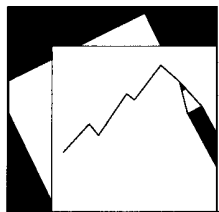


# Working Paper

---

INTERNATIONAL MONETARY FUND



# IMF Working Paper

---

## Elasticity Optimism

*Jean Imbs and Isabelle Mejean*

## IMF Working Paper

Research Department

### Elasticity Optimism

**Prepared by Jean Imbs and Isabelle Mejean<sup>1</sup>**

Authorized for distribution by Gian Maria Milesi-Ferretti

December 2009

#### Abstract

**This Working Paper should not be reported as representing the views of the IMF.**

The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate.

In most macroeconomic models, the substitutability between domestic and foreign goods is calibrated using aggregated data. This imposes homogeneous elasticities across goods, and the calibration is only valid under this assumption. If elasticities are heterogeneous, the aggregate substitutability is a weighted average of good-specific elasticities, which in general cannot be inferred from aggregated data. We identify structurally the substitutability in US goods using multilateral trade data. We impose homogeneity, and find an aggregate elasticity similar in value to conventional macroeconomic estimates. It is more than twice larger with sectoral heterogeneity. We discuss the implications in various areas of international economics.

JEL Classification Numbers: F41, F32, F21

Keywords: Trade Elasticities, Heterogeneity, Calibration, Global Imbalances, International Portfolio, Monetary Policy

Author's E-Mail Addresses: [jimbs@unil.ch](mailto:jimbs@unil.ch), [imejean@imf.org](mailto:imejean@imf.org)

---

<sup>1</sup> We thank the Fondation Banque de France and the National Center of Competence in Research “Financial Valuation and Risk Management” for financial support. The National Centers of Competence in Research (NCCR) are a research instrument of the Swiss National Science Foundation. This paper was begun while Imbs was visiting the Hong Kong Institute for Monetary Research, whose hospitality is gratefully acknowledged. We thank audiences at the Paris School of Economics, the European Central Bank, the Stockholm School of Economics, Ente Einaudi, the NBER 2008 Summer Institute, UCLA, UC Berkeley, UC Santa Cruz, the San Francisco Fed, HEC Paris, Insead, Sciences-Po and Cambridge, as well as George Alessandria, Nicolas Coeurdacier, Jonathan Eaton and Alessandro Rebucci for fruitful exchanges. The views expressed herein are those of the authors and should not be attributed to the IMF, its Executive Board, or its management.

# Contents

	Page
I Introduction . . . . .	<a href="#">4</a>
II Trade Elasticities: Practice ahead of Theory . . . . .	<a href="#">7</a>
A Practice . . . . .	<a href="#">8</a>
B Theory . . . . .	<a href="#">10</a>
1 No Firm Entry . . . . .	<a href="#">12</a>
2 Allowing for Firm Entry . . . . .	<a href="#">14</a>
3 The Price Elasticity of Imports . . . . .	<a href="#">16</a>
III Identification . . . . .	<a href="#">17</a>
A Microeconomic Estimates . . . . .	<a href="#">17</a>
B Homogeneous Estimates . . . . .	<a href="#">20</a>
C Data . . . . .	<a href="#">21</a>
IV Results and Relevance . . . . .	<a href="#">22</a>
A Microeconomic Results . . . . .	<a href="#">23</a>
B Macroeconomic Results . . . . .	<a href="#">25</a>
C Relevance . . . . .	<a href="#">28</a>
D Stability . . . . .	<a href="#">30</a>
V Conclusion . . . . .	<a href="#">32</a>
References . . . . .	<a href="#">33</a>
A Aggregation and Homogeneity . . . . .	<a href="#">36</a>
B Heterogeneous Supply . . . . .	<a href="#">37</a>
C The Price Elasticity of Imports . . . . .	<a href="#">40</a>
D Derivation of the Supply Curve . . . . .	<a href="#">41</a>
E Estimated Variances . . . . .	<a href="#">42</a>
F A Two-Sector Version of BKK <sup>2</sup> . . . . .	<a href="#">43</a>
Tables	
1 Estimation with common correlated effects . . . . .	<a href="#">26</a>
2 Variants on the weights . . . . .	<a href="#">31</a>
3 Estimation without common correlated effects . . . . .	<a href="#">31</a>
F.1 Benchmark Parameter Values taken from BKK (1994) . . . . .	<a href="#">45</a>

---

<sup>2</sup>We are grateful to Jean-Olivier Hairault who kindly shared his codes to the one-sector version of the BKK model.

## Figures

1	Microeconomic estimates of the elasticity of substitution . . . . .	<a href="#">23</a>
2	The J-curve in a two-sector BKK model . . . . .	<a href="#">29</a>

## I. INTRODUCTION

The substitutability between domestic and foreign goods is central to most calibrated models in international economics. Depending on the value assigned to the parameter, the predictions of virtually any calibration exercise with an international dimension change quantitatively, sometimes dramatically. The parameter value is key in models that seek to quantify the magnitude of a change in international prices consistent with a rebalancing of external balances as in Obstfeld and Rogoff (2005). It is also a crucial element to the mechanisms that underpin the international sharing of risk in Cole and Obstfeld (1991), and the composition of international portfolio holdings, as in Coeurdacier (2009). The same is true of the importance of exchange rates in the optimal conduct of monetary policy, as in Galí and Monacelli (2005).

This is quite simply one of the most important parameters in international economics. Unsurprisingly, its calibrated value draws from literally decades of empirical work. Unfortunately, little consensus has emerged from the effort, except for two broad conclusions. First, elasticity estimates inferred from aggregate data are barely positive. Second, there are enormous differences between goods. Long time ago, Orcutt (1950) referred to an “elasticity pessimism”, which he related to the low observed volatilities in aggregate quantities, and the high volatility of international relative prices. Here we ask whether the vast heterogeneity in microeconomic elasticity estimates cannot actually be cause for optimism. We show inferring the parameter from aggregate data actually imposes homogeneity. Such homogeneity constraint can create a heterogeneity bias, which we show is the main reason why aggregate elasticity estimates are close to zero in aggregate US data. We propose a correction of the parameter that accounts for heterogeneity at the microeconomic level.

In virtually all macroeconomic models, the elasticity of substitution between domestic and foreign macroeconomic quantities is calibrated on the basis of estimates that arise from aggregate data. But an elasticity obtained from aggregated data implicitly imposes homogeneity across sectors, because the estimation is in fact performed on quantities and price indices that have been aggregated up to the country level. With aggregate data, the estimation can only hope to identify *the* elasticity of substitution between the domestic and foreign bundles of goods produced in different sectors. For example, with conventional Constant Elasticity of Substitution (CES) preferences, the demand  $M$  for the imported bundle of goods is given by

$$M = \left( \frac{P_M}{P} \right)^{1-\sigma} PC,$$

where  $\frac{P_M}{P}$  is the price of imported goods relative to the domestic price index,  $PC$  are total domestic nominal expenditures, and  $C = \left( C_H^{\frac{\sigma-1}{\sigma}} + C_F^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$ .  $C_H$  and  $C_F$  capture demand for domestic and foreign goods, respectively, aggregated over the sectors in each country. Sectoral heterogeneity is aggregated away in the very construction of these quantities. By definition,  $P_M$  aggregates sector specific prices, using the elasticity of substitution between sectors; and  $P$  aggregates domestic and import prices, using the international elasticity  $\sigma$ . Given the very structure of this demand system, it is impossible for  $\sigma$  to have a sector dimension. This is because cross-sector aggregation is performed first, so that the only international substitutability that can prevail is between aggregate bundles of goods. Homogeneity is forced onto the estimation, and could have severe consequences on the end results. In the words of Kreinin (1967): “It is a major shortcoming of most estimates that they are concerned with demand for manufactured products as a whole. Elasticity estimates for individual

commodities are rare. This lack imposes a severe constraint on attempts to quantify the effects of policy measures on the volume of trade and economic welfare. Of necessity most such studies apply a common elasticity figure to all commodity groups.” (p.510). Aggregating the *data* is fundamentally different from aggregating elasticity *estimates*, as we do in this paper.

Identification issues notwithstanding, the aggregate elasticity of substitution  $\sigma$  can be inferred from (one minus) the observed response of imported aggregates to appropriately identified shifts in price indices. But the inference is *only* valid under the restriction that the sector-specific elasticities of substitution,  $\sigma_k$ , are equal across sectors. It is undisputable that the price elasticity of aggregate imports is close to zero in the data; but jumping to the conclusion that  $\sigma$  is, as well, is only warranted under the restriction that  $\sigma_k = \sigma$  for all sectors  $k$ .<sup>3</sup> Given the overwhelming evidence that estimates of  $\sigma_k$  vary by sector, we argue this is an implausible assumption.

With heterogeneous elasticities of substitution across sectors and CES preferences, we show the aggregate elasticity is given by an appropriately weighted average of sector-specific elasticities  $\sigma_k$ , say  $\sum_k \omega_k \sigma_k$ , with weights  $\omega_k$  given by theory. The weighted average should be understood as the representative agent’s elasticity of substitution, computed in a way that accounts for the sector-level heterogeneity in  $\sigma_k$ . Whenever the estimates of  $\sigma$  based on aggregate data are significantly different from  $\sum_k \omega_k \sigma_k$ , the homogeneity constraint implicit in aggregate data effectively creates a heterogeneity bias. Then, the only option available to pin down an aggregate substitutability that accounts for heterogeneity is to estimate  $\sigma_k$  from disaggregated data, and infer the weights used in aggregation from theory. Using aggregate data to estimate the price elasticity of aggregate imports, as is customary in macroeconomics, will only identify an aggregate substitutability that imposes homogeneity.

Now it is in fact possible the estimates of  $\sigma$  obtained from aggregated data happen to pinpoint precisely the adequately weighted average of  $\sigma_k$ . Then aggregate data imply estimates of  $\sigma$  that happen to account for sectoral heterogeneity. Sector-level elasticities are still constrained to homogeneity, but at a level that happens to be the one implied by the theory with heterogeneity, i.e.,  $\sigma = \sum_k \omega_k \sigma_k$ . In this case, the price elasticity of aggregate imports can still be used to infer an aggregate substitutability that accounts for heterogeneity. The question is therefore empirical: Do aggregate data imply an estimate of  $\sigma$  that happens to concord with the value of  $\sum_k \omega_k \sigma_k$  implied by sectoral data? In this paper, we show the answer for US data is negative.

We develop a model of CES preferences telling us how to properly aggregate microeconomic elasticities. Households have CES preferences over a continuum of goods, or sectors. Consumption in each sector is in turn a CES aggregate of different varieties, produced in different countries. Crucially, the elasticity of substitution between varieties is allowed to be sector-specific. We construct a measure of aggregate substitutability consistent with a representative agent choosing between country-level aggregates of domestic and foreign quantities. The model accommodates the well documented fact that substitutability is different across sectors. We show there is potentially a discrepancy between conventional macroeconomic estimates of the elasticity of substitution, imposing equal elasticities across goods, and aggregate estimates allowing for heterogeneity.

---

<sup>3</sup>We show this formally for CES preferences in Appendix A

When firm supply decisions are modeled explicitly, the observed response of quantities to prices is an imperfect reflection of substitutability, since it conflates both demand and supply phenomena. Traded quantities respond to altered price conditions in a way that does not only reflect consumers' preferences, but also entry decisions on the part of supplying firms, an argument recently developed in Chaney (2008). Our theory allows for this possibility. In the CES setup, we show the discrepancy between constrained and unconstrained elasticities persists whether firm entry decisions are permitted, or not. In fact, with CES preferences, we are able to spell out the conditions under which firm dynamics act to reinforce a heterogeneity bias in conventional macroeconomic estimates of the elasticity of substitution.

Our second goal is to estimate disaggregated elasticities. To do so, we borrow from a methodology introduced by Feenstra (1994a) and recently implemented by Broda and Weinstein (2006). In the context of our model, the approach can be used to identify the parameter of interest. Demand at a sectoral level is given by a CES aggregator of domestic and foreign goods varieties. Under an Armington assumption, the substitutability of domestic and foreign varieties is the same as the substitutability between two foreign varieties. Domestic and foreign varieties do differ in terms of transport costs and preferences, but not substitutability. As a consequence, we are able to identify sectoral elasticities of substitution using the observed cross-country variation in the trade flows towards a given importer. In comparison with conventional approaches, this provides estimates that are structural, and do not fall victim to the endogeneity concerns that plague any regression of (relative) quantities on (relative) prices.

The Armington assumption makes it possible to identify the substitutability between domestic and foreign goods using *imported* prices and quantities only. This contrasts with much of the existing empirical work, where it is the response of quantities to changes in the price of foreign goods relative to *domestic* varieties that is studied. It is important that we validate this identifying assumption. We do so by comparing our results to those obtained in past studies where prices are measured relative to domestic goods. At the sector level, our median estimate across manufacturing is 5.1. This is virtually identical to the estimates obtained, at the same level of aggregation and for similar activities, by Houthakker and Magee (1969) or Kreinin (1967). This similarity in estimates across decades suggests there is nothing special to our data, or our approach, relative to existing empirical work. At sector level, we seem to identify the object of interest to many international economists.

Armed with an econometric methodology to estimate  $\sigma_k$  and a theory to aggregate them, we conduct two experiments. First, we constrain the estimates to equality across sectors. We expect the results to be in line with the macroeconomic literature that has used aggregate data to infer the aggregate elasticity of substitution. This puts our theory to the test of its ability to reproduce conventional macroeconomic estimates under conventional macroeconomic assumptions. Second, we relax the homogeneity constraint, and obtain estimates of  $\sum_k \omega_k \sigma_k$ . If a significant difference exists with constrained estimates, a heterogeneity bias prevails in the data. Then, macroeconomic data cannot be used to obtain a substitutability measure that reflects the sectoral heterogeneity in  $\sigma_k$ .

In macroeconomic applications, calibration exercises typically favor values of the parameter that are inferred from aggregate data, and chosen as “plausible mid-points” to the wide range of estimates the literature has uncovered. For instance, Obstfeld and Rogoff (2005) use a value of 2; Backus, Kehoe and Kydland (1994) use 1.5. The elasticity is (crucially) unitary in Cole and Obstfeld (1991), 0.9 in Heathcote and Perri (2008), 1.5 in Chari, Kehoe and McGrattan (2002), and set between 0.6 and 2 in Coeurdacier, Kollmann and Martin (2007). When all elasticities are forced to be equal across



sectors, our approach generates aggregate estimates between 2.5 and 3 for the US. The range is within the ballpark of the calibrated values used in the macroeconomic literature, and within the range of estimates obtained from aggregate data surveyed in Goldstein and Kahn (1985). With heterogeneity however, aggregate elasticity estimates more than double, with values in the US up to 6 or 7. Such significant difference demonstrates a heterogeneity bias is at play in aggregate US data. It also tells us an aggregate elasticity of substitution that accounts for sectoral heterogeneity is closer to 7 than to 2.

The comparison between constrained and unconstrained estimates focuses squarely on the importance of heterogeneity. We implement the same estimator, on the same data set, with the sole difference that we impose or not a homogeneity constraint. Potential alternative explanations for a discrepancy between the elasticities inferred from microeconomic or macroeconomic data are effectively held constant. When we perform our estimation on an aggregated version of our data, we find estimates for the elasticity of substitution near one. This is close to the results obtained with a homogeneity constraint, and close to conventional macroeconomic estimates. But the comparison is less valuable, since aggregating the data obscures the *ceteris paribus* nature of our main experiment. For instance, aggregation substantially reduces the dimensionality of the data relevant to identification.

Does heterogeneity matter economically? We discuss some illustrations in areas as diverse as the resolution of global imbalances, the dynamics of the trade balance, international risk sharing, portfolio choice, and optimal monetary policy. In short, the parameter is central to most of international macroeconomics. Whether it is in fact 2 or 7 does make a quantitative difference, and calibrated equilibrium responses often change sizably. In some instances, the correction is also relevant qualitatively: for instance, the justification for a home bias in portfolio investment, or what macroeconomic policies are Pareto-superior can all change profoundly with an elasticity substantially above one.

The impact of heterogeneity is significant and robust. Differences between homogeneous and heterogeneous elasticity estimates are statistically significant at conventional confidence levels, and they prevail for a variety of alternative measures or econometric procedures. Our conclusions withstand controls for common components in prices and quantities, and a battery of alternative data sources used to aggregate up microeconomics estimates. Point estimates are minimally affected, and heterogeneity always retains an economically and statistically significant impact.

In what follows, we first present the model used to guide the aggregation of industry specific elasticities. Section 3 discusses the identification of sector specific parameters, their aggregation and the data involved. We also describe our adaptation of Feenstra's approach. Section 4 reports our results, and document their relevance in recent standard models of international trade in goods and assets. Section 5 concludes.

## II. TRADE ELASTICITIES: PRACTICE AHEAD OF THEORY

We open with a summary of the empirical literature concerned with estimating trade elasticities to infer substitutability. We discuss common practice, and how end estimates tend to be heterogeneous in disaggregated data. We then lay out the theory we use to map out conventional elasticity estimates with the utility parameter of interest in macroeconomic calibration exercises. The model tells us how

to compute aggregate elasticities in ways that may or may not allow for heterogeneity at the good's or sector's level. It is constructed to also accommodate firm entry decisions.

### A. Practice

Estimating import price elasticities is an old business in economics. Historically, estimates arising from sectoral or aggregate data have virtually always been the object of separate studies. The dichotomy is natural, since microeconomic trade elasticities typically consider the effect of a shift in sectoral price, holding constant the price of close substitutes, whereas macroeconomic elasticities consider shocks to all relative prices. In what follows, we are careful to consider the response of aggregate quantities to the same, macroeconomic, shock, allowing or not for heterogeneity in the response of quantities to a given shock in prices. Whether we impose homogeneity or not, we do so on the same data, characterized by an identical variance-covariance matrix, and on the basis of the same estimator. The comparison brings the focus squarely on the importance of heterogeneity in elasticity estimates.

Goldstein and Kahn (1985) survey cross-country evidence about trade elasticities, estimated on the basis of aggregated data. Their meta-analysis summarizes the results implied by aggregate data for a variety of countries, time periods and econometric techniques. They report estimates for import price elasticities rarely above 2 in absolute value. Such small estimates have largely remained unchanged since, even though the econometrics involved have become considerably more sophisticated. Thus, Marquez (1990) implements a frequency domain estimator, Gagnon (2003) instruments import prices using the real exchange rate, and Hooper, Johnson and Marquez (1998) use co-integration techniques. In most instances, estimates of the import price elasticity arising from aggregated data are found to be weakly negative, not always significant, and rarely larger than 2 in absolute value. The conventional inference in international macroeconomics is to deduce that the elasticity of substitution between aggregate bundles of domestic and foreign goods is close to zero as well. For instance, Backus, Kehoe and Kydland (1994) note that “there is some uncertainty about what value of [the elasticity of substitution] is indicated by the data. The most reliable studies seem to indicate that for the United States the elasticity is between 1 and 2” (p.91).

Studies that provide disaggregated elasticity estimates have been on the rise over the past decades, because of improved data availability. Early contributions include Kreinin (1967) on three sectors for ten industrial countries, Huddle (1970) on Brazil, or Khan (1975) on Venezuela. Import elasticities at the sector level display considerable heterogeneity, ranging from  $-1.1$  to  $-5.0$  for instance in Kreinin (1967). Heterogeneity is largely confirmed in more recent contributions that merge disaggregated information on trade restrictions, traded quantities, and prices. For instance, Blonigen and Wilson (1999) document substitution elasticities between 0 and 3.52 in 146 US sectors. Microeconomic studies also open the door to instrumentation strategies where changes in international relative prices can be ascribed to events, such as trade liberalizations, whose magnitude and timing are arguably exogenous to each market's circumstances. Thus, using NAFTA and detailed information on good specific tariff changes, Head and Ries (2001) find disaggregated substitution elasticities between 8 and 12, while Romalis (2007) documents estimates between 4 and 13. In the words of Anderson and Van Wincoop (2004), “overall the literature leads us to conclude that the elasticity is likely to be in the range of 5 to 10” in disaggregated data (p.716).

Even the relatively infrequent papers that do estimate both macro and micro elasticities carefully avoid drawing a link between them. For instance, Houthakker and Magee (1969) dedicate a section to elasticities estimated on aggregate traded quantities and prices. They find import price elasticities close to zero for fifteen developed economies, and reject heterogeneity in estimates across countries. In a separate section, they estimate import elasticities for five US commodity classes broken down by degree of processing. They find substantial heterogeneity, with values ranging from 0 to  $-4.05$ . But they never explicitly or formally relate heterogeneity at the sectoral level and aggregate estimates. In footnote 8, they remark that “the price elasticity for total imports is considerably below a weighted average of the five separate elasticities; this is due to the fact that crude foods and crude materials, which have a low elasticity, are more variable than the other classes and consequently have a relatively greater influence on the relevant sums of squares and cross-products.” The intuition resemble a heterogeneity bias.

We do not propose to dispute the argument that micro-based elasticity estimates can be different objects than the values obtained from macroeconomic data. For instance, Ruhl (2005) argues cross-sectional elasticity estimates are naturally higher, for they embed long run phenomena, such as firm dynamics and the associated adjustments in the quantities produced. Time series data, in contrast, tend to focus on high or medium frequency developments, and may overlook most entry or exit decisions. But in practice, disaggregated datasets tend to be cross-sectional, whereas aggregate ones have information over time. The apparent importance of aggregation may therefore be an artefact of data availability and differences in econometric methodologies. The parameters estimated in micro- and macro-economic studies can in fact be fundamentally different, since in practice they may not capture the extensive margin to the same extent.

In what follows, we are careful to accommodate this possibility. Both our corrected and conventional elasticity estimates arise from the same dataset. Both are affected by a putative extensive margin to exactly the same extent. In fact, Feenstra (1994a) discusses how his empirical approach can accommodate time-varying number of firms in each exporting economy, an argument we clarify when we describe the estimation. In short, our homogeneous and heterogeneous estimations are conceptually similar, and they use the same dimension in the same data. Ruhl’s (2005) argument cannot explain the discrepancies we identify in structural estimates of the elasticity of substitution. In fact, the experiment we propose is *ceteris paribus* in a general sense. We estimate homogeneous and heterogeneous elasticities on the same data set, and within the same methodology. Our econometric approach holds constant all other potential explanations for a difference between macro and micro elasticity estimates. If a discrepancy subsists, it will have to stem from heterogeneity. We zoom in onto the heterogeneity question, and point to its potential importance in macroeconomics.

Macroeconomists have in fact recently recognized such potential importance. When calibrating the elasticity of substitution, an increasing number of international macroeconomists eschew estimates of the parameter that were obtained from aggregate data. They refer instead to microeconomic studies. For instance, Obstfeld and Rogoff (2000) choose a value of 6, arguing it is consistent with disaggregated estimates. In their theory of the international diffusion of technology shocks, Corsetti, Dedola and Deluc (2008) consider a value of 8, once again “based on the estimates in the trade literature” (p.460). Coeurdacier (2009) chooses 5, as “the lower bound of estimates from the trade literature” (p.88), to study the impact of trade costs on aggregate portfolio choice. All these models are macroeconomic in nature, and all build upon the CES preferences that we assume here as well. Such a recent trend in calibration choices draws an implicit link between the heterogeneous disaggregated estimates and their macroeconomic counterpart. In this paper, we formalize the link.

## B. Theory

Consumption in the domestic economy is an aggregate of imperfectly substitutable sectors  $k = 1, \dots, K$ . Utility is given by

$$C = \left[ \sum_{k \in K} (\alpha_k C_k)^{\frac{\gamma-1}{\gamma}} \right]^{\frac{\gamma}{\gamma-1}}$$

where  $\alpha_k$  denotes an exogenous preference parameter and  $\gamma$  measures the substitutability between sectors. Consumption in each sector is derived from a continuum of varieties of good  $k$ , that may be imported or not, as in

$$C_k = \left[ \sum_{i \in I, i \neq d} (\beta_{ki} C_{ki})^{\frac{\sigma_k-1}{\sigma_k}} + (\beta_{kd} C_{kd})^{\frac{\sigma_k-1}{\sigma_k}} \right]^{\frac{\sigma_k}{\sigma_k-1}}$$

where  $i \in I$  indexes varieties and  $d$  is the domestic variety of good  $k$ . Crucially, the elasticity of substitution  $\sigma_k$  is specific to each industry, and assumed identical across all varieties, imported or not.  $\beta_{ki}$  lets preferences vary exogenously across varieties, reflecting for instance differences in quality or home bias in consumption. The sectors that verify  $\beta_{ki} = 0$  for all  $i \neq d$  are effectively non-traded.

This structure of demand is classic in international economics. The key assumption for our purposes is equal substitutability between two varieties, no matter their origin. Introducing the assumption is largely what opened the door to the New Trade literature, pioneered by Krugman (1980), and laid the foundation for the more recent models of trade with heterogeneous firms, starting with Melitz (2003).

We allow for heterogeneity in the continuum of firms that produce each variety  $i \in I$  in each sector  $k$ . We have:

$$C_{ki} = \left[ \int_f c_{kif}^{\frac{\rho_k-1}{\rho_k}} df \right]^{\frac{\rho_k}{\rho_k-1}}$$

where  $f$  indexes the range of firms that are active in country  $i$  and sector  $k$ . The range of active firms may differ endogenously across countries and sectors.  $\rho_k$  denotes the elasticity of substitution between the varieties produced by firms in sector  $k$ , which we assume identical across all countries  $i$ . We further assume  $\rho_k > \sigma_k > \gamma > 1$  for all  $k$ . The varieties produced by two firms from the same country are more substitutable than those produced by firms located in two different countries. These, in turn, are more substitutable than goods from different sectors. Finally, we assume complementarities away.

The representative maximizing agent chooses her consumption allocation on the basis of Cost, Insurance, Freight prices, labeled in local currencies.<sup>4</sup> Utility maximization implies that demand for variety  $i$  in each sector  $k$  is given by

$$C_{ki} = \beta_{ki}^{\sigma_k-1} \left( \frac{P_{ki}}{P_k} \right)^{-\sigma_k} \alpha_k^{\gamma-1} \left( \frac{P}{P_k} \right)^{-\gamma} C \quad (1)$$

---

<sup>4</sup>Without loss of generality, we could introduce an additional price wedge, reflecting distribution costs that presumably affect both domestic and foreign varieties. This would merely add some notation, but no further insight. In the empirics, the price of each variety is measured Free on Board, i.e. net of both retail and transportation costs.

with

$$\begin{aligned}
P_{ki} &= \left[ \int_f p_{kif}^{1-\rho_k} df \right]^{\frac{1}{1-\rho_k}} \\
P_k &= \left[ \sum_{i \in I, i \neq d} \left( \frac{P_{ki}}{\beta_{ki}} \right)^{1-\sigma_k} + \left( \frac{P_{kd}}{\beta_{kd}} \right)^{1-\sigma_k} \right]^{\frac{1}{1-\sigma_k}} \\
P &= \left[ \sum_{k \in K} \left( \frac{P_k}{\alpha_k} \right)^{1-\gamma} \right]^{\frac{1}{1-\gamma}}
\end{aligned}$$

The demand addressed to each firm is given by:

$$c_{kif} = \left( \frac{p_{kif}}{P_{ki}} \right)^{-\rho_k} C_{ki} \quad (2)$$

where  $p_{kif}$  is the (cost, insurance, freight) local currency price charged by firm  $f$  producing good  $k$  in country  $i$ . The paper's empirics are performed at the level of good  $k$ , allowing for different varieties across exporters  $i$ . The firm-level dimension  $f$  is introduced so that the theory-implied estimates we develop can accommodate firm entry dynamics.

We now ask our model how the estimated response of *aggregate* quantities to changes in *aggregate* international relative prices is affected depending on whether heterogeneity in  $\sigma_k$  is permitted. For this to be a meaningful experiment in a model with multilateral trade at the industry level, we consider disturbances to international relative prices of a specific kind. First, we focus on changes in all relative prices, across all sectors  $k$ . This means reallocation of demand across industries is solely driven by the heterogeneous response of sectoral quantities to a uniform price shock. It is relative quantities whose responses may be heterogeneous, which in turn may obscure aggregate estimates.

Second, we focus on uniform shocks to the international price of domestic goods, across all exporters  $i$  in  $I$ . This assumes away reallocation of demand across source exporting economies, with relative prices changing identically in all markets. We do this for practical reasons, so that the multilateral dimension of the model collapses into a two-country version, and we can interpret our estimate as capturing the substitutability between composite goods in the domestic economy and in the rest of the world.<sup>5</sup> A natural candidate is a domestic shock to relative production costs, driven for instance by a productivity disturbance. It will change the international price of domestic goods by an identical amount across all sectors  $k$  and exporters  $i$ . As a short-hand, we label the shock a “domestic wage” disturbance  $w_d$ . An increase in  $w_d$  represents a real appreciation driven by a positive shift in relative domestic costs and ultimately prices across all sectors.

Consider the definition of an aggregate elasticity of substitution  $\sigma$  between bundles of domestic and

---

<sup>5</sup>The second assumption is made for convenience. The intuition remains the same if we focus on a change in (all) relative prices between the domestic economy and a specific exporter  $i$ . The data needed to perform aggregation are just slightly different, and identification becomes more complicated since it relies on the cross-section of all exporters to a given destination.

foreign goods. By definition,

$$\sigma = 1 + \frac{\partial \ln \sum_k \sum_{i \neq d} P_{ki} C_{ki} - \partial \ln \sum_k P_{kd} C_{kd}}{\partial \ln w_d}$$

The elasticity of substitution captures the relative response of demand for domestic or foreign bundles of goods. Demand is expressed in nominal terms because virtually all trade data are expressed in value, especially at a disaggregated level. Since the driving force to the shift in relative prices is aggregate, the difference between the elasticity of substitution arising from volume or value data is simply 1.

Using equation (1) and its counterpart for the domestic variety, simple algebra implies

$$\begin{aligned} \sigma - 1 &= \sum_k \sum_{i \neq d} n_{ki} (1 - \sigma_k) \frac{\partial \ln P_{ki}}{\partial \ln w_d} - \sum_k n_{kd} (1 - \sigma_k) \frac{\partial \ln P_{kd}}{\partial \ln w_d} \\ &\quad + \sum_k (n_k - n_{kd}) (\sigma_k - \gamma) \frac{\partial \ln P_k}{\partial \ln w_d} \end{aligned} \quad (3)$$

with

$$\begin{aligned} n_{ki} &\equiv \frac{P_{ki} C_{ki}}{\sum_{k \in K} \sum_{i \neq d} P_{ki} C_{ki}} \\ n_{kd} &\equiv \frac{P_{kd} C_{kd}}{\sum_{k \in K} P_{kd} C_{kd}} \\ n_k &= \sum_{i \neq d} n_{ki} \end{aligned}$$

The aggregate elasticity of substitution has two components. The first two terms correspond to appropriately weighted averages of the responses of individual price indices to the cost shock, at the variety level  $i$  or  $d$ . The third term captures the responses of the sector price index to the cost shock.

## 1. No Firm Entry

We clarify the separate influences of the issues raised by a homogeneity constraint, and supply response within each sector  $k$ . In particular, we first focus on the former and shut down firm entry. Since we are interested in long run estimates, in each sector  $k$  the representative domestic producer modifies her price according to the change in costs,  $\frac{\partial \ln P_{kd}}{\partial \ln w_d} = 1$ . As we assume away any changes in market structure, foreign producers do not respond to changes in domestic costs, and  $\frac{\partial \ln P_{ki}}{\partial \ln w_d} = 0$ .<sup>6</sup> As a result, the response of the sector price index is given by

$$\frac{\partial \ln P_k}{\partial \ln w_d} = \frac{\partial \ln P_k}{\partial \ln P_{kd}} \frac{\partial \ln P_{kd}}{\partial \ln w_d} = \frac{P_{kd} C_{kd}}{P_k C_k} = 1 - w_k^M$$

---

<sup>6</sup>The assumption of CES preferences is crucial here. With linear quadratic preferences, for instance, or more sophisticated market structure, price responses embed both demand elasticities and changes in markups.

where  $w_k^M = \frac{\sum_{i \neq d} P_{ki} C_{ki}}{P_k C_k}$  is the share of imports in total expenditures on good  $k$ . Equation (3) simplifies into

$$\sigma_{NoFirm} = \sum_k n_{kd} \sigma_k + \sum_k (n_k - n_{kd}) (1 - w_k^M) (\sigma_k - \gamma) \quad (4)$$

The aggregate elasticity  $\sigma_{NoFirm}$  contains two terms. First, a weighted average of industry-specific elasticities, with weights corresponding to the importance of sector  $k$  in overall domestic expenditures. The second term reflects the response of the industry specific price index  $P_k$ . Since by assumption the relative price of good  $k$  changes identically across all source economies  $i \neq d$ , the composition of the ideal price index in sector  $k$  changes significantly in response to the shock considered. We label  $\sigma_{NoFirm}$  a total elasticity, i.e. one that takes the response of price indices into account. In contrast, a partial elasticity assumes aggregate price responses away.<sup>7</sup> We note the second term in equation (4) is likely small. By definition of  $n_{kd}$ ,  $\sum_k (n_k - n_{kd}) (1 - w_k^M) (\sigma_k - \gamma) < \sum_k n_k \sigma_k (1 - w_k^M)$ . The difference between partial and total elasticities is bounded above. Relative to partial elasticity, the upper bound contains an extra multiplicative term smaller than 1, since  $w_k^M < 1$  for all  $k$ . Partial and total elasticities are therefore likely to differ by small amounts. In fact, if sector allocations of expenditures are similar for domestic and foreign goods,  $n_k \simeq n_{kd}$ , and partial and total elasticities are virtually identical.

Focus for now the partial elasticity and suppose  $\frac{\partial \ln P_k}{\partial \ln w_d} = 0$ . The aggregate substitutability between foreign and domestic varieties is given by a weighted average of each industry's corresponding preference parameter. Strictly speaking, this weighted average is the direct equivalent of the aggregate elasticity of substitution obtained from an aggregation of *estimates*, rather than an aggregation of the *data*. It captures the direct effect of a shock to  $w_d$  on  $P_{kd}$ , and the resulting immediate change in  $C_{kd}$  visible from equation (1). In macroeconomic data, traded quantities are summed up across sectors and exporting countries. The econometrician asks how these synthetic aggregates respond to changes in international prices. In other words, she estimates one single value for  $\sigma$ , which may well differ from  $\sum_k n_{kd} \sigma_k$ . If it does, a heterogeneity bias prevails in the data, and disaggregated data are necessary to calculate an aggregate elasticity accounting for heterogeneity.<sup>8</sup> In fact, equation (4) tells us the direction and magnitude of a heterogeneity bias depend on the cross-sector correlation between expenditures shares  $n_{kd}$  and goods' substitutability  $\sigma_k$ . If substitutable varieties tend to form a large share of domestic expenditures, the unconstrained elasticity  $\sigma_{NoFirm}$  will take larger values than its constrained counterpart,  $\bar{\sigma}_{NoFirm}$ .

Sectoral price indices also respond to macroeconomic shocks. Consider a positive shock to domestic costs, driving up domestic prices and therefore  $P_k$ , to an extent that increases with the share of domestic varieties,  $1 - w_k^M$ . The response of the price index affects the relative prices of both domestic and imported varieties. It acts to dilute the initial upwards change in  $\frac{P_{kd}}{P_k}$ , which dampens measured total elasticity. But it also drives a negative response in  $\frac{P_{ki}}{P_k}$ , with opposite end effects on measured total elasticity.

<sup>7</sup>The second term in equation (4) exists because of our focus on an aggregate shock in relative prices. If instead the shock considered were microeconomic in nature and focused on a specific exporter  $i$  - a change in tariff - then under standard atomistic assumptions and large  $I$ , the second term in equation (4) would disappear, and we would be left with a partial elasticity.

<sup>8</sup>In Appendix A, we show that in a CES world with heterogeneous  $\sigma_k$ , estimating the conventional specification on aggregated data effectively imposes an additional constraint, namely  $\sigma_k = \sigma = \gamma$ .

Estimates of the elasticity of substitution obtained from aggregate data provide potentially biased values of a preference parameter that accounts for heterogeneity in  $\sigma_k$ . The bias arises because differences in sector-specific substitutability are assumed away. This is true of partial and total elasticity, and we later report results for both concepts. In what follows, we obtain structural estimates of  $\sigma_k$  (and its counterpart constrained to homogeneity), and infer the corresponding, theory-implied (partial or total) values for  $\sigma_{NoFirm}$  and  $\bar{\sigma}_{NoFirm}$ . We will show that the implied values for  $\bar{\sigma}_{NoFirm}$  are in line with classic results from the macroeconomic empirical literature, whereas the implied values for  $\sigma_{NoFirm}$  are not.

## 2. Allowing for Firm Entry

We now reintroduce firm heterogeneity in the definition of the aggregate elasticity of substitution in equation (3). Exporter's prices  $P_{ki}$  and  $P_{kd}$  are now indices, reflective of the continuum of firms active in sector  $k$  across countries. This range will vary endogenously in response to changes in international relative prices, and this response itself may differ across sectors and exporters because of differences in transport costs, assumptions on firms distribution, or heterogeneity in substitutability parameters. The argument is similar to Chaney (2008). In Appendix B, we go through what is now the conventional model of supply decisions made by heterogeneous firms in each sector and each country. As we describe in the Appendix, the price charged in country  $i$  and sector  $k$  is given by

$$P_{ki} = \left[ \frac{M_{ki}}{1 - G_{ki}(\bar{\varphi}_{ki})} \int_{\bar{\varphi}_{ki}}^{\infty} p_{kif}(\varphi)^{1-\rho_k} dG_{ki}(\varphi) \right]^{\frac{1}{1-\rho_k}} \quad (5)$$

where  $M_{ki}$  is the mass of exporting firms active in sector  $k$  and country  $i$ .  $G_{ki}(\varphi)$  denotes the cumulative distribution function for firms' productivity in sector  $k$  and country  $i$ , and  $\bar{\varphi}_{ki}$  is an (endogenous) productivity cut-off level, above which firms are active on export markets. Equation (5) clearly implies that  $\frac{\partial \ln P_{ki}}{\partial \ln w_d}$ , and by analogy  $\frac{\partial \ln P_{kd}}{\partial \ln w_d}$ , can vary across sectors or countries, with end effects on the definition of the aggregate elasticity given in equation (3). Price indices can now react to domestic cost shocks because market structure changes endogenously. In particular, the response of prices now has three components. First the response of individual firm prices to the domestic shock,  $\frac{\partial \ln P_{kif}}{\partial \ln w_d}$ . This is zero for  $i$  in  $I$ ,  $i \neq d$  and 1 for  $i = d$ , since it reflects the pure intensive margin of adjustment. Second, the response of the productivity cut-off,  $\bar{\varphi}_{ki}$ , as firms enter markets in response to shifting relative prices. And third, the adjustment in the potential number of firms serving the market. These last two terms capture the extensive margin of price adjustments, set to zero in the previous section.

In the short run, the mass of firms in country  $i$  and sector  $k$  is usually taken as given, and the third effect is assumed equal to zero. In this case, the extensive margin works through the response of the productivity cut-off for a constant mass of potential exporters. This is the avenue followed in Chaney (2008), who shows the response of traded quantities to changes in costs is magnified by the extensive margin. Given our concern with long run elasticities, we depart from this approach and focus instead on the long run impact of changes in relative prices, when the potential number of exporters is endogenous.

When firms productivity follow a Pareto distribution, we show in the Appendix that the productivity cut-off  $\bar{\varphi}_{ki}$  only depends on exogenous parameters in a long run equilibrium where  $M_{ki}$  adjusts. The



extensive margin now works through the number of potential entrants. In the Appendix, we derive an expression for  $M_{ki}$ , which can be used to compute  $\frac{\partial \ln M_{ki}}{\partial \ln w_d}$  and  $\frac{\partial \ln M_{kd}}{\partial \ln w_d}$ . We then solve for the endogenous responses of foreign and domestic price indices,  $\frac{\partial \ln P_{ki}}{\partial \ln w_d}$  and  $\frac{\partial \ln P_{kd}}{\partial \ln w_d}$ . Using the definition of  $P_k$ , we solve equation (3) for the aggregate elasticity of substitution allowing for firm entry decisions in all markets:

$$\begin{aligned} \sigma_{Firm} - 1 &= \sum_k n_{kd}(\sigma_k - 1) \frac{\rho_k}{\rho_k - \sigma_k} + \sum_k (n_k - n_{kd})(1 - w_k^M)(\sigma_k - \gamma) \frac{\rho_k}{\rho_k - \sigma_k} \frac{\rho_k - 1}{\rho_k - \gamma} \\ &\quad + (\gamma - 1) \Delta \sum_k (n_k - n_{kd}) \frac{\gamma}{\rho_k - \gamma}, \end{aligned} \quad (6)$$

where  $\Delta = \frac{\sum_k w_k \frac{\rho_k}{\rho_k - \gamma} (1 - w_k^M)}{\sum_k w_k \frac{\rho_k}{\rho_k - \gamma}} = \frac{\partial \ln P}{\partial \ln w_d}$ , and  $w_k = P_k C_k / PC$  is the share of sector  $k$  in aggregate consumption.

The first term in the expression continues to contain an adequately weighted average of disaggregated elasticities  $\sigma_k$ . The magnitude of the potential heterogeneity bias continues therefore to prevail if  $n_{kd}$  and  $\sigma_k$  correlate positively across sectors, exactly as it did when we ignored any extensive margin. However, with an extensive margin, whenever  $\rho_k > \sigma_k$ , the summation term is larger than in equation (4). This happens because prices indices now react to the cost shock through both intensive and extensive margins. The implied estimate is higher because it now encapsulates the relative responses of quantities produced, an argument reminiscent of Ruhl (2005) or Chaney (2008).

Price adjustments at both margins will also affect the response of sectoral and aggregate price indices, as reflected in the last two terms of equation (6). Both terms can take either sign, depending on the relative magnitude of  $n_k$  and  $n_{kd}$ . They correspond to the response in price indices, and reflect the difference between a partial and a total elasticity. For reasons analogous to previously, they are small in magnitude, especially in a symmetric setup where  $n_k$  is close to  $n_{kd}$ .

Our main purpose in this section is to ascertain that supply responses across all markets do not alter the existence of an heterogeneity bias in estimates of the aggregate elasticity of substitution. In other words, the most important question here is whether the discrepancy between the estimate constrained to homogeneity,  $\bar{\sigma}_{Firm}$ , and its unconstrained counterpart,  $\sigma_{Firm}$ , differs sizably from that between  $\bar{\sigma}_{NoFirm}$  and  $\sigma_{NoFirm}$ . We noted in the previous section that the difference between  $\bar{\sigma}_{NoFirm}$  and  $\sigma_{NoFirm}$  increases in the cross-sector correlation between  $n_{kd}$  and  $\sigma_k$ . For positive correlations, a model-implied weighted average of  $\sigma_k$  will take larger values than its counterpart imposing homogeneity  $\sigma_k = \sigma$  for all  $k$ .

Equation (6) suggests the discrepancy will be even larger with firm entry. Under plausible conditions, a given positive value for the correlation between  $n_{kd}$  and  $\sigma_k$  implies larger values for  $\sigma_{Firm} - \bar{\sigma}_{Firm}$  than for  $\sigma_{NoFirm} - \bar{\sigma}_{NoFirm}$ . In other words, the theory implied bias with firm entry is larger than its counterpart assuming supply responses away. To see this, manipulate equation (6) to obtain

$$\begin{aligned} \sigma_{Firm} - \bar{\sigma}_{Firm} &= \sigma_{NoFirm} - \bar{\sigma}_{NoFirm} + \sum_k n_{kd} \left[ (\sigma_k - 1) \frac{\sigma_k}{\rho_k - \sigma_k} - (\bar{\sigma} - 1) \frac{\bar{\sigma}}{\rho_k - \bar{\sigma}} \right] \\ &\quad + \sum_k (n_k - n_{kd})(1 - w_k^M) \left[ (\sigma_k - 1) \frac{\sigma_k}{\rho_k - \sigma_k} - (\bar{\sigma} - 1) \frac{\bar{\sigma}}{\rho_k - \bar{\sigma}} \right], \end{aligned} \quad (7)$$

where  $\bar{\sigma}$  is a homogeneous elasticity of substitution, constrained to be identical across all sectors. Notice the last term in equation (6) drops because it is invariant to  $\sigma_k$ . Under mild conditions, the first summation increases in the correlation between  $n_{kd}$  and  $\sigma_k$ , and acts to augment the magnitude of  $\sigma_{NoFirm} - \bar{\sigma}_{NoFirm}$ .<sup>9</sup> The last summation term can take either sign, depending on values for  $n_k$  and  $n_{kd}$ . But by analogy with the previous section, that last summation term in equation (7) must be substantially smaller than the first, and cannot act to mitigate its impact, even if it takes negative values.

We later introduce an empirical approach to estimate  $\sigma_k$  (and its constrained counterpart  $\bar{\sigma}$ ). For lack of cross-country firm-level data, we cannot obtain any estimates of  $\rho_k$ . We are therefore only able to calculate a theory-implied heterogeneity bias without firm entry,  $\sigma_{NoFirm} - \bar{\sigma}_{NoFirm}$ . In our data,  $n_{kd}$  and  $\sigma_k$  are positively correlated so that the bias will take positive values. Under these conditions, and within the confines of our model of long run supply effects, we just showed that  $\sigma_{Firm} - \bar{\sigma}_{Firm}$  is greater than  $\sigma_{NoFirm} - \bar{\sigma}_{NoFirm}$ . This suggests the heterogeneity bias we document - on the basis of our estimates of  $\sigma_k$  - is a lower bound. It is in fact a conservative under-assessment of the bias that would obtain if an econometrician estimated directly  $\bar{\sigma}_{Firm}$  using aggregate data inclusive of firm entry decisions.

### 3. The Price Elasticity of Imports

In most of the literature, the elasticity of substitution is inferred from the price elasticity of imports, at any level of aggregation. Since we want to validate our assumptions by comparing our results to conventional estimates (at the micro or macroeconomic level), it is important that we verify the bias we discuss continues to prevail in estimates of the price elasticity of imports. Fortunately, the exercise follows arguments that are analogous to what was just discussed. The conventional approach to identifying the price elasticity of imports consists in estimating the response of imported quantities to changes in the relative price of imports (a negative number).<sup>10</sup> In the model, this is equivalent to

$$\eta = \frac{\partial \ln [\sum_k \sum_i P_{ki} C_{ki}]}{-\partial \ln w_d}$$

Demand continues to be given by equation (1). In Appendix C, we derive the following expression for  $\eta_{NoFirm}$ :

$$\eta_{NoFirm} = \gamma - \sigma_{NoFirm} + \sum_k n_{kd}(\sigma_k - \gamma)w_k^M - \gamma \sum_k w_k(1 - w_k^M) \quad (8)$$

The elasticity contains three terms: a direct linear function of the elasticity of substitution  $\sigma_{NoFirm}$ , and two summation terms, likely to be smaller in magnitude than  $\sigma_{NoFirm}$ , since  $n_{kd}, w_k^M < 1$ .

<sup>9</sup>This will be true as long as  $\text{corr}(n_{kd}, \rho_k) < \frac{\rho_k}{\sigma_k} \text{corr}(n_{kd}, \sigma_k)$ . The condition is plausible, since the correlation between  $\rho_k$  and the expenditure share of sector  $k$  is likely to be smaller than that between  $\sigma_k$  and the expenditure share. After all,  $\rho_k$  denotes a substitutability between firms within a sector, whereas  $\sigma_k$  captures a substitutability between countries.

<sup>10</sup>We focus on estimates of the price elasticity of imports abstracting from firm dynamics, since we just established they do not affect our main results.

Abstracting from both summations terms, the elasticity of imports is given by  $\gamma - \sigma_{NoFirm}$ , or by  $\gamma - \bar{\sigma}_{NoFirm}$  when homogeneity is imposed on  $\sigma_k$ . Constraining all elasticities to be the same, the partial price elasticity of imports and the elasticity of substitution are linearly related, a conventional result. With positive cross-sector correlation between  $n_k$  and  $\sigma_k$ , estimates of the (partial) price elasticity of imports that introduce heterogeneity will be larger in absolute value. In the model,  $\sigma_{NoFirm} > \bar{\sigma}_{NoFirm}$  implies that  $\eta_{NoFirm} < \bar{\eta}_{NoFirm}$ . This can explain why macroeconomic estimates of  $\eta$  are smaller in absolute value than those arising from disaggregated data.

For a given  $\gamma$ , equation (8) implies a value for  $\eta_{NoFirm}$  given structural estimates of  $\sigma_k$ , and a value for  $\bar{\eta}_{NoFirm}$  given a structural estimate for  $\bar{\sigma}$ . Whether a heterogeneity bias can account for the discrepancy between estimates of  $\eta$  obtained from micro and macroeconomic data clearly does not depend on a choice for  $\gamma$ . In fact, we later calibrate  $\gamma$  at its lowest value of 1, which corresponds to an upper bound (in absolute value) for the price elasticity of imports. Our hope is the implied value for  $\bar{\eta}_{NoFirm}$  continues to be in line with macroeconomic estimates, i.e. close to zero.

We cannot calculate  $\eta_{Firm}$ , the price elasticity of imports accounting for firm entry, since we do not have estimates of  $\rho_k$ . But by analogy with the previous section, our results suggest that  $\eta_{Firm} - \bar{\eta}_{Firm} < \eta_{NoFirm} - \bar{\eta}_{NoFirm}$ . This means the bias we can measure is an understatement of what would obtain from aggregate data inclusive of firm entry choices. In what follows we therefore limit ourselves to computing  $\eta_{NoFirm}$ , imposing homogeneity or not. We compare the implied values of  $\eta_{NoFirm}$  and  $\bar{\eta}_{NoFirm}$  with conventional estimates of  $\eta$  at both the macro and micro levels obtained from the literature.

### III. IDENTIFICATION

We review how the methodology in Feenstra (1994a) is adapted to our purposes. Identification is structural, but requires a CES demand system, with constant markups. Our results are predicated on these assumptions. We first discuss the econometrics involved in estimating  $\sigma_k$  for all sectors  $k$  in the US economy. We emphasize how we accommodate common effects across all sectors and measurement error. We then turn to the estimation of  $\bar{\sigma}$ , a measure of elasticity constrained to be identical across sectors. We close with a description of our data.

#### A. Microeconomic Estimates

We identify the substitutability between domestic and foreign varieties using the observed cross-section of traded quantities and prices across exporters to one destination. This is afforded by the crucial assumption of an Armington aggregator between varieties of each good, irrespective of their origin. The assumption is what makes it possible to use Feenstra's (1994a) methodology in the present context, even though we do not observe prices or quantities of domestically produced varieties. We now describe our implementation of his methodology, but keep the development concise and focused on the modifications we introduce.

Demand is given in equation (1), which after rearranging writes:

$$C_{kit} = \left( \frac{P_{kit}}{P_{kt}} \right)^{1-\sigma_k} \frac{\beta_{kit}^{\sigma_k-1} P_{kt} C_{kt}}{P_{kit}}$$

where  $t$  is a time index. Feenstra (1994a) or Broda and Weinstein (2006) impose a simple supply structure, with prices fixed in local currency and inclusive of trade costs

$$P_{kit} = \tau_{kit} \exp(v_{kit}) C_{kit}^{\omega_k}$$

where  $v_{kit}$  denotes a technological shock that can take different values across sectors and exporters,  $\tau_{kit}$  is a trading cost and  $\omega_k$  is the inverse of the price elasticity of supply in sector  $k$ .<sup>11</sup> The potential aggregate effects of the nominal exchange rate are soaked up by the shock  $v_{kit}$ . It will therefore be important to implement appropriate econometric tools to control for any potential common effects in our estimated system.<sup>12</sup> In Appendix D we show how this expression for supply can arise from the model with firm entry developed in the previous sections.

In practice, the approach uses expenditure shares to alleviate measurement error in unit values, following Kemp (1962). We define  $s_{kit} = \frac{P_{kit} C_{kit}}{P_{kt} C_{kt}}$  and rewrite demand as

$$s_{kit} = \left( \frac{P_{kit}}{P_{kt}} \right)^{1-\sigma_k} \beta_{kit}^{\sigma_k-1}$$

We do not observe domestically produced consumption. In addition, prices are measured Free on Board. We introduce tilded variables to denote the observed counterparts to theory-implied prices and quantities. We observe  $\tilde{P}_{kit} \equiv P_{kit}/\tau_{kit}$ . The empirical market shares are therefore given by

$$\tilde{s}_{kit} \equiv \frac{\tilde{P}_{kit} C_{kit}}{\sum_{i \neq d} \tilde{P}_{kit} C_{kit}} = \frac{s_{kit}}{\tau_{kit}} \left( 1 + \frac{P_{kdt} C_{kdt}}{\sum_{i \neq d} \tilde{P}_{kit} C_{kit}} \right) \equiv \frac{s_{kit}}{\tau_{kit}} \mu_{kt}$$

Taking logarithms, it is straightforward to rewrite demand as

$$\Delta \ln \tilde{s}_{kit} = (1 - \sigma_k) \Delta \ln \tilde{P}_{kit} + \Phi_{kt} + \varepsilon_{kit} \quad (9)$$

with  $\Phi_{kt} \equiv (\sigma_k - 1) \Delta \ln P_{kt} + \Delta \ln \mu_{kt}$ , a time-varying intercept common across all varieties, and  $\varepsilon_{kit} \equiv (\sigma_k - 1) \Delta \ln \beta_{kit} - \sigma_k \Delta \ln \tau_{kit}$  an error term that captures random trade cost and taste shocks, via changes in  $\tau_{kit}$  and  $\beta_{kit}$ . Feenstra (1994a) shows this implies the demand system is robust to quality changes in variety  $i$  of good  $k$  - or indeed to time-varying number of firms producing good  $k$  in country  $i$ . The estimation is robust to the presence of an extensive margin within exporting economies.

After rearranging, substituting in log-linearized supply yields

$$\Delta \ln \tilde{P}_{kit} = \Psi_{kt} + \frac{\omega_k}{1 + \omega_k \sigma_k} \varepsilon_{kit} + \delta_{kit} \quad (10)$$

with  $\Psi_{kt} \equiv \frac{\omega_k}{1 + \omega_k \sigma_k} \left[ \Phi_{kt} + \Delta \ln \sum_i (\tilde{P}_{kit} C_{kit}) \right]$  a time-varying factor common across varieties, which subsumes sector specific prices and quantities.  $\delta_{kit} \equiv \frac{1}{1 + \omega_k \sigma_k} \Delta v_{kit}$  is an error term encapsulating movements in the exchange rate or aggregate technological developments in country  $i$  and sector  $k$ .

<sup>11</sup>We follow Feenstra (1994a) and assume all exporters have the same supply elasticity. Whether prices are inclusive of transport costs or not is innocuous for the end estimates, as  $\tau_{kit}$  enters the residuals of the estimated equation.

<sup>12</sup> $v_{kit}$  will also absorb any heterogeneity in the extent of the exchange rate pass-through.

Under standard assumptions on taste shocks  $\beta_{kit}$  and technology shocks  $v_{kit}$ , it is possible to identify the system formed by equations (9) and (10). Identification rests on the cross-section of exporters  $i$  to the domestic economy, and is achieved in relative terms with respect to a reference country  $r$ .<sup>13</sup> We follow Feenstra (1994a) and summarize the information contained in the system with the following estimable regression

$$Y_{kit} = \theta_{1k}X_{1kit} + \theta_{2k}X_{2kit} + u_{kit} \quad (11)$$

where  $Y_{kit} = (\Delta \ln \tilde{P}_{kit} - \Delta \ln \tilde{P}_{krt})^2$ ,  $X_{1kit} = (\Delta \ln \tilde{s}_{kit} - \Delta \ln \tilde{s}_{krt})^2$ ,  $X_{2kit} = (\Delta \ln \tilde{s}_{kit} - \Delta \ln \tilde{s}_{krt})(\Delta \ln \tilde{P}_{kit} - \Delta \ln \tilde{P}_{krt})$  and  $u_{kit} = (\varepsilon_{kit} - \varepsilon_{krt})(\delta_{kit} - \delta_{krt}) \frac{(\sigma_k - 1)(1 + \omega_k)}{1 + \omega_k \sigma_k}$ . Estimates of equation (11) map directly with the parameters of interest, since

$$\begin{aligned} \theta_{1k} &= \frac{\omega_k}{(\sigma_k - 1)(1 + \omega_k)} \\ \theta_{2k} &= \frac{\omega_k \sigma_k - 2\omega_k - 1}{(\sigma_k - 1)(1 + \omega_k)} \end{aligned}$$

Equation (11) still suffers from an endogeneity issue. We follow Feenstra (1994a), instrument the regressors with country-sector specific fixed effects, and correct the estimation for heteroskedasticity across exporters  $i$ . As in Feenstra, identification is therefore based on the cross-sectional dimension of equation (11), and is only valid under CES preferences of the type assumed here. We include an intercept to account for the measurement error arising from using unit values to approximate prices. Given the origin of potential measurement error, we let it prevail at the most granular level afforded by our data.<sup>14</sup>

The system summarized by equation (11) can accommodate developments that are specific to each sector  $k$ . But in macroeconomic applications where the universe of economic activities that form Gross Domestic Product is considered, it is important to allow for more general, aggregate influences. Aggregate technology shocks for instance, or movements in the nominal exchange rate, presumably affect prices and quantities jointly in all sectors. If it were a shock in the exporting economy, that would correspond to a common component of  $v_{kit}$  across all  $k$ . We allow for such correlated effects in as general and parsimonious a manner as possible. We implement a correction suggested by Pesaran (2006) to purge all “Common Correlated Effects” (CCE) from sector level data, and estimate

$$Y_{kit} = \theta_0 + \theta_{1k}\hat{X}_{1ki} + \theta_{2k}\hat{X}_{2ki} + \theta_{3k}X_{1it} + \theta_{4k}X_{2it} + u_{kit} \quad (12)$$

where the intercept allows for HS6-specific measurement error, hatted variables are the instrumented versions of  $X_{1kit}$  and  $X_{2kit}$ , and  $X_{1it}$  and  $X_{2it}$  control for the time-varying component of  $Y_{kit}$  that is common across all sectors. In particular, following Pesaran (2006),  $X_{1it}$  and  $X_{2it}$  are the cross-sector arithmetic averages of  $X_{1kit}$  and  $X_{2kit}$ .

<sup>13</sup>In the empirics, we choose a reference country that is present in the US market during the whole observed period.

<sup>14</sup>For the instrumentation to be consistent, there must be some cross-country differences in the relative variance of the demand and supply curves. For an intercept to capture measurement error, its variance must be equal across exporting countries. Of course, it may still be specific to each sector. See Feenstra (1994a, 1994b).

Armed with consistent (and sector-specific) estimates of  $\theta_{1k}$  and  $\theta_{2k}$ , it is straightforward to infer elasticities. In particular, the model implies

$$\begin{aligned}\hat{\sigma}_k &= 1 + \frac{\hat{\theta}_{2k} + \Delta_k}{2\hat{\theta}_{1k}} \text{ if } \hat{\theta}_{1k} > 0 \text{ and } \hat{\theta}_{1k} + \hat{\theta}_{2k} < 1 \\ \hat{\sigma}_k &= 1 + \frac{\hat{\theta}_{2k} - \Delta_k}{2\hat{\theta}_{1k}} \text{ if } \hat{\theta}_{1k} < 0 \text{ and } \hat{\theta}_{1k} + \hat{\theta}_{2k} > 1\end{aligned}$$

with  $\Delta_k = \sqrt{\hat{\theta}_{2k}^2 + 4\hat{\theta}_{1k}}$ . Appendix E details how these are also used to infer standard deviations around these point estimates.

As is apparent, there are combinations of estimates in equation (12) that do not correspond to any theoretically consistent estimates of  $\hat{\sigma}_k$ . This is a problem we encounter in our data, as Broda and Weinstein (2006) did. We follow their approach, and use a search algorithm that minimizes the sum of squared residuals in equation (12) over the intervals of admissible values of the supply and demand elasticities. We use this approach whenever direct estimates of  $\theta_{1k}$  and  $\theta_{2k}$  cannot be used to infer  $\hat{\sigma}_k$ . Whenever CCE are included, we hold constant the estimates of  $\theta_{3k}$  and  $\theta_{4k}$  obtained from the direct instrumental variable regression, and search the combination of values for  $\sigma_k$  and  $\omega_k$  that minimizes the sum of squared residuals in equation (12). The corresponding standard errors are obtained via bootstrapping of the procedure using 1,000 repetitions.

## B. Homogeneous Estimates

We invoke equation (4) to aggregate adequately the estimates of  $\hat{\sigma}_k$  just obtained. Our purpose is to compare these results to what is obtained when sectoral elasticities are constrained to be homogeneous, as they would in conventional regression analysis based on macroeconomic data. To do so, we impose  $\sigma_k = \sigma$  and modify equation (12) into

$$Y_{kit} = \theta_0 + \theta_1 \hat{X}_{1ki} + \theta_2 \hat{X}_{2ki} + \theta_3 X_{1it} + \theta_4 X_{2it} + u_{kit} \quad (13)$$

We maintain the assumption of a HS6-specific intercept, to continue to accommodate the possibility that measurement error varies at the disaggregated level. Not doing so will conflate two potential sources of bias, and the one we are pursuing relates only to the estimates of  $\theta_1$  and  $\theta_2$ . These are now constrained to equality across all sectors  $k$ . Identification continues to rest on the cross-section of exporters  $i$ , but equation (13) is now estimated on the pooled dataset formed by observations on all sectors. It is noteworthy that identification in equations (12) and (13) rests in practice on the *same* dimension of the *same* dataset. It is therefore difficult to ascribe the discrepancy we find to a difference in the extent of an extensive margin. Whether our data (and procedure) capture or not firm dynamics, they do so equally in both estimations.

In the presence of heterogeneity in  $\sigma_k$ , a bias arises because imposing the constraint that  $\theta_{jk} = \theta_j$  for all  $k$  and  $j = 1, 2, 3, 4$  in equation (12) imposes homogeneity in coefficients that theory implies are heterogeneous across sectors. As in Pesaran and Smith (1995), ignored heterogeneity enters the residuals in equation (13), and creates a potential bias in the estimates of  $\theta_j$ ,  $j = 1, 2, 3, 4$ . As a result, empirical estimates of the constrained elasticity of substitution  $\hat{\sigma}$  may well be away from a simple average of its sectoral counterparts.

We continue to allow for the possibility that aggregate shocks in any country  $i$  should affect all sectors simultaneously, and include adequately modified CCE terms. The instrumentation and correction for heteroskedasticity are also modified accordingly. In particular, country-specific effects are used as instruments across the whole panel of sectors. Armed with estimates of  $\theta_1$  and  $\theta_2$ , it is easy to obtain a value for the constrained elasticity of substitution  $\hat{\sigma}$ . Our model then implies conventional macroeconomic estimates of the elasticity of substitution are given by

$$\bar{\sigma}_{NoFirm} = \hat{\sigma} + (\hat{\sigma} - \gamma) \sum_k (n_k - n_{kd}) (1 - w_k^M)$$

with a standard error given by a first-order approximation, as detailed in Appendix E.  $\bar{\sigma}_{NoFirm}$  is the constrained total elasticity, and  $\hat{\sigma}$  is a constrained partial elasticity.

### C. Data

The Armington assumption requires that imports from different countries be imperfectly substitutable varieties. The hypothesis is increasingly palatable as the granularity of the data augments. We choose to use disaggregated, multilateral trade data from the Base Analytique du Commerce International (BACI), released by the Centre d'Etudes Prospectives et d'Informations Internationales (CEPII), and available at the 6-digit level of the harmonized system (HS6). The data cover around 5,000 products over the 1996-2004 period for a large cross-section of countries. The database describes bilateral trade at the sectoral level, building on the United Nations ComTrade database with some added effort put in the harmonization of trade flows on the basis of both import and export declarations. The improvement acts to limit measurement error.

Unlike Feenstra (1994a) or Broda and Weinstein (2006), we do not estimate elasticities at the most disaggregated level in the main body of the text, but rather partition our data into 56 ISIC (Revision 3) industries where we implement our methodology. The constrained estimation given by equation (13) is effectively estimated on a panel of HS6 sectors, and our partition of HS6 sectors into 56 ISIC categories is entirely innocuous here. In the constrained estimation, homogeneity is imposed across all HS6 sectors, and whether or not they belong to the same ISIC sector is irrelevant. The partition becomes important for the unconstrained estimation, and is performed for lack of detailed information on  $n_k$ ,  $n_{kd}$ ,  $w_k$  or  $w_k^M$  at such high level of granularity. It corresponds to the assumption that all HS6 goods are equally substitutable within an ISIC category, but not between. This does presumably assume some heterogeneity away, and possibly creates a bias as a result. We conjecture that heterogeneity between ISIC industries is more sizable, and thus creates more of a bias. We do however perform some robustness in section D, using all HS6 goods in the unconstrained estimation. But to do so we have to maintain some rather stringent assumptions on the values for  $n_k$ ,  $n_{kd}$ ,  $w_k$  or  $w_k^M$ .

The approach adapted from Feenstra (1994a) requires relatively little information on traded flows. To estimate equation (12) we only need measures of  $\tilde{P}_{kit}$  and the expenditure shares  $\tilde{s}_{kit}$ . As is conventional, we use unit values to approximate bilateral prices, and divide values of bilateral trade flows by their volume. In BACI, values are denominated in USD and are Free On Board.<sup>15</sup>

---

<sup>15</sup>In general, trade data are collected by national customs offices in the currency of the declaring country. These data are then converted in US dollars by the United Nations, using the current nominal exchange rate.

Quantities are in tons. The empirical model described in section A is not sensitive to the currency denomination of trade data, nor to the treatment of trade costs, as both are passed into the residuals. Expenditure shares are measured as  $\tilde{s}_{kit} = \frac{\tilde{P}_{kit} C_{kit}}{\sum_{i \neq d} \tilde{P}_{kit} C_{kit}}$ .

We subject our data to sampling with a view to limiting the role of extreme outliers. These are notoriously frequent in approaches making use of unit values to approximate prices. For instance, tonnage is not always appropriate to capture the traded volumes of all HS6 goods, which can instill artificial (massive) volatility in the resulting time series on prices. In each sector, we exclude annual variations in prices and market shares that exceed five times the median value. In addition, we impose a minimum of 20 exporters for each HS6 good over the whole observed time period. The cross-section of exporters is what ultimately achieves identification. Measurement error may prevail in estimates of  $\hat{\sigma}_k$  that are based on too few exporters, which would translate into biased values for (unconstrained) aggregate elasticities. We require that at least 20 exporters be present to alleviate this concern. Our data ultimately represent 77 percent of the total value of US imports, across 56 ISIC sectors.

In the model,  $n_k$  and  $n_{kd}$  depend directly on the import share  $w_k^M$  and the expenditure share  $w_k$ . In particular, we have

$$n_k = \frac{w_k w_k^M}{\sum_k w_k w_k^M} \text{ and } n_{kd} = \frac{w_k (1 - w_k^M)}{\sum_k w_k (1 - w_k^M)}$$

Calibration is therefore only needed for  $w_k$  and  $w_k^M$ . In the main body of the text, we consider the following data sources. The expenditure shares  $w_k$  are obtained from the OECD STAN dataset, as the 1997 ratio of sectoral absorption (value added and imports net of exports) relative to the aggregate across sectors. The import shares  $w_k^M$  are computed from the US input/output (IO) tables, available in the ISIC (Revision 3) nomenclature, as the 1997 ratio of imports over domestic gross output. Values for  $n_k$  and  $n_{kd}$  are calculated accordingly.<sup>16</sup>

In section D, we verify our results do not depend on this specific choice of data sources. We discuss four alternatives. First, we compute  $w_k^M$  directly from the BACI dataset used in our main estimation, rather than the IO tables, normalized by a measure of domestic output taken from the OECD STAN data. But we continue to compute both  $n_k$  and  $n_{kd}$  on the basis of their model-implied values. Second, the IO tables provide enough information to compute  $n_k$  directly, rather than on the basis of a model-implied formula. In our second variant, we do so, and use IO tables to calibrate both  $w_k^M$  and  $n_k$ . But  $n_{kd}$  continues to be computed according to the model, since we do not have information on domestic production. Our third variant combines both insights. We infer  $w_k^M$  from the BACI and STAN dataset, but now also use BACI to calibrate  $n_k$ . Finally, we return to our original data sources in our fourth variant, get  $w_k$  from STAN and  $w_k^M$  from the IO tables. But now, we compute sectoral absorption on the basis of gross output rather than value added.

#### IV. RESULTS AND RELEVANCE

We first review the microeconomic estimates, obtained across 56 ISIC sectors, and relate them with existing evidence. We then aggregate the estimates, preserving heterogeneity. We compare the

---

<sup>16</sup> $w_k$  and  $w_k^M$  do not sum to one because of non-traded sectors. Since  $n_k$  and  $n_{kd}$  both sum to unity by definition, we normalize each definition so that it is the case.

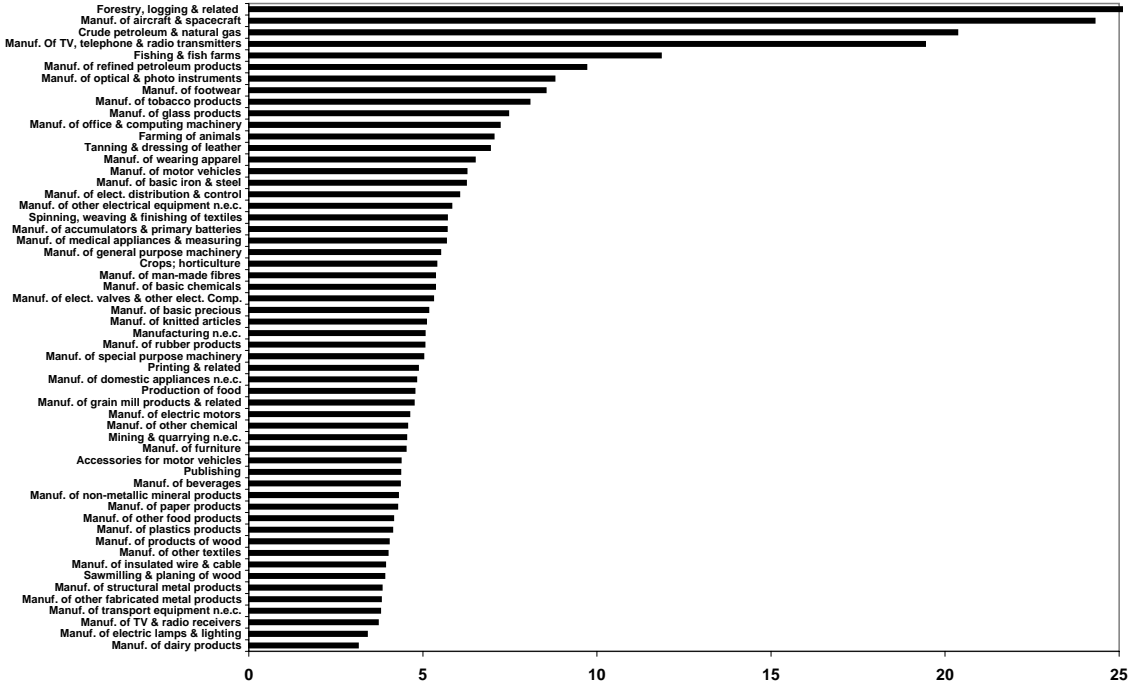


results with our estimation imposing homogeneity - performed using the same data but with the added constraint that  $\sigma_k = \sigma$ . We discuss the discrepancy, and in particular whether its magnitude is significant economically. We argue the heterogeneous estimates we uncover change dramatically the quantitative and qualitative predictions of a vast range of international macroeconomic models. We close with some robustness.

## A. Microeconomic Results

Figure 1 reports sectoral estimates of  $\hat{\sigma}_k$  for 56 ISIC sectors. On average,  $\hat{\sigma}_k$  is equal to 6.7, with values ranging from 3.1 to 28 and a standard deviation of 4.9. The median value is 5.1, reflecting a skewed distribution of elasticities: only 5 out of 56 estimates are above 10. How do our results compare with existing studies of the substitutability between foreign and domestic varieties, at similar aggregation levels? If anything, a median value of 5.1 lies at the low end of the range of estimates obtained in the empirical trade literature. Romalis (2007) finds elasticities of substitution between 4 and 13 at the HS6 level. Head and Ries (2001) find values between 7.9 and 11.4 at the 3-digit SIC level. Hanson (2005) finds estimates between 4.9 and 7.6 using data at the US county level. A common denominator across these studies is their focus on disaggregated, microeconomic information on traded quantities and/or tariffs.

Figure 1. Microeconomic estimates of the elasticity of substitution



How does the magnitude of our individual sectoral estimates compare with the literature? In theory, the parameters we estimate are comparable with the values obtained in the conventional approach regressing imported quantities on relative prices. The relative price of imports is typically measured with respect to domestically produced varieties, as in Houthakker and Magee (1969) or Kreinin

(1967). At a sectoral level, the resulting estimates of the price elasticity of imports map with  $\sigma_k$  according to  $\eta_k = 1 - \sigma_k$ . This follows from the conventional logic developed in the introduction, applied at the level of an individual sector.

Import elasticities were the object of a vast literature spread over the 1960s and 1970s. In most instances, the results focused on short-run elasticities, typically for reasons of data availability especially at a disaggregated level. In contrast our estimates correspond to long run elasticities, since the identification is in cross-section. The data available then were also coarser, and focused on just a few sectors. Still, in what follows we strive to ensure our disaggregated estimates are consistent with the existing estimates of import elasticities.

Houthakker and Magee (1969) report in their Table 6 a long run price elasticity in manufactures estimated at  $-4.05$ . This is virtually identical to the median value we obtain across our 56 manufacturing sectors, equal to  $-4.1$  ( $= 1 - 5.1$ ). Kreinin (1967) documents similar estimates, with an elasticity for manufactures equal to  $-4.71$ . It is remarkable that such different data sources, coverages and methodologies should yield strikingly similar median estimates.

The data in Houthakker and Magee (1969) and Kreinin (1967) are much coarser than ours, but they also discuss the relative magnitudes of elasticity estimates across the categories they observe. Manufactures have the higher estimates, followed by semi-manufactures and crude foods and materials. Similar relative rankings come out of the survey in Goldstein and Khan (1985). They summarize their Table 4.4 commenting that “the price elasticity of demand for manufactures is significantly larger than that for non-manufactures. Within non-manufactures, price elasticities for raw materials appear to be larger than those for food and beverages” (pages 1084-1085). A precise mapping is difficult given the differences in granularity, but the ranking is roughly prevalent in our results as well. There are exceptions, but our highest estimates concern finished manufactures, such as aircrafts, TVs, telephones, photo instruments, footwear, motor vehicles or office machinery. At the other end of the spectrum, we find relatively low elasticities for dairy, wood, food, beverages and semi-manufactures like wires or metal products.

Mapping our most disaggregated, individual sector estimates with the literature becomes quickly difficult, once again because of data availability as of 20 or 30 years ago. In fact, not many papers have attempted to estimate sector-specific price elasticities of imports, say at the two or three digit level of aggregation. We were able to identify two exceptions. Stone (1979) presents US estimates at the two digit level. On the whole, his estimates are lower than ours, but that can simply reflect his focus on short run elasticities. A few examples may nevertheless help illustrate the relative similarities in our results. For “Inorganic Chemicals”, Stone estimates an import price elasticity of  $-3.40$ , as against  $-3.60$  for “Other Chemicals” in Figure 1. He finds  $-2.22$ ,  $-2.32$ , and  $-3.71$  in “Rubber Products”, “Plastic Materials and Articles” and “Dyeing, Tanning and Coloring Agents”, as against  $-4.1$ ,  $-3.1$  and  $-5.9$  in “Rubber Products”, “Manufactures of Plastic Products” and “Tanning and Dressing of Leather” in Figure 1. Keeping in mind ours are estimates of long run elasticities, these values lie in similar ballparks.

Shiells (1991) estimates long run elasticities at the three digit SITC level, but only for 12 US sectors. Once again, an accurate mapping is impossible in most cases. Interestingly however, his estimate in “Newsprint” is  $-3.6$ , indistinguishable from our value of  $-3.4$  for “Publishing”. He also finds  $-3.5$  in “Steel Plate and Sheet”, relatively close to the estimate of  $-5.2$  we find for “Manufacturing of Basic

Iron and Steel”, even though sector definitions are different. The discrepancies become even less substantial when taking into account Shiells’ relatively large standard errors.

This comparison exercise is not meant to suggest we reproduce exactly sector-specific results that were obtained several decades ago in totally different data using drastically different methodologies. Rather, we seek to ascertain the identification strategy we follow is not fundamentally falsified. In particular, the Armington assumption is what exonerates us from having to observe any characteristics of domestically produced goods. With the Armington aggregator, the observed prices and quantities of imports originating from a cross-section of countries are sufficient to identify the elasticity of interest, between domestic and foreign varieties.

From this point of view, it is reassuring that our mean and median estimates should be strikingly close to seminal, fundamental contributions to the literature on imports price elasticities. Virtually all the papers there do make use of domestic prices in their estimations: import prices are evaluated relative to their domestically produced counterpart. That we should find similar results without any information on domestic prices brings support to the Armington assumption. The few punctual comparisons we report at the level of individual sectors do, as well.

There is of course an obvious comparison absent from our analysis so far. We have implemented a variant of the methodology introduced by Feenstra (1994a), just as Broda and Weinstein (2006) have. Our objectives are fundamentally different, as are some of our identifying restrictions and some of the corrections we introduce. Still, Broda and Weinstein estimate the universe of substitution elasticities in disaggregated US data, just as we do. Given the similarities in methodologies, it is not surprising that our estimates should be similar, and they are. Their mean estimate at the three digit level is 4.0, with a standard deviation of 7.9. “Petroleum Oils and Oils from Bituminous Minerals, Crude”, “Aircraft and Associated Equipment” or “Fuel Wood” are sectors with relatively high elasticities, whereas “Lighting Fixtures”, “Radio-Broadcast Receivers” or “Motorcycles and Cycles” all rank towards the bottom of their list.

But the comparison is not especially informative in terms of validating our assumptions. What is key here is the Armington assumption that the substitutability between two foreign varieties should equal that between domestic and foreign varieties. If this is true in the data, we can infer directly the price elasticity of imports. Since the latter has been the object of a vast empirical literature, it is with it that we have striven to compare our results.

## B. Macroeconomic Results

We now turn to macroeconomic estimates of the elasticity of substitution implied by the values in Figure 1. To do so, we apply the aggregation procedure spelled out in the previous sections. Our main point concerns the difference in estimates of  $\eta_{NoFirm}$  and  $\sigma_{NoFirm}$  where  $\sigma_k = \sigma$  as against those where  $\sigma_k$  is left unconstrained. Since we have

$$\sigma_{NoFirm} = \sum_k n_{kd} \sigma_k + \sum_k (n_k - n_{kd}) (\sigma_k - \gamma) (1 - w_k^M)$$

the heterogeneity bias increases in the correlations between  $\sigma_k$  and  $n_k$ . In our data, the correlation is mildly positive. This suggests constraining all elasticities to homogeneity will act to lower estimates of both  $\sigma_{NoFirm}$  and  $\eta_{NoFirm}$ . The second summation has second order effects only, so that the

Table 1. Estimation with common correlated effects

	Import Elasticity $\eta_{NoFirm}$	Substitution Elasticity $\sigma_{NoFirm}$
Constrained total elasticity	-1.980 <sup>a</sup> (.175)	4.124 <sup>a</sup> (0.300)
Constrained partial elasticity	-2.738 <sup>a</sup> (.262)	3.738 <sup>a</sup> (0.263)
Unconstrained total elasticity	-4.508 <sup>a</sup> (.745)	7.226 <sup>a</sup> (0.962)
Unconstrained partial elasticity	-6.553 <sup>a</sup> (1.100)	6.921 <sup>a</sup> (0.697)
Number of sectors	56	56
Number of grid searches	11	11

Note: Standard errors in parentheses (obtained by bootstrapping for grid searched sectors), <sup>a</sup> denotes significance at the 1% level.

choice of a value for  $\gamma$  has minimal effect on end estimates of  $\sigma_{NoFirm}$  and  $\bar{\sigma}_{NoFirm}$ . From its definition in equation (8), the calibration of  $\gamma$  does however have first order effects on the level of  $\eta_{NoFirm}$ , although not on the discrepancy between  $\eta_{NoFirm}$  and  $\bar{\eta}_{NoFirm}$ . We set  $\gamma$  at its minimal value of 1, so that our results correspond to upper bounds (in absolute values) of  $\eta_{NoFirm}$  and  $\bar{\eta}_{NoFirm}$ .

Table 1 reports estimates of both aggregate elasticities. We first report estimates of  $\bar{\eta}_{NoFirm}$ , the total price elasticity of imports, when we impose that  $\sigma_k$  be equal across all sectors. Our point estimate suggests a value for the parameter of -1.98. A confidence interval at standard significance levels implies values ranging roughly from -1.6 to -2.3. This is at the high end of the range of values obtained in conventional estimates of the elasticity based on macroeconomic data. For instance, Goldstein and Kahn (1985) claim that “Harberger’s (1957) judgment of 25 years ago that the price elasticity of import demand for a typical country lies in or above the range of -0.5 to -1.0 still seems on the mark”. In their Table 4.1, they report estimates for the US between -1.03 and -1.76. The interval is not significantly distinct from the one we estimate, even though the calibration of  $\gamma$  implies an upper bound estimate of  $\bar{\eta}_{NoFirm}$ . Choosing higher values for  $\gamma$  would only bring our results closer to the literature.

We obtain an aggregate estimate with nothing but import prices, that is not significantly different from one obtained on the basis of relative prices computed with domestic price indices. More importantly, -1.98 is consistent with the choices made in the vast majority of calibration exercises in international macroeconomics. Following the reasoning we develop in the introduction, macroeconomic calibrations infer the elasticity of substitution from  $1 - \eta$ , where the import elasticity is estimated on aggregate *data*. Here, the implied elasticity of substitution is between two and three, which includes the ranges of values used in, say, Obstfeld and Rogoff (2005) or Backus, Kehoe and Kydland (1994).

Constrained estimates stand in contrast with the value of  $\eta_{NoFirm}$  obtained when  $\sigma_k$  is left unconstrained across sectors. As shown in the Table, the parameter jumps to -4.5, with standard errors that guarantee a significant difference at conventional confidence levels. The right panel of the Table reports the theoretical values for  $\sigma_{NoFirm}$ , which we find is in fact around 7. We argue this is the value that should enter the utility of a representative agent with heterogeneous preferences across

sectors. Given the overwhelming evidence that substitutability is heterogeneous across goods or sectors, we contend a value around 7 is preferable from a calibration standpoint. Such a high value characterizes adequately the average substitutability of a representative agent who has heterogeneous preferences across goods. Aggregating the data, instead of aggregating estimates, gives rise to a heterogeneity bias. With heterogeneity, the response of aggregate quantities estimated from aggregated data is not indicative of the average elasticity of substitution.

It is useful to check the heterogeneity bias we document continues to prevail in estimates of the partial import elasticity. We report its constrained and unconstrained values in Table 1. As is patent, a bias continues to prevail, with a constrained estimate at -2.7, jumping to -6.5 when sector specific elasticities are permitted. The adjustment of price indices does not explain our results away. But as expected, it tends to dilute measured elasticity, as partial import elasticities are systematically higher.

We finally verify an aggregated version of our data continues to imply estimates of  $\sigma$  that are commensurate with results in the literature. We implement our estimator on our dataset, aggregated to the country level. Aggregate U.S. imports are given by a simple sum across sectors of the values imported from a given country, and aggregate import prices are computed as the chained Tornqvist index of HS6 specific prices. Identification continues to rest on the cross-section of exporters to the U.S, and requires that aggregate bundles of imported goods be different varieties of the same good, with elasticity of substitution  $\sigma$ . The assumption is palatable at a disaggregated level, but hard to maintain for macroeconomic data, especially if countries are specialized. The bundle of goods exported to the U.S. by developing economies is likely to differ fundamentally from that exported by the developed world, a difference that goes beyond imperfect substitutability between varieties of the same good. We try and alleviate the concern, and focus on a cross-section of 24 high income OECD exporters to the U.S, where the composition of exports is presumably relatively homogeneous.

Based on this cross-section of 24 countries, we obtain an estimate for  $\sigma$  equal to 1.34. The corresponding price elasticity of imports is close to zero, consistent with conventional aggregate results. The result is reassuring, for it confirms an aggregated version of our data yields perfectly standard estimates. Our data therefore mirror the general conclusions of the literature we have described, with large differences between micro and macroeconomic estimates. Two words of caution are however in order. First, aggregating the data obscures considerably the ceteris paribus nature of our experiment. While estimates of  $\sigma_{NoFirm}$  and  $\bar{\sigma}_{NoFirm}$  cannot differ for other reasons than a heterogeneity bias, a number of other parameters are affected by the aggregation of data, rather than of estimates. The dimensionality and sample size of our data are different, as are the variance-covariance properties of import prices and quantities.

Second, aggregating the data effectively forces the substitutabilities between varieties and between sectors to be equal. In our notation, this implies  $\sigma = \gamma$ , as we show in Appendix A. As all sectors sum up to aggregate imports, it is by construction impossible to distinguish substitution between two varieties of the same sector across countries, or between two different sectors across countries. There are therefore two hypotheses implicit in the use of aggregate data: the homogeneity constraint  $\sigma_k = \sigma$  we have discussed, and the additional assumption that  $\sigma = \gamma$ . On the basis of the vast heterogeneity in existing estimates of  $\sigma_k$ , we have so far speculated the former is quantitatively important. Since our methodology is not equipped to provide structural estimates of  $\gamma$ , we have ignored the latter. In our data, the difference between  $\sigma_{NoFirm}$  and  $\bar{\sigma}_{NoFirm}$  - the heterogeneity bias - dwarves that between  $\bar{\sigma}_{NoFirm}$  and 1.34. Sample issues notwithstanding, it is the heterogeneity bias that dominates.

### C. Relevance

Is the correction we document relevant in economic terms? We now discuss the quantitative and qualitative consequences of using heterogeneous estimates to calibrate macroeconomic models with an international dimension. We are careful to focus on theories with CES preferences, since our estimates are predicated on the assumption. A straightforward implication concerns models directly dealing with the resolution of global imbalances. Most prominently, Obstfeld and Rogoff (2005) use a calibrated model to argue a reversal of the US current account is compatible with a 30% depreciation of the real exchange rate. The calibration sets substitutability at 2. In a slightly simplified two-country version, we obtained depreciation rates of 22 and 21% for values of the parameter of 5 and 7, respectively, down from 31% with an elasticity of 2.<sup>17</sup> The parameter is quantitatively important, and shaves off one third of the “required” depreciation, almost all the way to the 19.3% that obtains for an elasticity of 100. This is true even though Obstfeld and Rogoff’s calibration gives prominence to another parameter, the elasticity of substitution between traded and non-traded goods, important in this instance because the US is a largely closed economy. Still, the effects are sizeable and probably important in terms of welfare as well.

Cole and Obstfeld (1991) show the endogenous response of the terms of trade can deliver perfect insurance against country-specific shocks when the elasticity of substitution between domestic and foreign goods is exactly unitary. This is a necessary but not sufficient condition, as perfect insurance also imposes restrictions on other parameters, for instance the intertemporal elasticity of substitution. The result is meant as an illustrative special case of a powerful mechanism. Still, models of international portfolio holdings have drastically different qualitative predictions depending on the magnitude of the parameter. For instance, Heathcote and Perri (2008) show that a model with CES preferences can generate a home equity bias for low enough values of  $\sigma$ . A positive domestic productivity shock will increase the relative return on domestic stocks as long as the terms of trade do not respond too strongly. In a sensitivity analysis, they show portfolio home bias obtains for values of the elasticity of substitution below 4. A contrario, in Coeurdacier (2009), domestic consumers choose to hold foreign assets to insure against shocks to domestic consumption, provided the terms of trade respond strongly enough in response to real shocks, which in his calibration happens for  $\sigma = 5$ . His sensitivity analysis suggests the conclusion remains robust for values of  $\sigma > 1$ . Both papers then move on to introduce labor income risk and endogenous production, or incomplete markets. Our purpose here is not to settle the question of the origins of an equity home bias. This depends of course on other calibration choices, not least the intertemporal elasticity of substitution or the existence of trade costs. It is however interesting that most papers in this literature conduct extensive sensitivity analyses to the choice of  $\sigma$ , with far-reaching implications on the models’ end predictions.

The policy consequences of international price differences will presumably also depend on the substitutability between domestic and foreign goods. The relevance of the exchange rate in the monetary policy rule developed in Galí and Monacelli (2005) is affected by the parameter. Galí and Monacelli focus on unitary elasticity, so the result is not directly apparent there. But De Paoli (2009) introduce a generalization of their model, and her conclusions point to that direction. As in Cole and Obstfeld (1991) with unitary elasticity, a marginal reduction in the utility value of output is accompanied by an exactly offsetting reduction in the utility value of consumption. This insulates

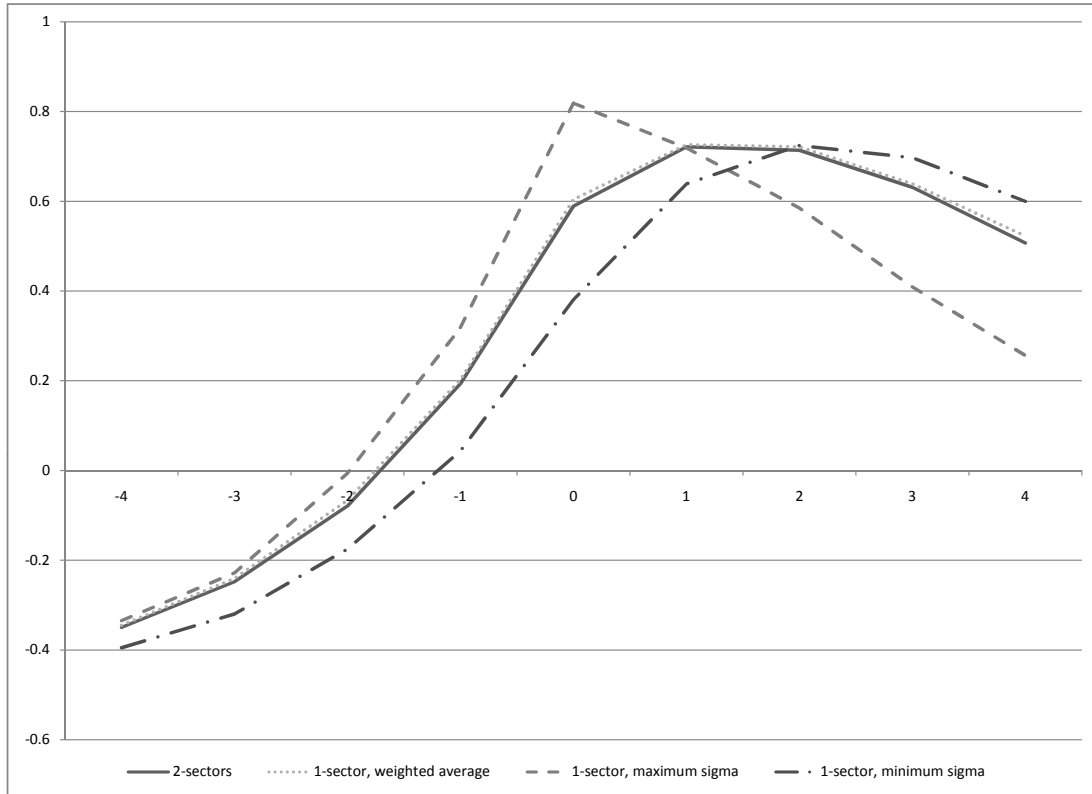
---

<sup>17</sup>We are grateful to Cedric Tille for graciously giving us the simulation code.

the economy from terms of trade movements. With non unitary elasticity however, policy shocks that affect the terms of trade also affect welfare, in a way that crucially depends on whether the calibrated parameter is above or below one. The fact that we find an aggregate estimate substantially above one must therefore have important policy implications.<sup>18</sup>

The elasticity of substitution of a representative agent living in a one-sector model should be aggregated adequately on the basis of the microeconomic elasticities that we know are heterogeneous. We have shown estimating that average on the basis of macroeconomic data can be misleading. Our recommendation is that  $\sigma_{NoFirm}$  should be preferred to  $\bar{\sigma}_{NoFirm}$  in calibrating one-sector theoretical economies if one wants to capture the fact that  $\sigma_k$  are heterogeneous in the data. We close with a simple exercise to validate this claim. We construct a two-sector version of a classical model in international economics, due to Backus, Kehoe and Kydland (1994) [BKK], where the sources of heterogeneity are sector-specific elasticities of substitution between domestic and foreign goods, along with  $w_k$  and  $w_k^M$ . We calibrate this version of the model using our sectoral results, and simulate a J-curve from it. We then compare this prediction with the conventional one-sector model. We ask what value of the (single) elasticity of substitution in the one sector model reproduces the J-curve that is implied by the calibrated two-sector version. We expect an adequately weighted average of the calibrated values of  $\sigma_k$  to come closest to the dynamics implied by the multi-sector version.

Figure 2. The J-curve in a two-sector BKK model



<sup>18</sup> Admittedly, other parameters will also determine welfare in this context - not least once again the intertemporal elasticity of substitution.

Since the workings of the model are well known, we leave a description of the details to Appendix F. We compare the relative performances of four models. First, a two-sector version of BKK, calibrated on our data and our results. In particular, we choose  $(\sigma_1; \sigma_2) = (4.8; 12.9)$ , which corresponds to below- and above-mean averages of  $\sigma_k$ . The second and third models use  $\sigma_1$  and  $\sigma_2$ , respectively, in conventional one-sector versions of BKK. Both versions are meant to capture a macroeconomic calibration fully ignorant of heterogeneity issues.<sup>19</sup> Finally, we calibrate a one-sector version of BKK using a weighted average of  $\sigma_1$  and  $\sigma_2$ ,  $\sum_{k=1,2} n_{kd} \sigma_k$ , consistent with the allowance for heterogeneity we have argued matters quantitatively. In all models, we calibrate  $w_k$  and  $w_k^M$  using our data. Figure 2 reports the J-curves implied by the four models. As is patent, the one-sector version of BKK that best matches the dynamics of the trade balance implied by the two-sector model is one that accounts for heterogeneity in the manner that we have described in this paper.

## D. Stability

This section verifies the robustness of our results in three dimensions. First, we ascertain our results do not depend on a particular choice of data source in computing  $w_k$  and  $w_k^M$ . Second, we investigate the importance of “Common Correlated Effects” in obtaining estimates of  $\sigma_k$ . Third, we relax our assumption that elasticities of substitution be identical across the HS6 categories regrouped in each ISIC sector. Instead, like Broda and Weinstein (2006) we estimate a value of  $\sigma_k$  for each HS6 category. We discuss the necessary shortcuts this requires in terms of aggregation.

Table 2 compares the constrained and unconstrained values of the total elasticities  $\sigma_{NoFirm}$  and  $\eta_{NoFirm}$  using different weighting vectors. The data sources and computations behind the four alternative variants we present in the Table are discussed in Section 3.3. The first line repeats the results implied by the benchmark weights we have used so far. Across the four variants, constrained estimates of  $\eta_{NoFirm}$  range around  $-2$  and are not significantly distinguishable from conventional estimates, for instance in Goldstein and Kahn (1985). Unconstrained estimates reach  $-5$ . The bias continues to be quantitatively important across these four alternatives.

The inclusion of Common Correlated Effects in the estimation of  $\sigma_k$  is justified by our interest in the macroeconomic implications of the microeconomic values we obtain. After all, the quantities traded at sector level, and their prices, do presumably respond to common, aggregate, macroeconomic influences. When estimating the sector-specific substitutability between domestic and foreign varieties, one wants to ascertain one is not capturing aggregate dynamics. This would amount to double-counting at the time of aggregation. Does this matter in our estimations? Table 3 provides a mixed answer, using again our benchmark weights. Without CCE, the constrained estimate of  $\eta$  decreases slightly, to  $-2.17$ , whereas the unconstrained estimate increases slightly, to  $-4.08$ . These changes are not strongly significant relative to our benchmark results, and they do not alter the conclusion of a significant heterogeneity bias. But they nevertheless suggest the introduction of a CCE term in equation (12) is not innocuous.

Finally, we relax our assumption that the substitutability between two HS6 categories be identical within each ISIC sector. In other words, we allow for heterogeneity in  $\sigma_k$  even within each ISIC

---

<sup>19</sup>We also experimented with the arithmetic average of  $\sigma_1$  and  $\sigma_2$  with similar conclusions. The arithmetic average is not necessarily an adequate proxy for  $\bar{\sigma}$ . We have discussed in section 3.2 the possibility that macroeconomic estimates of  $\bar{\sigma}$  should suffer from an heterogeneity bias of an econometric nature.



Table 2. Variants on the weights

	Import elasticity		Substitution Elasticity	
	Unconstrained	Constrained	Unconstrained	Constrained
Benchmark	-4.51	-1.98	7.22	4.12
Variant 1	-5.17	-2.21	6.93	4.05
Variant 2	-4.38	-2.08	7.36	4.02
Variant 3	-4.60	-2.15	6.77	4.06
Variant 4	-4.41	-2.10	7.27	4.12

Note: Benchmark:  $w_k^M$  using imports and output from IO tables,  $\alpha_k$  using STAN sectoral interior demand. Variant 1:  $w_k^M$  using imports from BACI and output from STAN,  $\alpha_k$  using STAN sectoral interior demand. Variant 2:  $n_k$  and  $w_k^M$  using imports and output from IO tables,  $\alpha_k$  using STAN sectoral interior demand. Variant 3:  $n_k$  and  $w_k^M$  using imports from BACI and output from STAN,  $\alpha_k$  using STAN sectoral interior demand. Variant 4:  $w_k^M$  using imports and output from IO tables,  $\alpha_k$  using STAN sectoral interior demand (absorption in terms of value added).

Table 3. Estimation without common correlated effects

	Import Elasticity	Substitution Elasticity
	$\eta_{NoFirm}$	$\sigma_{NoFirm}$
Constrained total elasticity	-2.166 <sup>a</sup> (.150)	4.442 <sup>a</sup> (.257)
Constrained partial elasticity	-3.016 <sup>a</sup> (.225)	4.016 <sup>a</sup> (.225)
Unconstrained total elasticity	-4.075 <sup>a</sup> (.112)	6.584 <sup>a</sup> (0.145)
Unconstrained partial elasticity	-5.946 <sup>a</sup> (.209)	6.321 <sup>a</sup> (0.138)
Number of sectors	56	56
Number of grid searches	12	12

Note: Standard errors in parentheses (obtained by bootstrapping for grid searched sectors), <sup>a</sup> denotes significance at the 1% level.

sector. Like Broda and Weinstein (2006), we estimate an elasticity of substitution for each HS6 sector, and then use equation (4) to aggregate them at the macroeconomic level. This raises the question of what values for  $w_k$  and  $w_k^M$  to use: we do not observe any of these weights at such a refined aggregation level. We choose to impose similar values of the weights for all HS6 categories that belong to one ISIC sector. Clearly, this assumes away some possible source of a heterogeneity bias, but there is simply no alternative. But we know choosing other values for  $w_k$  and  $w_k^M$  do not matter at the ISIC level. At least, this tests whether the heterogeneity in estimates of  $\sigma_k$  within each ISIC category can be such that our conclusions are altered.

Naturally, the constrained estimates of  $\eta_{NoFirm}$  continue to be identical, for instance at -2.17 without a CCE corrective term. After all, this is an estimation that constrains all coefficients to be identical, within and between ISIC categories. The difference arises for unconstrained estimates. We estimated values for  $\sigma_k$  in 4,021 HS6 categories, and aggregated them using our benchmark (ISIC) weights. We obtained a value of -5.23 for the unconstrained elasticity. Computing standard error bands across this point estimate is not tractable, but the aggregation bias appears to be even stronger. We conclude ignoring heterogeneity within ISIC sectors is not important to our conclusions.

## V. CONCLUSION

The elasticity of substitution between domestic and foreign varieties is central in international economics. But no clear consensus has emerged from a vast empirical literature seeking to pin down the parameter, except for one essential finding. On average, microeconomic data tend to imply substantially more heterogeneous values than macroeconomic aggregates. We propose that this heterogeneity is the reason why aggregate results are close to zero. We compute structural estimates of aggregate substitutability allowing or not for heterogeneity at the sectoral level. We find that imposing homogeneity is enough to obtain aggregate estimates in line with the macroeconomic evidence, even using a disaggregated dataset. Allowing for heterogeneity results in an aggregate parameter value of up to 7. This discrepancy validates the conjecture of a heterogeneity bias in elasticity estimates that goes back at least to Houtakker and Magee (1969). Such high parameter values change dramatically the conclusions of calibrated models in areas of international economics as varied as global imbalances, international risk sharing, portfolio choice and optimal monetary policy.

## REFERENCES

- Ai, Chunrong and Edward Norton, 2003. "Interaction terms in logit and probit models," *Economics Letters*, Vol. 80, No. 1, pp. 123-129.
- Anderson, J. and E. Van Wincoop, 2004. "Trade Costs", *Journal of Economic Literature*, Vol. 84, No. 1, September.
- Backus, D., P. Kehoe and F. Kydland, 1994. "Dynamics of the Trade Balance and the Terms of Trade: The J-Curve?" *American Economic Review*, Vol. 84, No. (1), pp. 84-103.
- Blonigen, B. and W. Wilson, 1999, "Explaining Armington: What Determines Substitutability Between Home and Foreign Goods?" , *Canadian Journal of Economics*, Vol. 32, No. 1, pp. 1-21.
- Broda, C. and D. Weinstein, 2006, "Globalization and the Gains from Variety", *The Quarterly Journal of Economics*, Vol. 121, No. 2, pp. 541-585.
- Chaney T., 2008, "Distorted Gravity: The Intensive and Extensive Margins of International Trade", *American Economic Review*, Vol. 98, No. 4, pp. 1707-21.
- Chari, V.V., P. Kehoe and E. McGrattan, 2002, "Can Sticky Price Models Generate Volatile and Persistent Real Exchange Rates?", *Review of Economic Studies*, Vol. 69, No. 3, pp. 533-563.
- Coeurdacier, N., 2009, "Do trade costs in goods market lead to home bias in equities?", *Journal of International Economics*, Vol. 77, pp. 86-100.
- Coeurdacier N., R. Kollmann and P. Martin, 2007, "International Portfolios with Supply, Demand and Redistributive Shocks", NBER Working Papers 13424.
- Cole, H. and M. Obstfeld, 1991, "Commodity trade and international risk sharing : How much do financial markets matter?", *Journal of Monetary Economics*, Vol. 28, No. 1, pp. 3-24.
- De Paoli, B., 2009, "Monetary Policy and Welfare in a Small Open Economy", *Journal of International Economics*, Vol.77, No.1, February.
- Feenstra, R., 1994a, "New Product Varieties and the Measurement of International Prices", *American Economic Review*, Vol. 84, No. 1, pp. 157-177.
- Feenstra, R., 1994b, "New Goods and Index Numbers: U.S. Import Prices", NBER Working Papers 3610.
- Gagnon, J., 2003, "Productive capacity, product varieties, and the elasticities approach to the trade balance", International Finance Discussion Papers 781.
- Gali, J. and Monacelli, T., 2005, "Monetary Policy and Exchange Rate Volatility in a Small Open Economy", *Review of Economic Studies*, Vol. 72, No. 3, pp. 707-734.
- Goldstein M. and M. Kahn, 1985, "Income and price effect in foreign trade", In: J. Ronald and P. Kennen, Editors, *Handbook of International Economics*, North-Holland, Amsterdam, pp. 1042-1099.

- Hanson, G., 2005, "Market Potential, Increasing Returns and Geographic Concentration", *Journal of International Economics*, Vol. 67, No. 1, pp. 1-24.
- Harberger, A., 1957, "Some Evidence on the International Price Mechanism", *Journal of Political Economy*, Vol. 65, No. 6, pp. 506-521.
- Head, K. and J. Ries, 2001, "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade", *American Economic Review*, Vol. 91, No. 4, pp. 858-876.
- Heathcote, J. and F. Perri, 2008, "The International Diversification Puzzle Is Not As Bad As You Think", NBER Working Papers 13483.
- Hooper, P., K. Johnson and J. Marquez, 1998, "Trade elasticities for G-7 countries", International Finance Discussion Papers 609.
- Houthakker, H. and S. Magee, 1969, "Income and Price Elasticities in World Trade", *The Review of Economics and Statistics*, Vol. 51, No. 2, pp. 111-25.
- Huddle, D., 1970, "Some Evidence on Price Elasticities in International Trade", *Oxford Economic Papers*, Vol. 22, No. 2, pp. 235-42.
- Kahn, M., 1975, "The Structure and Behavior of Imports of Venezuela", *The Review of Economics and Statistics*, Vol. 57, No. 2, pp. 221-24.
- Kemp, M., 1962, "Error of Measurements and Bias in Estimates of Import Demand Parameters", *Economic Record*, Vol. 38, pp. 369-372.
- Kreinin, M., 1967, "Price Elasticities in International Trade", *Review of Economics and Statistics*, Vol. 49, pp. 510-16.
- Krugman, P., 1980, "Scale Economies, Product Differentiation, and the Pattern of Trade", *American Economic Review*, Vol. 70, No. 5, pp. 950-59.
- Marquez, J., 1990, "Bilateral Trade Elasticities", *The Review of Economics and Statistics*, Vol. 72, pp. 70-77.
- Melitz, M., 2003, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, Vol. 71, No. 6, pp. 1695-1725.
- Obstfeld, M. and K. Rogoff, 2000, "The Six Major Puzzles in International Finance: Is There a Common Cause?", *NBER Macroeconomics Annual*, 15.
- Obstfeld, M. and K. Rogoff, 2005, "Global Current Account Imbalances and Exchange Rate Adjustments", *Brookings Papers on Economics Activity*, Vol. 2005, No. 1, pp. 67-146.
- Orcutt, G., 1950, "Measurement of Price Elasticities in International Trade", *The Review of Economics and Statistics*, Vol. 32(2): 117-132.
- Pesaran, H., 2006, "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure", *Econometrica*, Vol. 74, No. 4, pp. 967-1012.
- Pesaran, H. and R. Smith, 1995, "Estimating Long Run Relationships from Dynamic Heterogeneous Panels", *Journal of Econometrics*, Vol. LXVIII, pp. 79-113.

- Romalis, J., 2007, "NAFTA's and CUSFTA's Impact on International Trade", *The Review of Economics and Statistics*, Vol. 89, No. 3, pp. 416-435.
- Ruhl, K., 2005, "Solving the Elasticity Puzzle in International Economics", mimeo.
- Segura-Cayuela, R. and J. Vilarrubia, 2008, "Uncertainty and entry into export markets", Banco de Espana Working Papers 0811.
- Shiells, J., 1991, "Errors in Import-Demand Estimates Based upon Unit-Value Indexes", *The Review of Economics and Statistics*, Vol. 73, No. 2, pp. 378-382.
- Stone, J., 1979, "Price Elasticities of Demand for Imports and Exports: Industry Estimates for the U.S., the E.E.C. and Japan", *The Review of Economics and Statistics*, Vol. 61, No. 2, pp. 306-312.

## APPENDIX A. AGGREGATION AND HOMOGENEITY

We want to replicate what is implied by an estimation of the substitutability between the aggregate bundles of domestic and foreign goods, performed on aggregate data. The nominal demand for the bundle of goods imported from country  $i$  is expressed in terms its relative price, and overall demand. In our notation:

$$P_i C_i \equiv \sum_k^K P_{ki} C_{ki} = \left( \frac{P_i}{P} \right)^{1-\sigma} PC, \quad (\text{A.1})$$

where  $PC = \sum_k^K P_k C_k$ ,  $P_i$  is the ideal price index corresponding to the aggregate quantities imported from country  $i$ , and  $P$  is the aggregate domestic price index. By definition,  $P_i$  aggregates sector specific prices, using the elasticity of substitution between sectors  $\gamma$ ; and  $P$  aggregates domestic and import prices, using the international elasticity  $\sigma$ . By construction, it is impossible for  $\sigma$  to have a sector dimension. Cross-sector aggregation is performed first, so that the only international substitutability that can prevail is between aggregate bundles of goods. The very notion of an ideal price index for aggregate imports precludes sector heterogeneity in the cross-country elasticity  $\sigma$ . Equation (A.1) is based on an aggregation of the *data*. Instead, our approach in this paper is based on an aggregation of elasticity *estimates*, which is flexible enough that it can preserve heterogeneity in all elasticities, and in particular in  $\sigma_k$ .

Suppose now the world is well characterized by heterogeneous  $\sigma_k$ . We seek to establish what parametric constraints are imposed through the use of aggregate data. We start from our CES model, where goods produced by different exporting countries are substitutable at the industry level. The nominal demand addressed to producers of good  $k$  located in country  $i$  writes

$$P_{ki} C_{ki} = \left( \frac{P_{ki}}{P_k} \right)^{1-\sigma_k} P_k C_k, \quad (\text{A.2})$$

where  $\frac{P_{ki}}{P_k}$  is the relative price of the variety produced in country  $i$ ,  $\sigma_k$  is the sector-specific elasticity of substitution and  $P_k C_k$  is nominal demand for good  $k$ . For simplicity, we set the preference shocks  $\beta_{ki}$  to zero. We seek to replicate the setup in equation (A.1) in the context of our disaggregated model. In logarithms, equation (A.2) implies

$$\begin{aligned} \ln P_i C_i &\equiv \ln \sum_k^K P_{ki} C_{ki} \\ &= \ln \sum_k^K \left( \frac{P_{ki}}{P_k} \right)^{1-\sigma_k} P_k C_k \\ &\approx \ln \frac{\sum_k^K P_{ki}^{1-\sigma_k}}{\sum_k^K P_k^{1-\sigma_k}} + \ln \sum_k^K P_k C_k, \end{aligned}$$

where we use (twice) the property that  $\ln \sum_k^K x_k \approx \sum_k^K \ln x_k$  for large  $K$ . We seek to recognize the relative price of aggregate imports in the ratio  $\frac{\sum_k^K P_{ki}^{1-\sigma_k}}{\sum_k^K P_k^{1-\sigma_k}}$ . A well defined ideal price index for aggregate imports exists if we have  $\sigma_k = \gamma$ . This effectively embeds two constraints. First, the international elasticities of substitution  $\sigma_k$  are homogeneous across sectors. Second, international and cross-sector elasticities are equal. In the body of the paper, we constrain the estimates of  $\sigma_k$  to

homogeneity. But our estimation approach is silent about  $\gamma$ , and so we cannot investigate the effect of that additional constraint. We speculate however heterogeneity in  $\sigma_k$  is empirically most relevant, if only because it is documented in decades of applied work. We verify the claim in our data: The estimate for  $\sigma$  we obtain from an aggregated version of our data is in fact close to the results imposing  $\sigma_k = \sigma$  reported in the body of the paper.

## APPENDIX B. HETEROGENEOUS SUPPLY

Here we present the now classic model of production with heterogeneous firms making entry decisions, inspired from Melitz (2003). The theory is written for a given sector  $k$  located in country  $i$ , and indexes are dropped to facilitate the exposition. All prices expressed in the exporter's currency are denoted with an asterisk \*. Firms in each sector  $k$  and country  $i$  are monopolistically competitive, produce differentiated varieties and are heterogeneous in terms of productivity. They decide whether to pay a fixed cost  $f^e$  prior to knowing their own productivity, which allows them to export to foreign markets. Once the cost is sunk, productivity  $\varphi$  is drawn from a distribution with probability density function  $g(\varphi)$ , and an associated cumulated distribution  $G(\varphi)$ . Productivity is revealed, and the firm decides whether to produce for the foreign market, with a technology featuring constant marginal costs and a fixed overhead export cost  $f$ . Both costs are expressed in terms of the unique composite factor of production, whose price is denoted with  $w^*$ . Finally, exporting involves an iceberg cost  $\tau$ , potentially specific to each sector and country. Note that for  $i \in I$  this setup departs slightly from the conventional Melitz (2003) model, where the decision to produce for the domestic market predates that to export. We assume there are fixed costs involved in gathering information on export markets, that are quite separate from those involved in a purely domestic production decision.<sup>20</sup>

Total costs in exporter's currency are given by

$$TC^*(\varphi) = w^*f + \frac{w^*\tau c(\varphi)}{\varphi}$$

where  $c(\varphi)$  is the equilibrium demand addressed to the firm by the foreign market. Individual costs are decreasing in  $\varphi$ , as firms with higher productivity can produce the same amount of output with fewer workers. Firms face a probability  $\delta$  of exogenous exit in each period. Under these assumptions, the decision to enter export markets and pay the sunk cost  $f^e$  is governed by the following maximization of the (probability) discounted value of profits

$$v(\varphi) = \max \left\{ 0, \sum_{t=s}^{\infty} (1-\delta)^{t-s} \pi^*(\varphi) \right\} = \max \left\{ 0, \frac{1}{\delta} \pi^*(\varphi) \right\} \quad (\text{B.1})$$

where profits, expressed in the exporter's currency, are given by

$$\pi^*(\varphi) = \left( p^*(\varphi) - \frac{\tau w^*}{\varphi} \right) c - w^*f$$

There is a unique threshold productivity  $\bar{\varphi}$  such that  $v(\varphi) > 0$  if and only if  $\varphi > \bar{\varphi}$ . Firms that are productive enough remain in the market after having drawn their productivity and produce in every

<sup>20</sup>See for instance Segura-Cayuela and Villarubia (2008).

period. Their export price is set by maximizing profits under a demand constraint given in the body of the paper by equation (2), i.e.

$$p^*(\varphi) = \frac{\rho}{\rho - 1} \frac{\tau w^*}{\varphi} \quad (\text{B.2})$$

As Melitz (2003), we introduce a weighted average productivity measure for firms active in the market

$$\tilde{\varphi} = \left[ \frac{1}{1 - G(\bar{\varphi})} \int_{\bar{\varphi}}^{\infty} \varphi^{\rho-1} g(\varphi) d\varphi \right]^{\frac{1}{\rho-1}}$$

Then define average profits  $\tilde{\pi}^*$ :

$$\begin{aligned} \tilde{\pi}^* &\equiv \frac{1}{1 - G(\bar{\varphi})} \int_{\bar{\varphi}}^{\infty} \pi^*(\varphi) g(\varphi) d\varphi \\ &= \frac{1}{1 - G(\bar{\varphi})} \frac{1}{\rho} c \left( \frac{p}{E} \right)^{\rho} \left( \frac{\rho}{\rho - 1} \tau w^* \right)^{1-\rho} \int_{\bar{\varphi}}^{\infty} \varphi^{\rho-1} g(\varphi) d\varphi - w^* f \end{aligned}$$

where  $E$  is the bilateral exchange rate, expressed in terms of the importer's currency. Since the threshold firm makes no profit, average profits simplify into

$$\tilde{\pi}^* = w^* f \left[ \left( \frac{\bar{\varphi}}{\tilde{\varphi}(\bar{\varphi})} \right)^{1-\rho} - 1 \right] \quad (\text{B.3})$$

Equation (B.3) is a “Zero Cutoff Profit” (ZCP) condition. It delineates a relation between average profits  $\tilde{\pi}^*$  and the productivity cut-off  $\bar{\varphi}$ . In particular, when the distribution of productivity follows a Pareto distribution,  $G(\varphi) = 1 - \left( \frac{b}{\varphi} \right)^{\kappa}$ , the ZCP schedule is flat and given by

$$\tilde{\pi}^* = w^* f \left( \frac{\kappa}{\rho - \kappa + 1} - 1 \right)$$

The ZCP condition characterizes a short run equilibrium, where the threshold productivity  $\bar{\varphi}$  is determined in function of the model's parameters. It is a short run equilibrium in that it takes firms' location decisions, and thus the potential number of firms active in the market, as given.

A long run equilibrium obtains when a free entry condition holds, that ensures the expected discounted value of profits for a potential entrant equals the fixed cost of entry in the export market. The condition will pin down both threshold productivity and the number of firms active in the export market. Rearranging the condition  $\int_0^{\infty} v(\varphi) g(\varphi) d\varphi = f^e$  implies

$$\tilde{\pi}^* = \frac{\delta f^e}{1 - G(\bar{\varphi})}$$

When productivity is distributed according to a Pareto distribution, the Free Entry (FE) condition becomes

$$\tilde{\pi}^* = \delta w^* f^e \left( \frac{\bar{\varphi}}{b} \right)^{\kappa}$$

The ZCP and FE conditions define a system of two equations in two unknowns,  $\tilde{\pi}^*$  and  $\bar{\varphi}$ . In the particular case of a Pareto distribution, the two conditions give an expression for the productivity



cut-off  $\bar{\varphi}$  as a function of exogenous parameters:

$$\bar{\varphi} = b \left( \frac{f}{\delta f^e} \right)^{1/\kappa} \left( \frac{\kappa}{\rho - \kappa + 1} - 1 \right)^{1/\kappa} \quad (\text{B.4})$$

Equation (B.4) shows that, once adjustments in the mass of entering firms are accounted for, the productivity cut-off is independent from cost shocks. In the long run, firms relocate in response to price changes. When productivity is distributed according to a Pareto-distribution, the threshold productivity  $\bar{\varphi}$  remains unchanged. The adjustment along the extensive margin happens exclusively via firm relocation and adjustment in the mass of potential exporters. This differs from Chaney (2008) where the total mass of potential entrants is taken as given.

Finally, the mass  $M$  of firms is obtained using the definition of the sectoral price index

$$p = \left[ \frac{M}{1 - G(\bar{\varphi})} \int_{\bar{\varphi}}^{\infty} (p^*(\varphi)E)^{1-\rho} g(\varphi) d\varphi \right]^{\frac{1}{1-\rho}}$$

Using optimal pricing, this implies

$$p^{1-\rho} = M \left( \frac{\rho}{\rho - 1} \tau E w^* \right)^{1-\rho} \bar{\varphi}^{\rho-1}$$

Substituting the expression into the definition for average profits implies

$$\tilde{\pi}^* = \frac{1}{\rho} \frac{pc}{EM} - w^* f$$

Using the ZCP condition finally gives

$$M = \frac{\rho - \kappa + 1}{\kappa} \frac{1}{\rho} \frac{1}{w^* f} \frac{pc}{E}$$

We now derive how prices respond to the cost shock. Since we identify separately domestic and foreign entities, we re-introduce sector and country indices. Foreign price indices  $P_{ki}$ ,  $i \in I$ ,  $i \neq d$ , can now react to domestic cost shocks because market structure responds endogenously. For all  $i$  in  $I$ ,  $i \neq d$ , we have

$$\frac{\partial \ln P_{ki}}{\partial \ln w_d} = 0 + \frac{1}{1 - \rho_k} \left( \frac{p_{kif}(\bar{\varphi}_{ki})}{P_{ki}} \right)^{1-\rho_k} M_{ki} g(\bar{\varphi}_{ki}) \bar{\varphi}_{ki} \frac{\partial \ln \bar{\varphi}_{ki}}{\partial \ln w_d} + \frac{1}{1 - \rho_k} \frac{\partial \ln M_{ki}}{\partial \ln w_d} \quad (\text{B.5})$$

while the response of the domestic price index is slightly different,

$$\frac{\partial \ln P_{kd}}{\partial \ln w_d} = 1 + \frac{1}{1 - \rho_k} \left( \frac{p_{kdf}(\bar{\varphi}_{kd})}{P_{kd}} \right)^{1-\rho_k} M_{kd} g(\bar{\varphi}_{kd}) \bar{\varphi}_{kd} \frac{\partial \ln \bar{\varphi}_{kd}}{\partial \ln w_d} + \frac{1}{1 - \rho_k} \frac{\partial \ln M_{kd}}{\partial \ln w_d} \quad (\text{B.6})$$

because each domestic firm prices at a fixed markup above its marginal costs  $w_d$ . The first term in both equations measures the response of individual firm prices to the domestic shock,  $\frac{\partial \ln P_{kif}}{\partial \ln w_d}$ . This is zero for  $i$  in  $I$ ,  $i \neq d$  and 1 for  $i = d$ , since it reflects the pure intensive margin of adjustment. The second term captures the response of the productivity cut-off,  $\bar{\varphi}_{ki}$ , and the third term corresponds to

the adjustment in the potential number of firms serving the market. In the long run, and under a Pareto distribution, equation (B.4) implies  $\frac{\partial \ln \bar{\varphi}_{ki}}{\partial \ln w_d} = \frac{\partial \ln \bar{\varphi}_{kd}}{\partial \ln w_d} = 0$ .

Using equilibrium values for  $p_c$ , straightforward algebra implies

$$\frac{\partial \ln M_{ki}}{\partial \ln w_d} = (1 - \sigma_k) \frac{\partial \ln P_{ki}}{\partial \ln w_d} + (\sigma_k - \gamma) \frac{\partial \ln P_k}{\partial \ln w_d} + \gamma \frac{\partial \ln P}{\partial \ln w_d}$$

and

$$\frac{\partial \ln M_{kd}}{\partial \ln w_d} = (1 - \sigma_k) \frac{\partial \ln P_{kd}}{\partial \ln w_d} + (\sigma_k - \gamma) \frac{\partial \ln P_k}{\partial \ln w_d} + \gamma \frac{\partial \ln P}{\partial \ln w_d} - 1$$

where  $\frac{\partial \ln P_{ki}}{\partial \ln w_d}$  is now different from zero because of the response of market structure in all foreign economies. Substituting in equations (B.5) and (B.6), respectively, we obtain

$$\frac{\partial \ln P_{ki}}{\partial \ln w_d} = \frac{\sigma_k - \gamma}{\sigma_k - \rho_k} \frac{\partial \ln P_k}{\partial \ln w_d} + \frac{\gamma}{\sigma_k - \rho_k} \frac{\partial \ln P}{\partial \ln w_d} \quad (\text{B.7})$$

and

$$\frac{\partial \ln P_{kd}}{\partial \ln w_d} = -\frac{\rho_k}{\sigma_k - \rho_k} + \frac{\sigma_k - \gamma}{\sigma_k - \rho_k} \frac{\partial \ln P_k}{\partial \ln w_d} + \frac{\gamma}{\sigma_k - \rho_k} \frac{\partial \ln P}{\partial \ln w_d} \quad (\text{B.8})$$

Notice this implies

$$\frac{\partial \ln P_{ki}}{\partial \ln w_d} = \frac{\partial \ln P_{kd}}{\partial \ln w_d} + \frac{\rho_k}{\sigma_k - \rho_k}$$

By definition, and since  $P_{ki}$  can now change in all countries  $i$ , we have

$$\frac{\partial \ln P_k}{\partial \ln w_d} = w_k^M \frac{\partial \ln P_{ki}}{\partial \ln w_d} + (1 - w_k^M) \frac{\partial \ln P_{kd}}{\partial \ln w_d} \quad (\text{B.9})$$

Equations (B.7), (B.8) and (B.9) form a system that can be solved for the response of the price index  $P_k$  to a domestic cost shock, allowing for long run firm entry decisions in both the domestic and all foreign markets. This writes

$$\frac{\partial \ln P_k}{\partial \ln w_d} = \frac{\rho_k}{\rho_k - \gamma} (1 - w_k^M) - \frac{\gamma}{\rho_k - \gamma} \frac{\partial \ln P}{\partial \ln w_d}$$

which completes the derivation in the text.

## APPENDIX C. THE PRICE ELASTICITY OF IMPORTS

The price elasticity of imports is defined as

$$\eta_{NoFirm} = \frac{\partial \ln [\sum_k \sum_i P_{ki} C_{ki}]}{-\partial \ln w_d}$$

Using equilibrium traded values, this simplifies into

$$\eta_{NoFirm} = \sum_k \sum_i n_{ki} \left[ (\sigma_k - 1) \frac{\partial \ln P_{ki}}{\partial \ln w_d} + (\gamma - \sigma_k) \frac{\partial \ln P_k}{\partial \ln w_d} - \gamma \frac{\partial \ln P}{\partial \ln w_d} \right]$$

Using the definitions of all price indices, we obtain

$$\begin{aligned}
\eta_{NoFirm} &= \sum_k n_k (\gamma - \sigma_k) (1 - w_k^M) - \gamma \frac{\partial \ln P}{\partial \ln w_d} \\
&= \sum_k n_k (\gamma - \sigma_k) (1 - w_k^M) - \gamma \sum_k w_k (1 - w_k^M) \\
&= \gamma - \sigma_{NoFirm} + \sum_k n_{kd} (\sigma_k - \gamma) w_k^M - \sum_k w_k (1 - w_k^M)
\end{aligned}$$

which is the expression in the text.

#### APPENDIX D. DERIVATION OF THE SUPPLY CURVE

Omitting once again indexes for sector  $k$  and country  $i$ , the definition of the sector price index implies

$$\begin{aligned}
p^{1-\rho} &= M \left( \frac{\rho}{\rho-1} \tau w^* E \right)^{1-\rho} \frac{1}{1 - G(\bar{\varphi})} \int_{\bar{\varphi}}^{\infty} \varphi^{\rho-1} d\varphi \\
&= M \left( \frac{\rho}{\rho-1} \tau w^* E \right)^{1-\rho} \bar{\varphi}(\bar{\varphi})^{\rho-1}
\end{aligned}$$

Now in the long run we have

$$M = \left( \frac{\bar{\varphi}}{\bar{\varphi}(\bar{\varphi})} \right)^{\rho-1} \frac{1}{\rho} \frac{1}{w^* f} \frac{pc}{E}$$

Substituting  $M$  out gives

$$p = \bar{\varphi}^{\frac{1-\rho}{\rho}} (w^* \rho f E)^{\frac{1}{\rho}} c^{-\frac{1}{\rho}} \left( \frac{\rho}{\rho-1} \tau w^* E \right)^{1-\rho}$$

Now recognize that  $\bar{\varphi}$  is constant, and define  $\exp(v_{kit}) = (w^* \rho f E)^{\frac{1}{\rho}} \left( \frac{\rho}{\rho-1} w^* \right)^{1-\rho} (\tau E)^{-\rho}$  to obtain the expression in the text. Note that we have assumed constant returns to scale in our theory, but the effective estimation procedure allows for non-constant returns to scale, i.e for non negative values of  $\omega$ .

## APPENDIX E. ESTIMATED VARIANCES

The variance of  $\hat{\sigma}_k$  is computed using the second-order moments of  $\hat{\theta}_{1k}$  and  $\hat{\theta}_{2k}$  and a first-order approximation of  $\sigma_k$  around its true value:

$$\begin{aligned}\sigma_k &= \hat{\sigma}_k + \frac{\partial \sigma_k}{\partial \theta_{1k}} \Big|_{\theta_{1k}=\hat{\theta}_{1k}} (\theta_{1k} - \hat{\theta}_{1k}) + \frac{\partial \sigma_k}{\partial \theta_{2k}} \Big|_{\theta_{2k}=\hat{\theta}_{2k}} (\theta_{2k} - \hat{\theta}_{2k}) \\ \Rightarrow Var(\hat{\sigma}_k) &= \left( \frac{\partial \sigma_k}{\partial \theta_{1k}} \Big|_{\theta_{1k}=\hat{\theta}_{1k}} \right)^2 Var(\hat{\theta}_{1k}) + 2 \frac{\partial \sigma_k}{\partial \theta_{1k}} \Big|_{\theta_{1k}=\hat{\theta}_{1k}} \frac{\partial \sigma_k}{\partial \theta_{2k}} \Big|_{\theta_{2k}=\hat{\theta}_{2k}} Cov(\hat{\theta}_{1k}, \hat{\theta}_{2k}) \\ &\quad + \left( \frac{\partial \sigma_k}{\partial \theta_{2k}} \Big|_{\theta_{2k}=\hat{\theta}_{2k}} \right)^2 Var(\hat{\theta}_{2k})\end{aligned}$$

where:

$$\begin{aligned}\frac{\partial \sigma_k}{\partial \theta_{1k}} &= \frac{1}{\theta_{1k}} \left[ 1 - \sigma + / - \frac{1}{\sqrt{\theta_{2k}^2 + 4\theta_{1k}}} \right] \\ \frac{\partial \sigma_k}{\partial \theta_{2k}} &= \frac{1}{2\theta_{1k}} \left[ 1 + / - \frac{\theta_{2k}}{\sqrt{\theta_{2k}^2 + 4\theta_{1k}}} \right]\end{aligned}$$

Using the same reasoning, the first-order approximation of the aggregate elasticity around its estimated value gives:

$$\sigma_{NoFirm} = \hat{\sigma}_{NoFirm} - \sum_{k \in K} [n_k + (n_k - n_{kd})(1 - w_k^M)](\sigma_k - \hat{\sigma}_k)$$

The variance of  $\hat{\sigma}_{NoFirm}$  is then defined as:

$$\begin{aligned}Var(\hat{\sigma}_{NoFirm}) &\equiv E(\sigma_{NoFirm} - \hat{\sigma}_{NoFirm})^2 \\ \Rightarrow Var(\hat{\sigma}_{NoFirm}) &= \sum_{k \in K} [n_k + (n_k - n_{kd})(1 - w_k^M)]^2 Var(\hat{\sigma}_k) \\ &\quad + \sum_{k \in K} \sum_{k' \neq k} [n_{k'} + (n_{k'} - n_{k'd})(1 - w_{k'}^M)] [n_k + (n_k - n_{kd})(1 - w_k^M)] Cov(\hat{\sigma}_k, \hat{\sigma}_{k'})\end{aligned}$$

Since we control for common correlated effects in the estimation of the  $\sigma_k$ s,  $Cov(\hat{\sigma}_k, \hat{\sigma}_{k'})$  is effectively zero, and the estimated variance is given by

$$Var(\hat{\sigma}_{NoFirm}) = \sum_{k \in K} [n_k + (n_k - n_{kd})(1 - w_k^M)]^2 Var(\hat{\sigma}_k)$$

In the constrained case, the same reasoning gives

$$Var(\hat{\hat{\sigma}}_{NoFirm}) = \left[ \sum_{k \in K} [n_k + (n_k - n_{kd})(1 - w_k^M)] \right]^2 Var(\hat{\sigma})$$

The variance of the constrained and unconstrained import elasticities are given by:

$$\begin{aligned} Var(\hat{\eta}_{NoFirm}) &= \sum_{k \in K} n_k^2 (1 - w_k^M)^2 Var(\hat{\sigma}_k) \\ Var(\hat{\eta}_{NoFirm}) &= \left[ \sum_{k \in K} n_k (1 - w_k^M) \right]^2 Var(\hat{\sigma}) \end{aligned}$$

## APPENDIX F. A TWO-SECTOR VERSION OF BKK<sup>21</sup>

Two countries  $i = 1, 2$  face aggregate productivity shocks. Each country is inhabited by a large number of identical agents and labor is internationally immobile. Our main departure from BKK is that each country produces two goods,  $a$  and  $b$ . Preferences of the representative agent in country  $i$  are characterized by utility functions of the form

$$E_0 \sum_{t=0}^{\infty} \beta^t U(c_{it}, 1 - n_{it})$$

where  $U = \log c + \xi \frac{(1-n_t)^{1-\eta}}{1-\eta}$  and  $c_{it}$  ( $n_{it}$ ) denote aggregate consumption (hours worked). Aggregate consumption is a Cobb-Douglas function of sector-specific consumption

$$c_{i,t} = \frac{c_{i,t}^a{}^{\alpha_i} c_{i,t}^b{}^{1-\alpha_i}}{\alpha_i^{\alpha_i} (1 - \alpha_i)^{1-\alpha_i}}$$

where  $c_{i,t}^a$  and  $c_{i,t}^b$  are the consumption baskets of good  $a$  and  $b$ , and  $\alpha_i$  is the share of sector  $a$  in nominal aggregate consumption. The same structure prevails for aggregate investment:

$$i_{i,t} = \frac{i_{i,t}^a{}^{\alpha_i} i_{i,t}^b{}^{1-\alpha_i}}{\alpha_i^{\alpha_i} (1 - \alpha_i)^{1-\alpha_i}}$$

Sectoral output is produced with capital  $k$  and labor  $n$  following a Cobb-Douglas function:

$$y_{i,t}^k = z_{i,t} (k_{i,t}^k)^\theta (n_{i,t}^k)^{1-\theta}, \quad i = 1, 2, \quad k = a, b$$

The quantity  $y_{i,t}^k$  denotes country  $i$ 's production of good  $k$ , in units of the local good. In equilibrium, it is equal to domestic sales  $c_{i,t}^{ik} + i_{i,t}^{ik}$  plus exports  $c_{i,t}^{i'k} + i_{i,t}^{i'k}$ . The vector  $z_t = (z_{1,t}, z_{2,t})$  is a stochastic shock to productivity. Importantly, productivity shocks are assumed symmetric across sectors. The cross-sector symmetry assures that, in each country, producer prices are homogenous. In what follows, domestic prices are normalized to unity and the relative price of foreign goods is denoted  $P$ .

---

<sup>21</sup>We are grateful to Jean-Olivier Hairault who kindly shared his codes to the one-sector version of the BKK model.

Sectoral consumption and investment,  $c_{i,t}^k$  and  $i_{i,t}^k$  are composites of foreign and domestic goods:

$$\begin{aligned} c_{i,t}^k &= \left[ \left( \beta_i^k c_{i,t}^{i'k} \right)^{\frac{\sigma_k-1}{\sigma_k}} + \left( (1 - \beta_i^k) c_{i,t}^{ik} \right)^{\frac{\sigma_k-1}{\sigma_k}} \right]^{\frac{\sigma_k}{\sigma_k-1}} \\ i_{i,t}^k &= \left[ \left( \beta_i^k i_{i,t}^{i'k} \right)^{\frac{\sigma_k-1}{\sigma_k}} + \left( (1 - \beta_i^k) i_{i,t}^{ik} \right)^{\frac{\sigma_k-1}{\sigma_k}} \right]^{\frac{\sigma_k}{\sigma_k-1}} \end{aligned}$$

The elasticity of substitution between foreign and domestic varieties  $\sigma_k$  is sector-specific. The weights  $\beta_i^k$  are related to the share of imports in the sectoral consumption of good  $k$ . In the calibration, they are assumed symmetric across countries but potentially different across sectors.

The aggregate capital stock evolves in each country according to:

$$k_{i,t+1} = (1 - \delta)k_{i,t} + i_{i,t}$$

where  $\delta$  is the depreciation rate. Adjustment costs for capital follow Obstfeld and Rogoff (1996):

$$\mathcal{C}_{it} = \frac{\Phi}{2} \frac{(k_{i,t+1} - k_{i,t})^2}{k_{i,t}}$$

Finally, fluctuations arise from persistent shocks to aggregate productivity:

$$z_{t+1} = Az_t + \varepsilon_{t+1}^Z$$

where  $\varepsilon^Z$  is distributed normally and independently over time with variance  $V^Z$ . The correlation between the technology shocks,  $z_1$  and  $z_2$  is determined by the off-diagonal elements of  $A$  and  $V^Z$ .

We can obtain national income and product accounts for each country. Aggregate GDP in country 1 in period  $t$ , in units of domestically produced goods, is  $y_{1t} = y_{1t}^a + y_{1t}^b$ . The resource constraint equates sectoral GDPs to the sum of (domestic and foreign) consumption and investment:

$$y_{1t}^k = c_{1,t}^{1k} + c_{2,t}^{1k} + i_{1,t}^{1k} + i_{2,t}^{1k}, \quad k = a, b$$

National output is related to expenditure components according to:

$$y_{1t} = c_{1t}^{1a} + c_{1t}^{1b} + i_{1t}^{1a} + i_{1t}^{1b} + P_t(c_{1t}^{2a} + c_{1t}^{2b} + i_{1t}^{2a} + i_{1t}^{2b})$$

Finally, the trade balance, defined as the ratio of net exports to output, both measured in current prices, is:

$$nx_t = \frac{c_{2t}^{1a} + c_{2t}^{1b} + i_{2t}^{1a} + i_{2t}^{1b} - P_t(c_{1t}^{2a} + c_{1t}^{2b} + i_{1t}^{2a} + i_{1t}^{2b})}{y_{1t}}$$

and the terms of trade  $P_t$  equal the sectoral marginal rate of transformation between the two varieties in country 1, evaluated at equilibrium quantities.

Table F.1 summarizes our calibration. Our only deviation from the classical BKK model pertains to the elasticities of substitution at the sector level. In particular, the two new parameters in the multi-sector version of BKK pertain to the calibration of the Armington and Cobb-Douglas aggregators for consumption. To calibrate these, we use the sectoral data in our estimation of the

Table F.1. Benchmark Parameter Values taken from BKK (1994)

<b>Preferences</b>	
	$\beta = 0.99$
	$\eta = 5$
	$\sigma_a = 12.9$
	$\sigma_b = 4.8$
	$\alpha_a = 1 - \alpha_b = 0.24$
	$w_a^M = 0.25$
	$w_b^M = 0.22$
<b>Technology</b>	
	$\theta = 0.36$
	$\delta = 0.025$
	$\phi = 10^{-6}$
	$n_{i,SS} = 0.34$
	$n_{i,SS}^a = \alpha_i n_{i,SS}$
<b>Forcing processes</b>	
	$A = \begin{bmatrix} 0.906 & 0.088 \\ 0.088 & 0.906 \end{bmatrix}$
	$Var \varepsilon_1^Z = Var \varepsilon_2^Z = 0.00852^2$
	$Corr(\varepsilon_1^Z, \varepsilon_2^Z) = 0.258$

aggregate substitution elasticity. In terms of the model presented in Section 2,  $\alpha_i$ , the share of sector  $a$  in nominal consumption, is directly related to  $w_k$ . In a symmetric steady state with  $P = 1$ , the  $\beta_i^k$  parameters are linked with the  $w_k^M$  parameters according to

$$\beta_i^k = \left[ \left( \frac{w_k^M}{1 - w_k^M} \right)^{\frac{1}{1-\sigma_k}} + 1 \right]^{-1}$$

As before,  $n_{kd}$  can be expressed as:

$$n_{kd} = \frac{(1 - w_k^M) w_k}{\sum_k (1 - w_k^M) w_k}$$