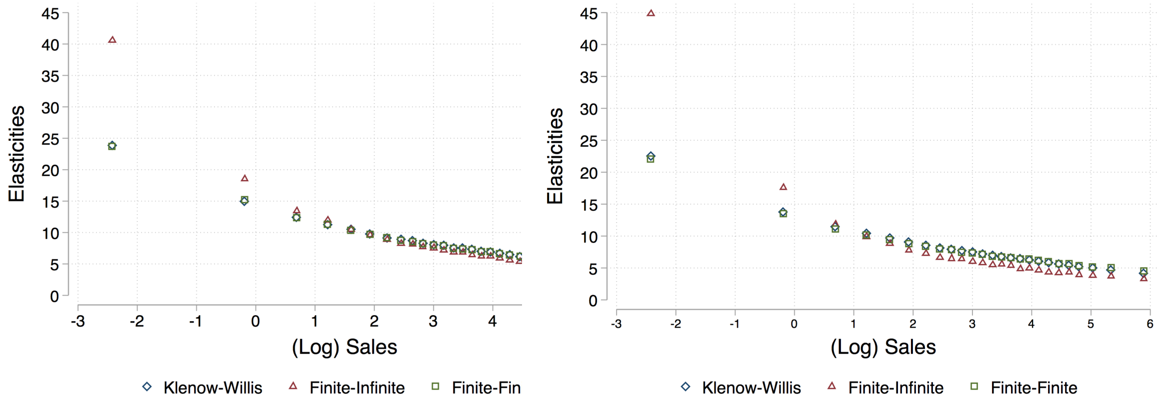
TABLE 3.19 – Testing the Identification Assumption

	Horse Power		Space		Miles/Dollar		Suv		Van		Truck	
Lagged	Level	Z-score	Level	Z-score	Level	Z-score	Level	Z-score	Level	Z-score	Level	Z-score
Price	0.0057	0.015	0.013	0.030	-0.024	-0.024	0.022	0.023	-0.015	-0.027	0.0092	0.016
	(0.0088)	(0.023)	(0.0078)	(0.018)	(0.019)	(0.019)	(0.027)	(0.028)	(0.011)	(0.019)	(0.0092)	(0.016)
Horse Power	0.66	0.66	-0.0040	-0.0035	0.044	0.017	-0.022	-0.0086	0.019	0.013	-0.012	-0.0079
	(0.023)	(0.023)	(0.0098)	(0.0085)	(0.025)	(0.0098)	(0.021)	(0.0086)	(0.016)	(0.011)	(0.011)	(0.0076)
Space	-0.0097	-0.011	0.63	0.63	-0.069	-0.031	-0.016	-0.0073	0.029	0.023	0.0052	0.0039
	(0.015)	(0.017)	(0.030)	(0.030)	(0.029)	(0.013)	(0.038)	(0.018)	(0.020)	(0.016)	(0.011)	(0.0086)
Miles/Dollar	0.014	0.037	-0.011	-0.025	0.53	0.53	-0.00078	-0.00080	-0.00054	-0.00095	-0.0012	-0.0020
	(0.0084)	(0.021)	(0.0057)	(0.013)	(0.055)	(0.055)	(0.0088)	(0.0090)	(0.0047)	(0.0083)	(0.0036)	(0.0062)
Suv	-0.0070	-0.017	0.0094	0.020	0.027	0.026	0.21	0.21	0.013	0.023	-0.034	-0.056
	(0.0043)	(0.011)	(0.0055)	(0.012)	(0.011)	(0.010)	(0.065)	(0.065)	(0.026)	(0.044)	(0.018)	(0.030)
Van	0.0086	0.013	0.0056	0.0071	0.013	0.0074	-0.017	-0.0098	0.18	0.18	-0.12	-0.12
	(0.0099)	(0.014)	(0.0038)	(0.0048)	(0.0074)	(0.0042)	(0.031)	(0.018)	(0.072)	(0.072)	(0.040)	(0.038)
Truck	-0.0012	-0.0018	0.014	0.019	0.0042	0.0025	-0.14	-0.086	0.10	0.11	0.10	0.10
	(0.0065)	(0.0098)	(0.0064)	(0.0085)	(0.022)	(0.013)	(0.077)	(0.047)	(0.058)	(0.060)	(0.090)	(0.090)
Observations	8268	8268	8268	8268	8268	8268	8268	8268	8268	8268	8268	8268
Model & Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table reports the coefficients estimated using Equation (1). Each column refers to a given product characteristics. We consider horsepower, space, miles-per-dollar, truck, van and suv. For each characteristic, Equation (1) is estimated using level or z-score variables. Z-score variable refers to the set of coefficients estimated using Equation (1) after standardizing all variables. Standard errors are clustered at the model level.

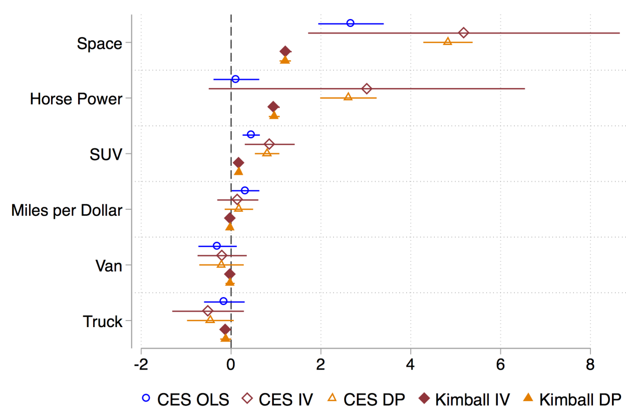
III Additional Tables and Figures

FIGURE 3.27 – Comparison across Kimball Specifications



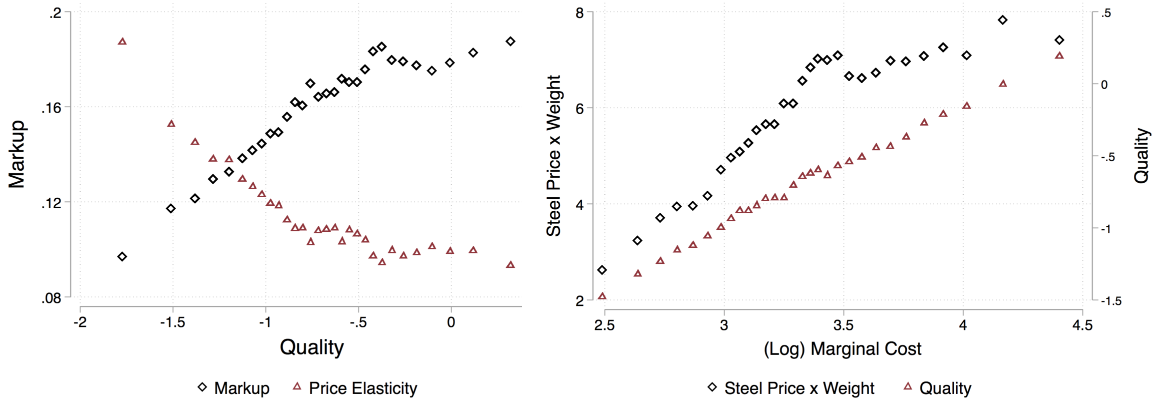
Note: The left panel shows a binscatter representation of the relationship between (log) sales and the Kimball price elasticities estimated using the DP approach. The right panel shows the relationship between (log) sales and Kimball price elasticities estimated using the IV approach. All three Kimball specifications (Finite-Finite, Finite-Infinite, and Klenow-Willis) are considered.

FIGURE 3.28 – Correlation between Inferred Quality and Product Characteristics



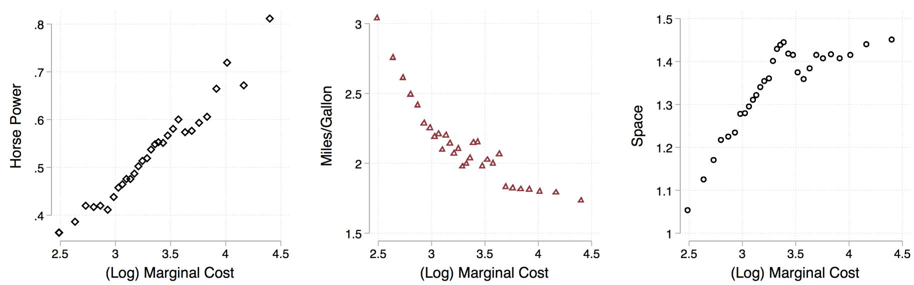
Note: The figure reports the relationship between product characteristics and quality inferred using Equation (5) for CES and Kimball demand systems. The coefficients referring to the DP approach (CES and Kimball) and the Kimball IV case are obtained from regression in Equation (7). We consider the following product characteristics: horsepower, space, miles-per-dollar and style (suv, truck, van). The coefficients referring to the OLS and IV estimates of the CES specification are obtained from Equation (6), where product characteristics are used to proxy for quality. All regressions use the entire sample and includes time and producer fixed effects. Standard errors are clustered at producer level, the bands around the estimates show the 95% confidence intervals.

FIGURE 3.29 – Markups and Marginal Cost



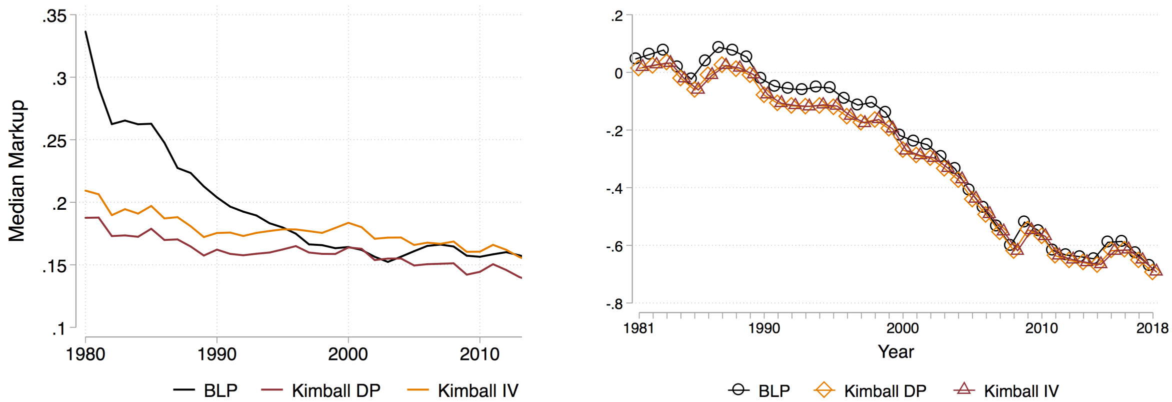
Note: The left panel shows the relationship between the measure of inferred quality and the price elasticity estimated from the Finite-Finite Kimball specification using the DP approach. Markups are computed under the assumption of monopolistic competition, $\mu_{it} = \frac{1}{\sigma_{it}-1}$, where σ_{it} is the estimated price elasticity. The right panel shows the relationship between: i) the implied marginal cost and a proxy of input costs; ii) the implied marginal cost and the measure of inferred quality estimated from the Finite-Finite Kimball specification using the DP approach. The marginal cost of each model is inferred as follow: $mc_{it} = \frac{p_{it}}{1+\mu_{it}}$. The input costs proxy is created multiplying the price of steel to the weight of each vehicle.

FIGURE 3.30 – Marginal Cost and Product Characteristics



Note: Each panel shows the relationship between the inferred marginal cost and a product characteristic. We consider horsepower (left), space (center) and miles-per-gallon (right). Marginal cost is inferred from $mc_{it} = \frac{p_{it}}{1+\mu_{it}}$, where μ_{it} is the markup computed under the assumption of monopolistic competition using the price elasticities estimated from the Finite-Finite Kimball specification using the DP approach.

FIGURE 3.31 – Trends in Markups and Marginal Cost



Note: The left panel shows the evolution of the median markup over the period 1980-2018. Markups are computed under the assumption of monopolistic competition, $\mu_{it} = \frac{1}{\sigma_{it}-1}$, where σ_{it} is the estimated price elasticity. The BLP and Finite-Finite Kimball specifications are considered. The right panel shows the estimated trend in the real marginal cost. The real marginal cost is computed from $mc_{it} = \frac{p_{it}}{1+\mu_{it}}$ and deflated using the CPI. The trend in the marginal cost is obtained regressing the inferred marginal cost at the model-year level on product characteristics and a time trend.

TABLE 3.20 – Ideal Price Index for the US Auto Market: CES vs Kimball

	Total		Decomposition				
			Price	Quality		Variety	
	Kimball	CES		Kimball	CES	Kimball	CES
Cumulative Change (%)	-127.3	-269.0	-60.1	-49.6	-175.9	-17.5	-32.9
Annual Change (%)	-3.35	-7.08	-1.58	-1.31	-4.63	-0.46	-0.87

Note: The Table reports the cumulative and the average annual change in the ideal import price indices for the auto market over the period 1980-2018 and its decomposition into the price, quality and variety channels. Prices are deflated using the CPI index from BLS. Quality is normalized such that the average change in quality of the set of continuing models that are not redesigned is zero. The price index is computed for both the Kimball and the CES specifications, estimated using the DP approach.

Appendix K Details on the US Import Data

I Further Examination of CES Estimates

Price Elasticities Across Different Levels of Aggregation Table 3.21 reports the mean and the median of the estimated elasticities using the DP approach for three different levels of product aggregation. As expected, we find lower elasticities when we aggregate products in broader categories. The average elasticity is 4.5 at the SITC3 level and it increases to 5.6 at the HS10 level. Even if the differences appear small, we can statistically reject the null hypothesis that the mean elasticities are the same across all level of aggregations. Note also that the median elasticities of substitution exhibit the same qualitative pattern, as their values increase from 2.9 to 3.4. The median estimates at more aggregate levels (three and five digit) statistically differ from the most disaggregated level.[160]

TABLE 3.21 – CES Elasticities based on the DP Approach at Different Levels of Aggregation

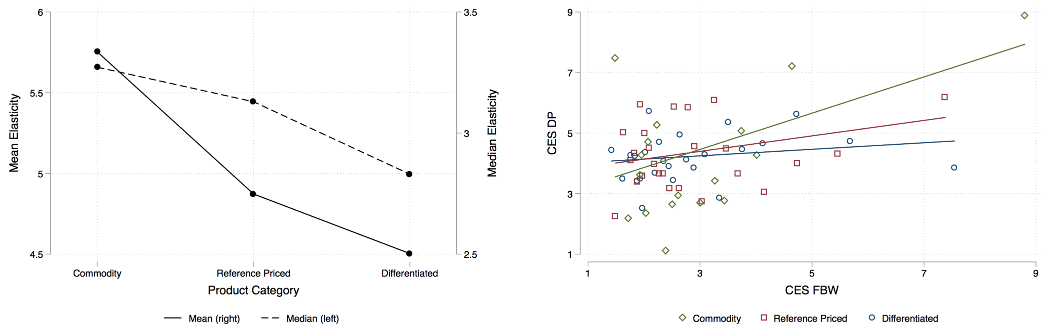
	HS10	SITC5	SITC3
Mean	5.65	5.09	4.49
(SE)	(0.09)	(0.21)	(0.40)
Median	3.37	3.13	2.87
(SE)	(0.05)	(0.10)	(0.23)
N	8508	1296	147
T-statistics		2.493	2.836
Pearson χ^2 p-value		0.043	0.025

Note: Mean and median of the elasticities of substitution estimated with the DP approach for the products defined at the HS10, SITC5 and SITC3 levels of aggregation. Only feasible estimates are reported. Values above 130 are censored. Standard errors for each statistics are bootstrapped. T-statistics refer to a t-test for differences in mean with respect to the HS10 level; p-values for Pearson difference in median tests with respect to the HS10 level.

Price Elasticities Across Different Rauch (1999) Product Classes We use the Rauch (1999) classification to distinguish products at the SITC4 level into three categories: commodities, referenced priced, and differentiated

[160] In contrast to the case of the mean estimates, we cannot statistically reject the hypothesis that the medians are the same at the SITC3 and SITC5 level.

FIGURE 3.32 – DP Elasticities and Rauch Conservative Classification



Note: The left panel displays the mean and the median of the elasticities of substitution estimated with the DP approach for each category of the Rauch Conservative Classification at the SITC4 level of aggregation. The right panel shows the correlation between the DP and FBW estimates for each category of the Rauch Conservative Classification at the SITC4 level of aggregation.

goods. Rauch (1999) provides two distinct classifications, "Liberal" and "Conservative", that only differ in a few products that can be classified in multiple ways. The left panel of Figure 3.32 shows both the mean and the median elasticity for each Rauch Conservative category. Both these statistics are ranked in increasing order between commodities, referenced priced, and differentiated products, as expected. We can reject the hypothesis that the combined set of commodities and referenced priced goods have the same mean or median than differentiated products.[161] Table 3.22 reports the corresponding values and their standard errors for Figure 3.32 and show that qualitative results holds also for the Liberal version of the classification.

In addition, again using the classification proposed by Rauch (1999), we can show that the quality bias in the conventional estimates is stronger among more differentiated products. Intuitively, quality differentiation is

[161] We statistically test the difference between differentiated products and the remaining categories pooled together. Differences are not statistically significant if the two categories are considered individually.

less likely among homogeneous goods, suggesting that the DP estimates in this case should on average be closer to, and more correlated with, the conventional estimates. Consistently with this intuition, the right panel of Figure 3.32 shows that the correlation between DP and FBW is stronger for commodities and the average difference between the two sets of estimates is smaller. As we consider less homogenous categories, referenced priced and differentiated products, the average quality bias increases while the correlation decreases.[162] Figure 3.33 shows that the qualitative pattern is robust to how products are grouped between homogenous and differentiated.

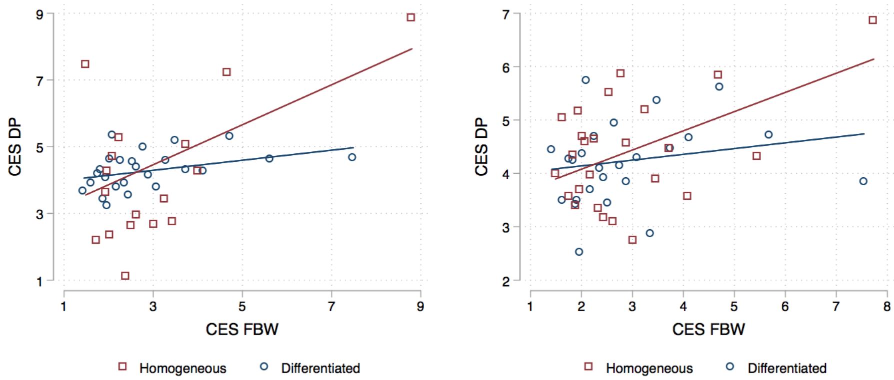
TABLE 3.22 – DP Estimates: Rauch Classifications

	Commodity	Reference Priced	Differentiated		Commodity	Reference Priced	Differentiated
Mean	5.75	4.87	4.50	Mean	5.28	4.77	4.58
(SE)	(0.86)	(0.42)	(0.25)	(SE)	(0.63)	(0.42)	(0.27)
Median	3.27	3.13	2.83	Median	3.24	3.10	2.82
(SE)	(0.69)	(0.18)	(0.18)	(SE)	(0.37)	(0.18)	(0.21)
N	50	168	317	N	75	162	298

Note: For each category of the Rauch Classification (commodity, reference priced and differentiated), the tables report the mean and the median CES elasticity estimated using the DP approach at the SITC4 level. The left panel refers to the Conservative version of the classification (corresponding to Figure 3.32 in the main text) while the right one to the Liberal version. It can be show that differences in mean and median are statistically significant at standard levels if the more homogeneous categories (commodities and reference priced) are pooled together and compared to differentiated products.

[162] The average difference between group captures the average quality bias and is represented by the intercept of a linear regression (fitting line). The slope would capture instead the correlation across estimates.

FIGURE 3.33 – Correlation DP and FBW, Different Pooling of Rauch Categories



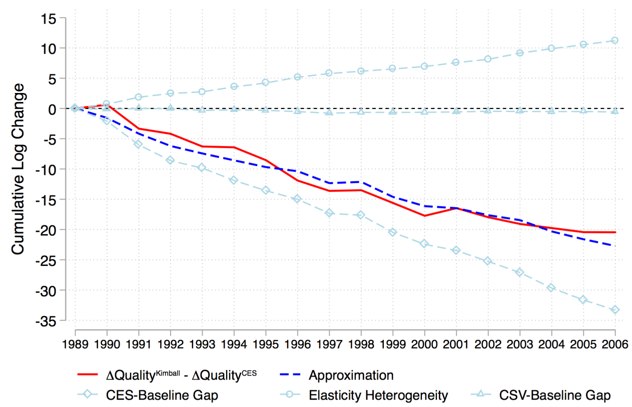
Note: The figure shows the correlation between the estimated elasticities using the DP and FBW methods at the SITC4 level using alternative breakdowns across products. Conservative Rauch classification is used. In the left panel, homogeneous products are defined as commodities only while, in the right panel, they include commodieties and reference priced goods.

II Further Results on Welfare and Quality Decomposition

II.1 Bias in Inferred Quality: CES vs. Kimball

Proposition 2 provides a decomposition of the gap between what Kimball and CES demand systems predict about the contribution of quality change to the aggregate price index. This gap is the sum of three terms: the gap in the love-of-variety proxies inferred by the two demand systems (CES-Baseline Gap), the contribution of reallocations of expenditure across products (Elasticity Heterogeneity) and the heterogeneity in own-price elasticities and the love-of-variety proxies between the common set of varieties and the set of baseline products (CVS-Baseline Gap).

FIGURE 3.34 – Quality Contribution: Kimball vs CES



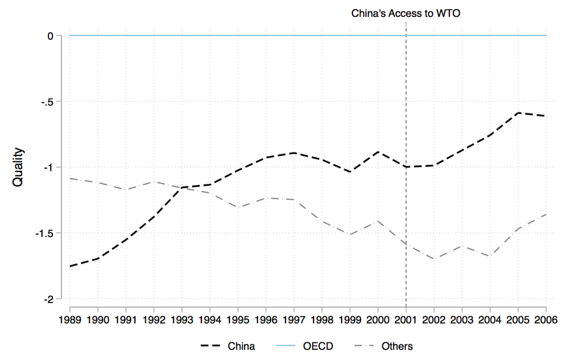
Note: The figure plots the decomposition of the gap in the Torqvist-weighted mean quality change between the inferred quality using Kimball and that under CES. The solid red line represents the estimate Kimball-CES gap in aggregate quality change. The dashed blue line represents the approximation of the gap according Proposition 2. The approximation is the sum of three components: the gap in the love-of-variety proxies (CES-Baseline Gap, diamonds line), the contribution of reallocations of expenditure across products (Elasticity Heterogeneity, circles line) and the heterogeneity in own-price elasticities and the love-of-variety proxies between the common set varieties and the set of baseline products (CVS-Baseline Gap, triangles line).

Figure 3.34 shows the cumulative gap between Kimaball and CES and its decomposition into the the three components based on Proposition 2. The contribution of the first term, the gap in the love-of-variety estimates between CES and the baseline varieties under Kimball, is negative and explains more than 100% of the gap. Since the market share of OECD countries within the common set of varieties is falling over time, the key reason for the overestimation of the contribution of quality by CES is simply that its estimated elasticities suffer from a downward heterogeneity bias. The contribution of the second term, the reallocation within the common set of varieties, is positive, suggesting that there are reallocations toward varieties with low price elasticities within each sector over time. Finally, the last term, which is the gap in elasticities between the common set of varieties and the baseline varieties, appears fairly small. The dashed blue line shows the sum of all the three terms in the approximation, which is fairly close to the overall gap implied by the estimated Kimball and CES specifications (red line).

II.2 Quality Decomposition

Figure 3.35 shows the evolution of the expenditure-weighted quality for each (group of) exporter(s), China, OECD economies and all other countries. The (expenditure-weighted) average quality of Chinese varieties has increased constantly since 1989 relative to the average OECD quality, which is normalized to zero over the entire time period. This supports the extensive evidence that Chinese goods have undergone a sophistication process, catching up with more advanced economies and largely contributing to the aggregate quality improvement of US imports.

FIGURE 3.35 – Decomposition of Quality across Countries



Note: The figure shows the evolution of the (expenditure weighted) average quality of each (group of) exporter(s), China, OECD economies and rest of the world. OECD (expenditure weighted) average quality is normalized to zero for exposition.

Table 3.23 shows that import quality has increased by around 28% over the time period from 1989 to 2006. This increase is exclusively driven by a rise in quality within each (group of) exporter(s) while compositional changes between exporters partially offset the within forces. This is consistent with the fact that Chinese products gained market share over the time period but still have lower quality compared to other exporters, even if they are catching up with the frontier. Notice also that the annual increase in quality is larger after China joined the WTO in 2001, suggesting that the trade liberalization shock boosted the sophistication process even more.

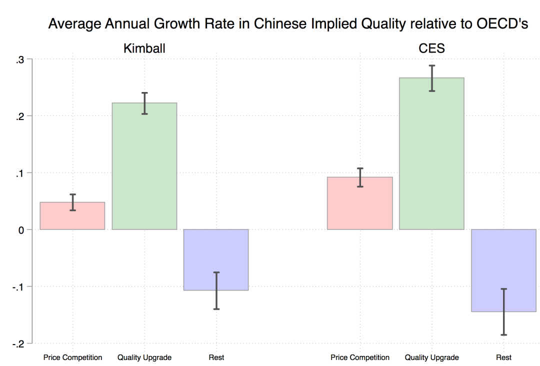
TABLE 3.23 – Between and Within Decomposition

	$\Delta\varphi$	Δ within	Δ between
Full Sample	0.283	0.424	-0.141
Before 2001	0.159	0.227	-0.068
After 2001	0.124	0.197	-0.073

Note: The Table shows a decomposition of the growth in aggregate product quality between and within exporters. We consider China, OECD economies and all the other exporters pool together. For each exporter, we compute the aggregate product quality as the expenditure-weighted average across varieties.

Finally, we also check how the quality of Chinese varieties relative to OECD's evolved for each category defined in Figure 3.43, quality upgrading, price competition and others. For each product category (HS8), we compute the average annual change in inferred quality of Chinese varieties relative to the set of advanced economies used for Figure 3.43. Consistent with their definition, the change in quality of Chinese products labelled as "quality upgrading" is four times larger than the change in quality of Chinese products labelled as "price competition". This confirms the intuition that quality improvements represent the key mechanism to explain the increase in Chinese import penetration and the simultaneous rise in relative prices. The specification of demand has first-order effect on our measurement of the role of quality as the quality improvements inferred from CES are larger for all three categories (right panel).

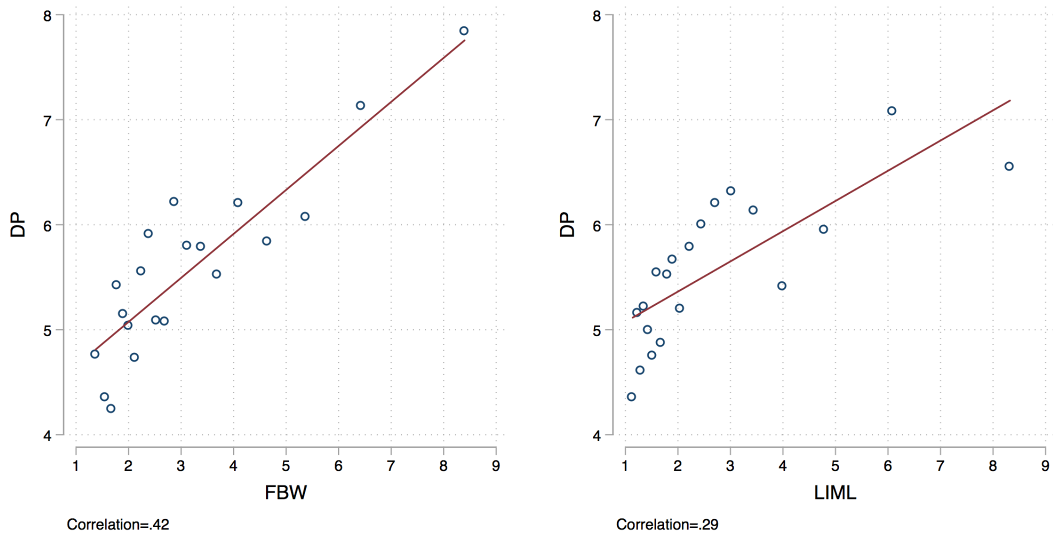
FIGURE 3.36 – Inferred Quality: Quality Upgrading and Price Competition



Note: The figure shows the average annual product quality growth rate across Chinese varieties defined at the HS8 level, relative to the average annual growth rate of the corresponding OECD variety. Left (right) panel uses inferred quality from Kimball (CES) specification. Product categories "Quality Upgrade", "Price Competition", and "Rest" are defined in Figure 3.43.

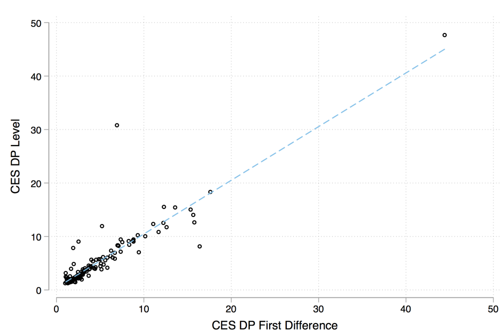
III Additional Tables and Figures

FIGURE 3.37 – Correlation between DP and FBW or LIML Estimates, HS10 level



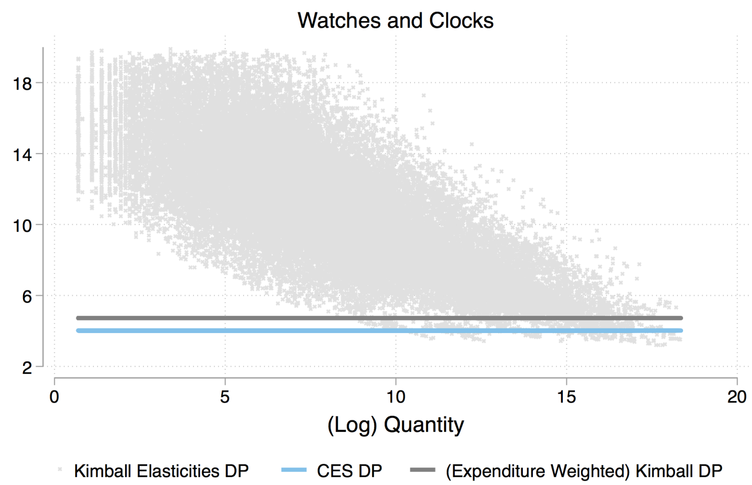
Note: The figure shows the binscatter plot of the relationship between the estimated elasticities using the DP approach and conventional methods like FBW (right panel) and LIML (left panel). The figures refers to the set of estimates at the HS10 level. Elasticities are censored at 10.

FIGURE 3.38 – Comparison CES Estimate: Level vs First Difference Moment



Note: The figure shows the correlation between the estimated CES elasticities obtained using CES as the limiting Kimball moment ($\sigma_o \equiv \sigma$) and the first difference moment used for the elasticities reported in Table 3.21. Dashed line represents the 45 degrees line.

FIGURE 3.39 – CES-Kimball Elasticity: Watch and Clocks



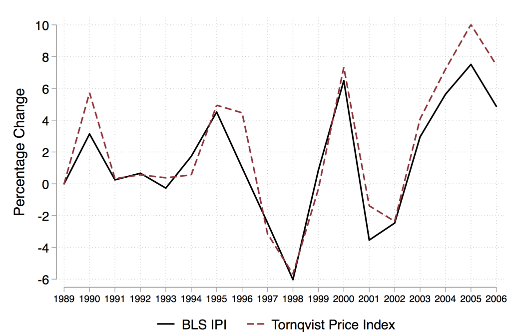
Note: The figure shows the entire set of Kimball price elasticities of each variety-time pair, σ_{it}, as a function of the (log) quantity imported for the sector Watches and Clocks (SITC3 884). The gray line represents the expenditure-weighted mean Kimball price elasticity while the blue line represents the CES estimated elasticity for the sector.

TABLE 3.24 – Kimball Parameters

	σ	σ_o	θ
Mean	1.99	651.2	0.73
	(0.13)	(156.1)	(0.23)
Median	1.12	8.13	0.16
	(0.054)	(1.12)	(0.017)

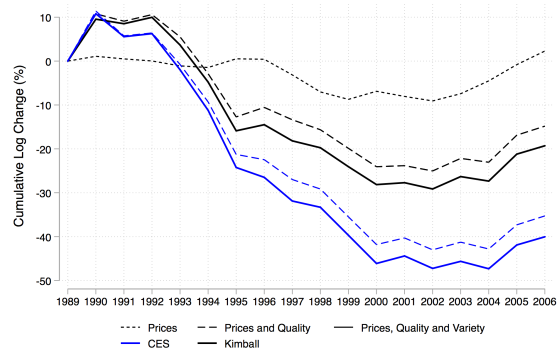
Note: The table displays the mean and the median across all SITC3, with the corresponding standard errors, of the estimated parameters of the Finite-Finite Kimball specification.

FIGURE 3.40 – Comparison with BLS Import Price Index



Note: The figure plots the year-to-year change in the BLS Import Price Index and a the price component of the aggregate import price constructed using the Tornqvist approximation.

FIGURE 3.41 – Dynamics of US Import Price Index - PPI Deflated



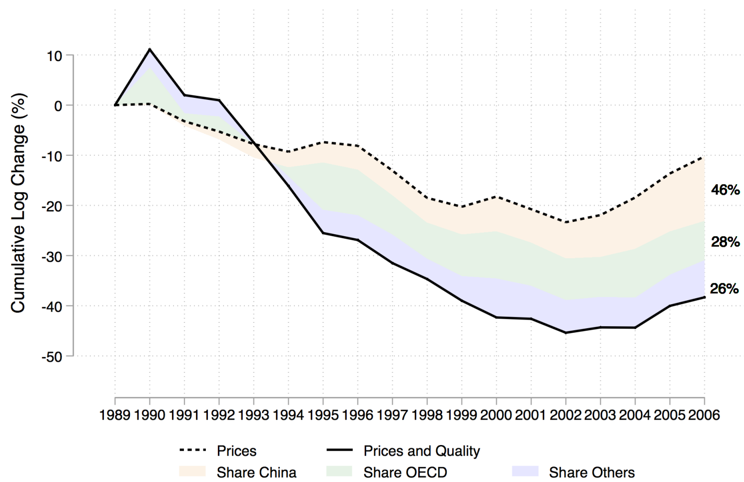
Note: The figure plots the aggregate import price indices for both the CES and the Kimball specifications and their decomposition into the price, quality and variety components, according Equations (19) and (23). Prices are deflated using the PPI index from BLS. The measure of inferred quality is normalized such that the average quality of the set of OECD varieties is zero. The solid lines represent the aggregate import price index including all three components. The dashed and dotted lines represent the price and quality components together and the price component only, respectively. Black (Blue) lines refer to the Kimball (CES) specification.

TABLE 3.25 – Welfare Gains from Trade - PPI Deflated

	Total		Decomposition				
			Price	Quality		Variety	
	Kimball	CES		Kimball	CES	Kimball	CES
Cumulative Change (%)	-19.3	-40.0	2.29	-17.1	-37.6	-4.48	-4.76
Annual Change (%)	-1.07	-2.22	0.13	-0.95	-2.09	-0.25	-0.26

Note: The Table reports the cumulative and the average annual change in the aggregate import price indices defined in Equations (**??**) and (**??**) and reported in Figure 3.41, and their decomposition. Prices are deflated using the PPI index from BLS. The measure of inferred quality is normalized such that the average quality of the set of OECD varieties is zero.

FIGURE 3.42 – Price Index, Decomposition of Quality across Countries: CES case



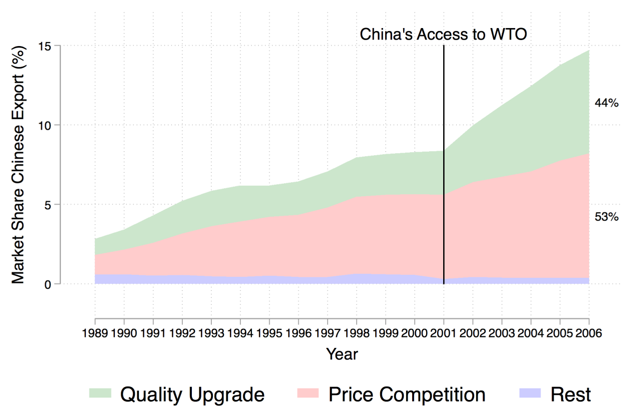
Note: The dashed line figure shows the price component of the aggregate import price index. The solid line shows the price component and the quality component of the aggregate import price index. The quality contribution is computed using the inferred quality from the CES specification. The difference between these two lines quantifies the role of product quality change and is decomposed into the role of Chinese varieties (orange area), OECD varieties (green area) and all other varieties pooled together (purple area).

Appendix L Additional Tables and Figure

I *Examining the Share of China in US Imports*

Figure 3.43 shows that the evolution of the aggregate import share of Chinese products, decomposing the change in the import share into three categories. We distinguish the market share of those products whose prices and market share have both increased relative to a set of benchmark origin countries ("quality upgrade" products), the market share of those products with rising market share but falling relative price ("price competition" products), and the market share of those products with falling market share and relative price ("rest" products). The aggregate import share of Chinese products increased up to 15% in 2006. Around 46% of the growth in the aggregate import share of Chinese products over the period 1989-2006 stems from the contributions of the first group ("quality upgrade" products), which represent 44% of the value of Chinese imports to the US by the end of this period.

FIGURE 3.43 – Decomposition of Chinese Export: Quality Upgrade and Price Competition

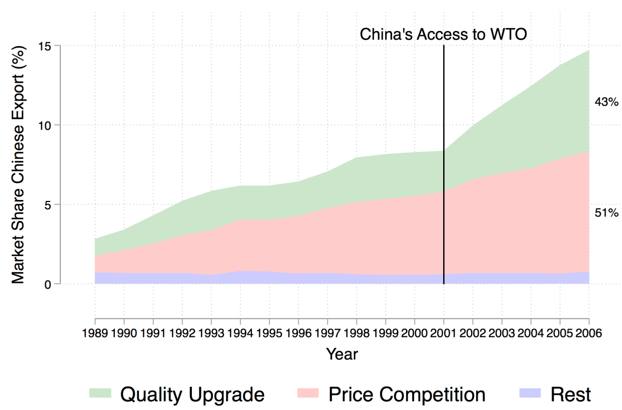


Note: The figure shows the decomposition of Chinese import share into three categories: Quality Upgrade, Price Competition and Rest. Quality Upgrade represents the market share of Chinese products whose both market share and relative price increased over time on average. Price Competition includes those products whose market share increased but relative price declined over time. The third category (Rest) includes the remaining products, that is, products whose market share declined over time. Relative prices are defined with respect to the average price across the varieties imported from a set of advanced economies, including Canada, Japan, Germany, United Kingdom, Switzerland, Italy, France, Belgium, Netherland, Spain, Austria, Denmark, Finland, Portugal, Sweden, Norway, Ireland, Iceland, Greece, Australia and New Zealand. Products classified at the 8-digit level of the Harmonized System classification.

Relative prices are defined with respect to the average price across the varieties imported from a set of advanced economies, including Canada, Japan, Germany, United Kingdom, Switzerland, Italy, France, Belgium, Netherland, Spain, Austria, Denmark, Finland, Portugal, Sweden, Norway, Ireland, Iceland, Greece, Australia and New Zealand. Products classified at the 8-digit level of the Harmonized System classification. Below, we show that the pattern in Figure 3.43 holds quantitatively when products are classified at the 10-digit level of the HS classification and considering alternative basket of countries as benchmark, such as all OECD economies, all advanced economies as classified by the IMF and individual countries like Germany or Japan.

Figure 3.45 and Table 3.28 show that the same pattern is stronger in industries where product quality and differentiation may play a stronger role, such as Machinery and Transportation. In addition, Table 3.26 shows that most of the growth due to quality upgrade took place after China's access to WTO in 2001.

FIGURE 3.44 – Decomposition Chinese Export, 10-digit level product codes



Note: The figure shows the decomposition of Chinese import share into three categories: Quality Upgrade, Price Competition and Rest. Quality Upgrade represents the market share of Chinese products whose both market share and relative price increased over time on average. Price Competition includes those products whose market share increased but relative price declined over time. The third category (Rest) includes the remaining products, that is, products whose market share declined over time. Relative prices are defined with respect to the average price across the varieties imported from a set of advanced economies, including Canada, Japan, Germany, United Kingdom, Switzerland, Italy, France, Belgium, Netherland, Spain, Austria, Denmark, Finland, Portugal, Sweden, Norway, Ireland, Iceland, Greece, Australia and New Zealand. Products classified at the 10-digit level of the Harmonized System classification.

TABLE 3.26 – Decomposition Chinese Export, Pre and Post 2001

	Full Sample	Before 2001	After 2001
Yearly Market Share Growth Rate	10.5	9.90	12.0
Share Quality Upgrade	46.1	31.7	58.7
Share Price Competition	55.6	73.4	39.9
Rest	-1.67	-5.13	1.37

Notes: Quality Upgrade, Price Competition and Rest. Quality Upgrade represents the market share of Chinese products whose both market share and relative price increased over time on average. Price Competition includes those products whose market share increased but relative price declined over time. The third category (Rest) includes the remaining products, that is, products whose market share declined over time. Relative prices are defined with respect to the average price across the varieties imported from a set of advanced economies, including Canada, Japan, Germany, United Kingdom, Switzerland, Italy, France, Belgium, Netherland, Spain, Austria, Denmark, Finland, Portugal, Sweden, Norway, Ireland, Iceland, Greece, Australia and New Zealand. Products classified at the 8-digit level of the Harmonized System classification.

FIGURE 3.45 – Decomposition Chinese Export, by Sector

Notes: Quality Upgrade, Price Competition and Rest. Quality Upgrade represents the market share of Chinese products whose both market share and relative price increased over time on average. Price Competition includes those products whose market share increased but relative price declined over time. The third category (Rest) includes the remaining products, that is, products whose market share declined over time. Relative prices are defined with respect to the average price across the varieties imported from a set of advanced economies, including Canada, Japan, Germany, United Kingdom, Switzerland, Italy, France, Belgium, Netherland, Spain, Austria, Denmark, Finland, Portugal, Sweden, Norway, Ireland, Iceland, Greece, Australia and New Zealand. Products classified at the 8-digit level of the Harmonized System classification. Food & Chemicals refers to the one digit industries 0 to 5 of SITC classification pooled together, Manufacturing Materials to industry 6, Machinery to industry 7, Miscellaneous Manufacture to industry 8. Industry 9 (Miscellanea) is dropped.

TABLE 3.27 – Decomposition Chinese Export, by Sector

	Aggregate	Food & Chemicals	Manufacturing Materials	Machinery	Misc Manufacture
Yearly Market Share Growth Rate	10.5	3.15	12.6	18.2	8.46
Share Quality Upgrade	46.1	62.2	51.3	54.2	27.3
Share Price Competition	55.6	118.9	51.0	45.1	71.7
Rest	-1.67	-81.1	-2.25	0.70	0.96

Notes: Quality Upgrade, Price Competition and Rest. Quality Upgrade represents the market share of Chinese products whose both market share and relative price increased over time on average. Price Competition includes those products whose market share increased but relative price declined over time. The third category (Rest) includes the remaining products, that is, products whose market share declined over time. Relative prices are defined with respect to the average price across the varieties imported from a set of advanced economies, including Canada, Japan, Germany, United Kingdom, Switzerland, Italy, France, Belgium, Netherland, Spain, Austria, Denmark, Finland, Portugal, Sweden, Norway, Ireland, Iceland, Greece, Australia and New Zealand. Products classified at the 8-digit level of the Harmonized System classification. Food & Chemicals refers to the one digit industries 0 to 5 of SITC classification pooled together, Manufacturing Materials to industry 6, Machinery to industry 7, Miscellaneous Manufacture to industry 8. Industry 9 (Miscellanea) is dropped.

TABLE 3.28 – Decomposition Chinese Export: Pre and Post 2001, by Sector

	Food & Chemicals	Manufacturing Materials	Machinery	Misc Manufacture
Full Sample				
Yearly Market Share Growth Rate	3.15	12.6	18.2	8.46
Share Quality Upgrade	62.2	51.3	54.2	27.3
Share Price Competition	118.9	51.0	45.1	71.7
Rest	-81.1	-2.25	0.70	0.96
Before 2001				
Yearly Market Share Growth Rate	1.50	11.4	16.5	8.32
Share Quality Upgrade	125.8	52.9	31.5	22.5
Share Price Competition	272.7	54.1	69.6	77.2
Rest	-298.4	-7.02	-1.06	0.33
After 2001				
Yearly Market Share Growth Rate	7.08	15.5	22.5	8.77
Share Quality Upgrade	40.1	50.1	65.1	32.9
Share Price Competition	65.7	48.9	33.3	65.4
Rest	-5.88	1.04	1.54	1.68

Notes: Quality Upgrade, Price Competition and Rest. Quality Upgrade represents the market share of Chinese products whose both market share and relative price increased over time on average. Price Competition includes those products whose market share increased but relative price declined over time. The third category (Rest) includes the remaining products, that is, products whose market share declined over time. Relative prices are defined with respect to the average price across the varieties imported from a set of advanced economies, including Canada, Japan, Germany, United Kingdom, Switzerland, Italy, France, Belgium, Netherland, Spain, Austria, Denmark, Finland, Portugal, Sweden, Norway, Ireland, Iceland, Greece, Australia and New Zealand. Products classified at the 8-digit level of the Harmonized System classification. Food & Chemicals refers to the one digit industries 0 to 5 of SITC classification pooled together, Manufacturing Materials to industry 6, Machinery to industry 7, Miscellaneous Manufacture to industry 8. Industry 9 (Miscellanea) is dropped.