Elasticity optimism

Jean Imbs, Isabelle Méjean

ECOLE POLYTECHNIQUE

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

ELASTICITY OPTIMISM

Jean IMBS
Isabelle MEJEAN

Février 2009

Cahier n° 2009-05

DEPARTEMENT D'ECONOMIE
Route de Saclay
91128 PALAISEAU CEDEX
(33) 1 69333033
http://www.enseignement.polytechnique.fr/economie/
mailto:chantal.poujouly@polytechnique.edu

ELASTICITY OPTIMISM[1]

Jean IMBS[2]
Isabelle MEJEAN[3]

Février 2009

Cahier n° 2009-05

Résumé: Les estimations de l'élasticité de substitution entre variétés domestiques et étrangères sont faibles dans les données macroéconomiques, mais nettement plus élevées dans les études sur données désagrégées. Cette divergence peut être la conséquence de l'hétérogénéité. Nous utilisons des données fines de commerce pour identifier structurellement les élasticités de substitution pour les biens américains. Nous construisons un modèle d'équilibre partiel pour agréger ces élasticités. Cet exercice nous permet de comparer les élasticités agrégées qui imposent une homogénéité inter-sectorielle et celles qui tiennent compte de l'hétérogénéité. Pour les premières, nous obtenons des valeurs similaires à ce que trouvent les études sur données macroéconomiques. Quand on tient compte de l'hétérogénéité, les élasticités sont plus de deux fois supérieures, jusqu'à 7. Ce paramètre est crucial dans la calibration de la plupart des modèles d'économie internationale. Nous discutons comment la correction que nous proposons affecte un certain nombre de résultats de la littérature, concernant notamment la dynamique des balances courantes, la transmission internationale des chocs, les choix internationaux de portefeuille, et la politique monétaire optimale.

Abstract: Estimates of the elasticity of substitution between domestic and foreign varieties are small in macroeconomic data, and substantially larger in disaggregated studies. This may be an artifact of heterogeneity. We use disaggregated multilateral trade data to structurally identify elasticities of substitution in US goods. We spell out a partial equilibrium model to aggregate them adequately at the country level. We compare aggregate elasticities that impose equality across sectors, to estimates allowing for heterogeneity. The former are similar in value to conventional macroeconomic estimates; but they are more than twice larger -up to 7- with heterogeneity. The parameter is central to calibrated models in most of international economics. We discuss the difference our corrected estimate makes in various areas of international economics, including the dynamics of external balances, the international transmission of shocks, international portfolio choice and optimal monetary policy.

Mots clés : Elasticités du Commerce, Agrégation, Calibration, Déséquilibres Globaux, Transmission Internationale, Politique Monétaire.

Key Words : Trade Elasticities, Aggregation, Calibration, Global Imbalances, International Transmission, International Portfolio, Monetary Policy.

Classification JEL: F41, F32, F21

---

[1] 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, and while Méjean was at CEPII. 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.

[2] HEC Lausanne, Swiss Finance Institute and CEPR.
[3]Department of Economics, Ecole Polytechnique.

1 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 diffusion of shocks, as in Kose and Yi (2006) or Corsetti, Dedola and Leduc (2008); or indeed to the extent of international risk sharing in Cole and Obstfeld (1991), and the composition of international portfolio holdings, as in Coeurdacier (2005). Models of international price differences have predictions that rest on the parameter value as well, as in Midrigan (forthcoming) or Atkeson and Burstein (2008). 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, finely disaggregated good-level quantities are more responsive to (international) relative prices than aggregates. Second, there are enormous differences between goods. Long time ago, Orcutt (1950) referred to an "elasticity pessimism", which he related to the gap between the low observed volatilities in aggregate quantities and the high volatility of international relative prices. He already conjectured that aggregates could obscure more responsive quantities at the microeconomic level. Here we ask whether this very heterogeneity may not actually be cause for optimism. We propose a correction of the elasticity of substitution between domestic and foreign macroeconomic quantities that accounts for heterogeneity at the microeconomic level.

Why should such a correction be meaningful? The robust finding that microeconomic studies uncover systematically higher estimates than aggregate data points to the possibility that aggregation matters. Suppose the price elasticity of imports varies by sector. An aggregate elasticity estimate implicitly imposes homogeneity and as a result may be biased for conventional econometric reasons. In other words, while it is the response of aggregate quantities that

macroeconomists are interested in, the fact they are undeniably much smoother than international prices may just be an artifact of aggregation. With heterogeneity, it does not follow that the aggregate elasticity of substitution is low. Orcutt (1950) reasoned that "goods with relatively low price elasticities can display the largest variation in prices and therefore exert a dominant effect on the estimated aggregate price elasticity, thereby biasing the estimate downwards." Aggregation can create a bias because the assumption that all quantities are equally substitutable and adjust homogeneously to changes in relative prices is not supported by the data. As a result there might be systematic differences between the responsiveness of aggregate traded quantities and the preference parameter it is meant to capture. It is of course the latter that should enter calibrated models. The bias we discuss matters for calibration purposes.

Our first goal is to develop a general model telling us how to properly aggregate microeconomic elasticities. We construct a measure of aggregate substitutability consistent with a representative agent choosing between aggregates of domestic and foreign quantities. At the same time, we accommodate the well documented fact that substitutability is vastly different across goods and sectors. We simulate the standard homogeneity assumption, under which elasticities are identical across sectors. This puts our theory to the test of its ability to reproduce conventional macroeconomic estimates under conventional macroeconomic assumptions. At the same time, the model maps out an index of substitutability between domestic and foreign aggregates, and a weighted average of its disaggregated counterparts. We let theory tell us which weights must be used for aggregation.

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 driven 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. There is a discrepancy between conventional macroeconomic estimates of the elasticity of substitution, imposing equal elasticities across goods, and aggregate estimates allowing for heterogeneity. The discrepancy persists whether firm entry decisions are permitted - or not. In fact, we are able to spell out the conditions under which firm dynamics act to reinforce an aggregation 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 an aggregator of domestic and foreign goods varieties. Under the Armington assumption we impose, 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 similaritiy in estimates across decades suggests there is nothing special to our data or our approach relative to existing empirical work. We seem to identify the object of interest to many international economists.

In macroeconomic applications, calibration exercises typically favor values of the parameter that are inferred from aggregate estimates of imports price elasticity. This is done for lack of a consensus, often because they are construed 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. This is within the ballpark of the calibrated values used in the macroeconomic literature, and within the range of estimates 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. This reflects the well known fact that microeconomic elasticities tend to be high on average, and also quite dispersed. It is also the result of an appropriate, theory-implied weighting of disaggregated estimates.

Does the correction matter economically? We discuss some illustrations in areas as diverse as the resolution of global imbalances, the dynamics of the trade balance, the international transmission of shocks, international risk sharing, portfolio choice, models of the real exchange rate and optimal monetary policy. In short, the parameter is central to most of international macroeconomics. Whether it is in fact 1 or 7 does make a quantitative difference, and calibrated equilibrium responses often change sizably. In some instances, the correction is also relevant qualitatively: the international response to productivity shocks, the justification for a home bias in portfolio investment or the Pareto-ranking of different macroeconomic policies can all alter profoundly with an elasticity substantially above one.

The impact of heterogeneity is significant and robust. Differences between standard and corrected 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.

2 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 larger on average but 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.

2.1 Practice

In a conventional framework, the substitutability between foreign and domestic goods maps directly with the price elasticity of imports. Estimating the latter, or trade elasticities in general, is an old business in economics. Early contributions used simple regression analysis to evaluate the response of aggregate imported quantities to changes in international relative prices. Houthakker and Magee (1969) for instance, estimate import elasticities between -0.5 and -1.5 depending on the economy considered.[1] The decades that followed saw a flurry of refinements, purporting to alleviate the econometric issues that go with having prices as independent variables. 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 based on macroeconomic data are found to be weakly negative, not always significant, and rarely larger than 2 in absolute value. The conventional inference in macroeconomics is to deduce that the elasticity of substitution between aggregate bundles of domestic and foreign goods is close to zero as well.

In contrast, the results obtained from disaggregated information are centered around higher averages, and substantial heterogeneity. For instance, Blonigen and Wilson (1999) document substitution elasticities between zero and 3.52 in 146 US sectors. Microeconomic studies also open the door to instrumentation strategies where changes in international relative prices can

[1] Most papers in this literature also take interest in the response of quantities traded to changes in income. We do not discuss this parameter here.

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).

In short, there is every indication that elasticities are higher on average at a disaggregated level. This comes as no surprise. More than fifty years ago, Orcutt (1950) already remarked that "it is widely recognized that the demand schedule for the product of an individual producer has, in general, far greater price elasticity than the aggregate demand schedule for the entire output of the product" (page 118). The importance of aggregation in the context of trade or substitution elasticities has a straightforward, economic intuition - one that seems to find support in the data. For instance, estimates in Hummels (2001) rise from 4.79 in one-digit SITC data to 8.26 for four-digit SITC data.

Ruhl (2005) proposes an interpretation. 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 are in fact fundamentally different, since in practice they do 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 estimators 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.

2.2 Theory

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

$$C = \prod_{k \in K} \frac{C_k^{\alpha_k}}{\alpha_k^{\alpha_k}}$$

where α_k denotes an exogenous preference parameter. 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 σ_k is specific to each industry, and assumed identical across all varieties, imported or not. β_{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 $i \neq d$ will effectively be 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. ρ_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 $\rho_k > \sigma_k$ 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.

The representative maximizing agent chooses her consumption allocation keeping in mind that all imported varieties incur a transport cost τ_{ki}, $i \in I$, $i \neq d$.[2] Utility maximization implies that demand for variety i in each sector k is given by

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

with

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

The demand addressed to each firm is given by:

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

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 σ_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,

[2]We could introduce without loss of generality 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.

we focus on changes in all relative prices, across all sectors k. This assumes away reallocation of demand across *industries*, with relative prices changing by the same amount in all sectors. 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.[3] A natural candidate is a domestic shock to relative production costs, driven for instance by a productivity disturbance that is not reflected in wages. 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 σ between bundles of domestic and foreign goods. By definition,

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

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.

[3]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.

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

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

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.

2.2.1 No Firm Entry

We clarify the separate influences of the issues raised by aggregation across all sectors k in K, and firm dynamics 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, $\frac{\partial \ln P_{kd}}{\partial \ln w_d} = 1$. Since we assume away any changes in market structure, foreign producers do not respond to changes in domestic costs, and $\frac{\partial \ln P_{ki}}{\partial \ln w_d} = 0$. As a result, the response of the sector price index is given by

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

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

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

The aggregate elasticity σ_{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 σ_{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.[4] We note the difference between partial and total elasticities is likely small. The second term in equation (4) is one order of magnitude smaller than the first. In fact, with similar sector allocation of expenditures for domestic and foreign goods, $n_k \simeq n_{kd}$, and partial and total elasticities are virtually identical.

Focus for now on partial elasticities 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, although disaggregated data are necessary to calculate it. 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 to the country level, and the econometrician asks how these aggregates respond to changes in international prices. In other words, she assumes $\sigma_k = \sigma$ for all k, which goes against most available empirical evidence. But equation (4) tells us the constraint may well be innocuous, if n_{kd} and σ_k happen to be uncorrelated. On the other hand, if substitutable varieties tend to form a large share of domestic expenditures, the unconstrained elasticity σ_{NoFirm} will take larger values than its constrained counterpart, $\bar{\sigma}_{NoFirm}$. The direction and magnitude of an aggregation bias depends therefore on the cross-sector correlation between expenditures shares n_{kd} and goods' substitutability σ_k.

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

[4]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.

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.

Estimates of the elasticity of substitution obtained from aggregate data provide potentially biased values of the true preference parameter that reflects the aggregate substitutability between domestic and foreign varieties. The bias arises because differences in good-specific substitutability are assumed away. This is true of partial and total elasticity, and we later report results for both concepts, correcting for heterogeneity or not. In what follows, we obtain structural estimates of σ_k (and its counterpart constrained to homogeneity), and infer the corresponding, theory-implied (partial or total) values for σ_{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 σ_{NoFirm} are not.

2.2.2 Allowing for Firm Entry

We now reintroduce firm heterogeneity in the definition of the aggregate elasticity of substitution, given 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 A, 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

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

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 $\frac{\partial \ln P_{ki}}{\partial \ln w_d}$ may 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 $\bar{\varphi}_{ki}$ depends only on exogenous parameters in a long run equilibrium where M_{ki} adjusts. The extensive margin works only 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:

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

The aggregate elasticity of substitution continues to be given by an adequately weighted average of disaggregated elasticities σ_k. An aggregation bias continues therefore to prevail in σ_{Firm} if n_{kd} and σ_k are correlated across sectors, exactly as it did when we ignored any extensive margin.

But there are two additional summation terms. The first one is positive under our assumptions and for $\sigma_k > 1$, both of which are empirically plausible. This positive correction is a direct consequence of the supply responses across all markets to a shift in relative costs. The estimate is higher because it now encapsulates the relative responses of quantities produced. This is reminiscent of Ruhl (2005) or Chaney (2008). The second correction can take either sign, depending on the relative magnitude of n_k and n_{kd}. But it is small in magnitude, especially in a symmetric setup where n_k is close to n_{kd}. We note this correction corresponds to the response in price indices at the sector level, and thus to the difference between a partial and a total elasticity.

Our main purpose in this section is to ascertain that supply responses across all markets do not alter the existence of an aggregation issue 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, σ_{Firm}, differs sizably from that between $\bar{\sigma}_{NoFirm}$ and σ_{NoFirm}. We noted in the previous section that the difference between $\bar{\sigma}_{NoFirm}$ and σ_{NoFirm} increases in the cross-sector correlation bewteen n_{kd} and σ_k. For positive values of this correlation, a model-implied weighted average of σ_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 σ_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, consider equation (6):

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

where σ and ρ are homogeneous elasticities of substitution, constrained to be identical across all sectors. Under mild conditions, the first summation increases in the correlation between n_{kd}

and σ_k.[5] The second summation, in turn, may take either sign, but it is one order of magnitude smaller, and corresponds to a bias in the relative responses in price indices. It will in fact be zero for symmetric models where the allocation of expenditures is identical for domestic and imported goods.

We later introduce an empirical approach to estimate σ_k (and its constrained counterpart σ). For lack of cross-country firm-level data, we cannot obtain any estimates of ρ_k or indeed ρ. We are therefore only able to calculate a theory-implied aggregation bias without firm entry, $\sigma_{NoFirm} - \bar{\sigma}_{NoFirm}$. In our data, n_{kd} and σ_k are positively correlated so that the bias will also take positive values. Under these conditions, we just showed that $\sigma_{Firm} - \bar{\sigma}_{Firm}$ is greater than $\sigma_{NoFirm} - \bar{\sigma}_{NoFirm}$. This suggests the aggregation bias we document - on the basis of our estimates of σ_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.

2.2.3 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).[6] In the model, this is equivalent to

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

Demand continues to be given by equation (1). In Appendix B, we derive the following

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

[6] 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.

expression for η_{NoFirm}:

$$\eta_{NoFirm} = 1 - \sigma_{NoFirm} + \sum_k n_{kd}(\sigma_k - 1)w_k^M - \sum_k \alpha_k \left(1 - w_k^M\right) \tag{7}$$

The elasticity contains three terms. First, (one minus) the elasticity of substitution σ_{NoFirm}. The second and third summations reflect the responses of industry specific and aggregate price indices, P_k and P. They correspond to the difference between partial and total elasticities. We note they are likely to be an order of magnitude smaller than σ_{NoFirm}, since $n_k, \alpha_k, w_k^M < 1$, but $\sigma_k > 1$.

The partial price elasticity of imports equals $1 - \sum_k n_k \, \sigma_k$. Imposing homogeneity in σ_k, it simply equals $1 - \bar{\sigma}$, where $\bar{\sigma}$ is the homogeneous elasticity of substitution between domestic and foreign varieties in *all* sectors. 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 σ_k, estimates of the (partial) price elasticity of imports that introduce heterogeneity will be larger in absolute value. In other words, the possibility that estimates based on macroeconomic data should be biased continue to prevail for the price elasticity of imports.

This helps us relate our results with the literature. First, the bias we discuss prevails in import elasticities as well since $\sigma_{NoFirm} > \bar{\sigma}_{NoFirm}$ implies that $\eta_{NoFirm} < \bar{\eta}_{NoFirm}$. This can explain why macroeconomic estimates of η are smaller in absolute value than those arising from disaggregated data. Second, using our structural estimates of σ_k in equation (7).gives us an estimate of η_{NoFirm}. We cannot calculate η_{Firm}, the price elasticity of imports accounting for firm entry, since we do not have estimates of ρ_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 (partial and total versions of) η_{NoFirm}, imposing homogeneity or not. We compare these results with conventional estimates of η at both the macro and micro levels obtained from the literature. Finding similar values will suggest the Armington assumption is well supported by the data.

3 Identification

We review how the methodology in Feenstra (1994a) is adapted to our purposes. We first discuss the econometrics involved in estimating σ_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.

3.1 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, and assume

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

where υ_{kit} denotes a technological shock that can take different values across sectors and exporters, τ_{kit} is a trading cost and ω_k is the inverse of the price elasticity of supply in sector k.[7] This assumes production decisions are taken on the basis of the price net of transport costs, and labeled in local currency. The potential aggregate effects of the nominal exchange rate are soaked up by the shock υ_{kit}. It will therefore be important to implement appropriate econometric tools to control for any potential common effects in our estimated system.[8] In

[7] We follow Feenstra (1994a) and assume all exporters have the same supply elasticity.

[8] υ_{kit} will also absorb any heterogeneity in the extent of the exchange rate pass-through.

Appendix C 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

$$\Delta \ln \tilde{s}_{kit} = (1 - \sigma_k)\Delta \ln \tilde{P}_{kit} + \Phi_{kt} + \varepsilon_{kit} \tag{8}$$

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 τ_{kit} and β_{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

$$\Delta \ln \tilde{P}_{kit} = \Psi_{kt} + \frac{\omega_k}{1 + \omega_k \sigma_k} \varepsilon_{kit} + \delta_{kit} \tag{9}$$

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. $\delta_{kit} \equiv \frac{1}{1+\omega_k\sigma_k} \Delta \upsilon_{kit}$ is an error term encapsulating movements in the exchange rate or aggregate technological developments in country i and sector k.

Under standard assumptions on taste shocks β_{kit} and technology shocks υ_{kit} , it is possible to identify the system formed by equations (8) and (9). 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.[9] We follow Feenstra (1994a) and summarize the information contained in the system with the following estimable regression

$$Y_{kit} = \theta_{1k} X_{1kit} + \theta_{2k} X_{2kit} + u_{kit} \tag{10}$$

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 (10) 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 (10) still suffers from an endogeneity issue. We follow Feenstra (1994a) and 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 (10). We also 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.[10]

The system summarized by equation (10) 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 υ_{kit} across all k.

[9] In the empirics, we choose a reference country that is present in the US market during the whole observed period.

[10] 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).

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

$$Y_{kit} = \theta_0 + \theta_{1k}\hat{X}_{1ki} + \theta_{2k}\hat{X}_{2ki} + \theta_{3k}X_{1it} + \theta_{4k}X_{2it} + u_{kit} \tag{11}$$

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}.

Armed with consistent (and sector-specific) estimates of θ_{1k} and θ_{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 D details how these are also used to infer standard deviations around these point estimates.

As is apparent, there are combinations of estimates in equation (11) 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 (11) over the intervals of admissible values of the supply and demand elasticities. We use this approach whenever direct estimates of θ_{1k} and θ_{2k} cannot be used to infer $\hat{\sigma}_k$. Whenever CCE are included, we hold constant the estimates of θ_0, θ_{3k} and θ_{4k} obtained from the direct instrumental variable regression, and search the combination of values for σ_k and ω_k that minimizes the sum of squared residuals in equation (11). The corresponding standard errors are obtained via bootstrapping of the procedure using 1,000 repetitions.

3.2 Macroeconomic 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 (11) into

$$Y_{kit} = \theta_0 + \theta_1 \hat{X}_{1ki} + \theta_2 \hat{X}_{2ki} + \theta_3 X_{1it} + \theta_4 X_{2it} + u_{kit} \tag{12}$$

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 θ_1 and θ_2. These are now constrained to equality across all sectors k. Identification continues to rest on the cross-section of exporters i, but equation (12) is now estimated on the pooled dataset formed by observations on all sectors. It is noteworthy that identification in equations (11) and (12) 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 σ_k, equation (12) suffers from a conventional aggregation bias, of an econometric nature, distinct from and non-competing with the argument developed in the previous sections. The bias arises because imposing the constraint that $\theta_{jk} = \theta_j$ for all k and $j = 1, 2, 3, 4$ in equation (11) will impose homogeneity in coefficients that theory implies are heterogeneous across sectors. As in Pesaran and Smith (1995), ignored heterogeneity enters the residuals in equation (12), and creates a potential bias in the estimates of θ_j, $j = 1, 2, 3, 4$. The direction of this bias is ambiguous, and depends on the correlation between the heterogeneity in the residuals of equation (12) and its regressors. A careful analysis of this possibility is beyond the scope of this paper. But we note that, 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 θ_1 and θ_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

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

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

3.3 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 (12) is effectively estimated on a panel of HS6 sectors, and our partition of HS6 sectors into 56 ISIC categories is entirely innocuous here. Homogeneity is imposed across all sectors, and two HS6 goods have to be equally substitutable whether or not they belong to the same ISIC sector. The partition is only relevant for the unconstrained estimation, and is performed for lack of detailed information on n_k, n_{kd}, α_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 4.4, 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}, α_k or w_k^M.

The approach adapted from Feenstra (1994a) requires relatively little information on traded flows. To estimate equation (11) 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.[11] Quantities are in tons. The empirical model described in section 3.1 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 73 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 α_k. In particular, we have

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

Calibration is therefore only needed for α_k and w_k^M. In the main body of the text, we consider

[11] 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.

the following data sources. The expenditure shares α_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 (Revison 3) nomenclature, as the 1997 ratio of imports over domestic gross output. Values for n_k and n_{kd} are calculated accordingly.[12]

In section 4.4, 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 directly compute n_k, rather than on the basis of a model-implied formula. In our second variant, we do so, and therefore 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 α_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.

4 Results and Relevance

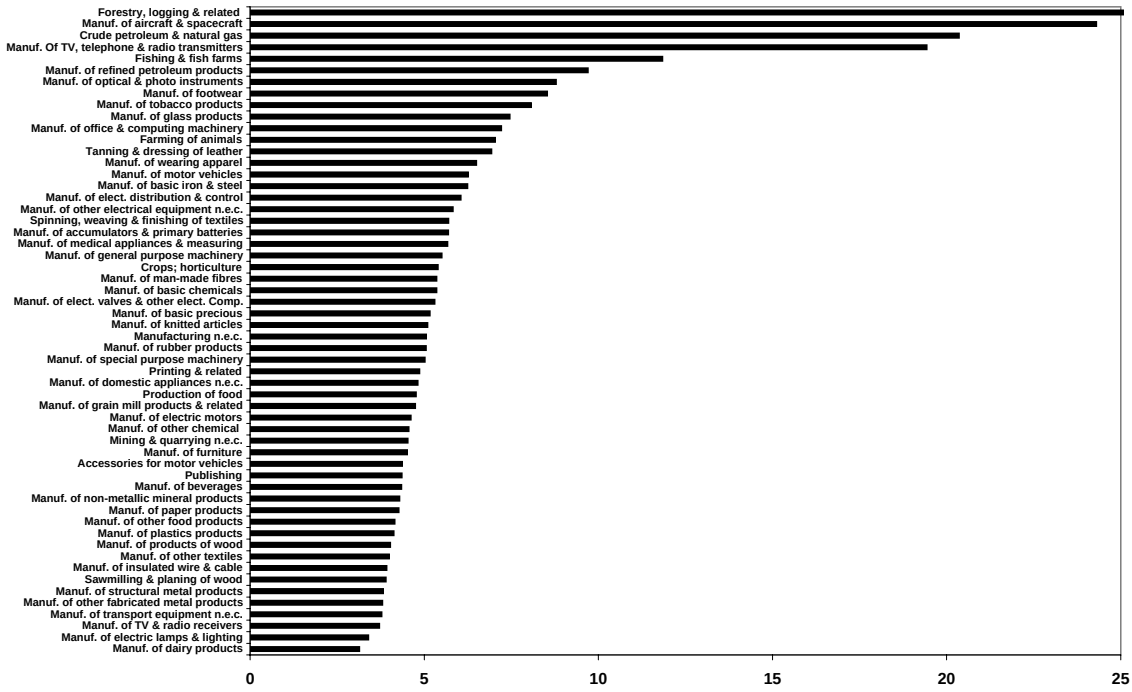
We first review the microeconomic estimates, obtained across 56 ISIC sectors, and relate them with existing evidence. We then aggregate them, first preserving heterogeneity. We compare the results with our estimation imposing homogeneity - using the same data sources to perform the aggregation. We discuss the discrepancy, and in particular whether its magnitude is significant *economically*. We argue the corrected estimates we uncover change dramatically the quantitative and qualitative predictions of a vast range of international macroeconomic models. We close with some robustness.

[12] α_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.

4.1 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 28 to 3.1 and a standard deviation of 4.9. The median value is only 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. In particular, they report cross-sector averages of microeconomic estimates, rather than estimates based on macroeconomic data.

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). The (partial) price elasticity of imports is then given by $1-\hat{\sigma}_k$. These elasticities were the object of a vast literature spread over the 1960s and 1970s, with two differences. First, in most instances, the results put forward focused on short-run elasticity estimates, 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. Second, the data used were more coarse, focusing 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. Our assumptions - and in particular the Armington aggregator - find comfort in this convergence of results.

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, sector-specific 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 elasticitites, 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 vindicates 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.

4.2 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 η_{NoFirm} and σ_{NoFirm} where $\sigma_k = \sigma$ as against those where σ_k is left unconstrained. Since we have

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

Table 1: Estimation with common correlated effects

	Import Elasticity η_{NoFirm}	Substitution Elasticity σ_{NoFirm}
Constrained total elasticity	-1.980^a (.175)	4.124^a (0.300)
Constrained partial elasticity	-2.738^a (.262)	3.738^a (0.263)
Unconstrained total elasticity	-4.508^a (.745)	7.226^a (0.962)
Unconstrained partial elasticity	-6.553^a (1.100)	6.921^a (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), ^{a} denotes significance at the 1% level.

the bias thus induced will increase in the correlations between σ_k and n_k. The second summation has second order effects only. In our data, the correlation is mildly positive. This suggests constraining all elasticities to be the same will act to lower estimates of both σ_{NoFirm} and η_{NoFirm}.

Table 1 reports estimates of both aggregate elasticities. We first report estimates of η_{NoFirm}, the total price elasticity of imports, when we impose that σ_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, an interval that is not significantly distinct from the one we estimate.

Once again, this vindicates the assumption of an Armington aggregator. 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. We refer to $1 - \eta_{NoFirm}$ as the natural proxy for aggregate substitutability, because macroeconomic calibrations would typically infer the elasticity of sub-

stitution from the price elasticity of imports, estimated *on the basis of macroeconomic data*. It is therefore the right comparison with the literature. 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). Note that in theory, it is only the partial version of σ_{NoFirm} that can be inferred as $1 - \eta_{NoFirm}$. The theoretical relation bewteen total elasticities is more complex, as illustrated in equation (7).

Constrained estimates stand in contrast with the value of η_{NoFirm} obtained when σ_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. This more than doubles the corresponding value for the estimated elasticity of substitution, at 5.5 - again, on the basis of a direct inference on σ_{NoFirm} as given by $1 - \eta_{NoFirm}$. The right panel of the Table reports the theoretical values for σ_{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.

Why the discrepancy? Does this not mean that aggregate quantities should be measurably more responsive to aggregate relative prices? Orcutt (1950) already speculated an explanation, reported by Goldstein and Khan (1985). On page 1070, they explain that, "in aggregate trade equations, goods with relatively low price elasticities can display the largest variation in prices and therefore exert a dominant effect on the estimated aggregate price elasticity, thereby biasing the estimate downwards." We view the results in Table 1 as a confirmation of this decade-old conjecture. With heterogeneity, the estimated response of aggregate quantities is not indicative of the true average elasticity of substitution.

It is useful to check the aggregation 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.

4.3 Relevance

Is the correction we document relevant in economic terms? We now discuss the quantitative and qualitative consequences of our corrected estimates. The most straightforward implication concerns models directly dealing with the responsiveness of traded quantities to relative prices, and in particular 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.[13] 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) showed 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. The result is meant as an illustrative special case of a powerful mechanism. Still, models of international portfolio holdings will have drastically different qualitative predictions depending on which side of one the parameter lies. For instance, Heathcote and Perri (2008) show that, with complete markets, complementarity between domestic and foreign goods can generate a home equity bias. 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: there will be portfolio home bias for values of the elasticity of substitution below one. A contrario, in Coeurdacier (2005), domestic consumers choose to hold foreign assets to insure against shocks to domestic consumption, provided the terms of trade respond

[13] We are grateful to Cedric Tille for graciously giving us the simulation code.

strongly enough in response to real shocks. This happens for values of the elasticity above one. 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. Rather, we stress the fact our estimated aggregate elasticity is substantially above one has far-reaching implications in models of international portfolio choice.

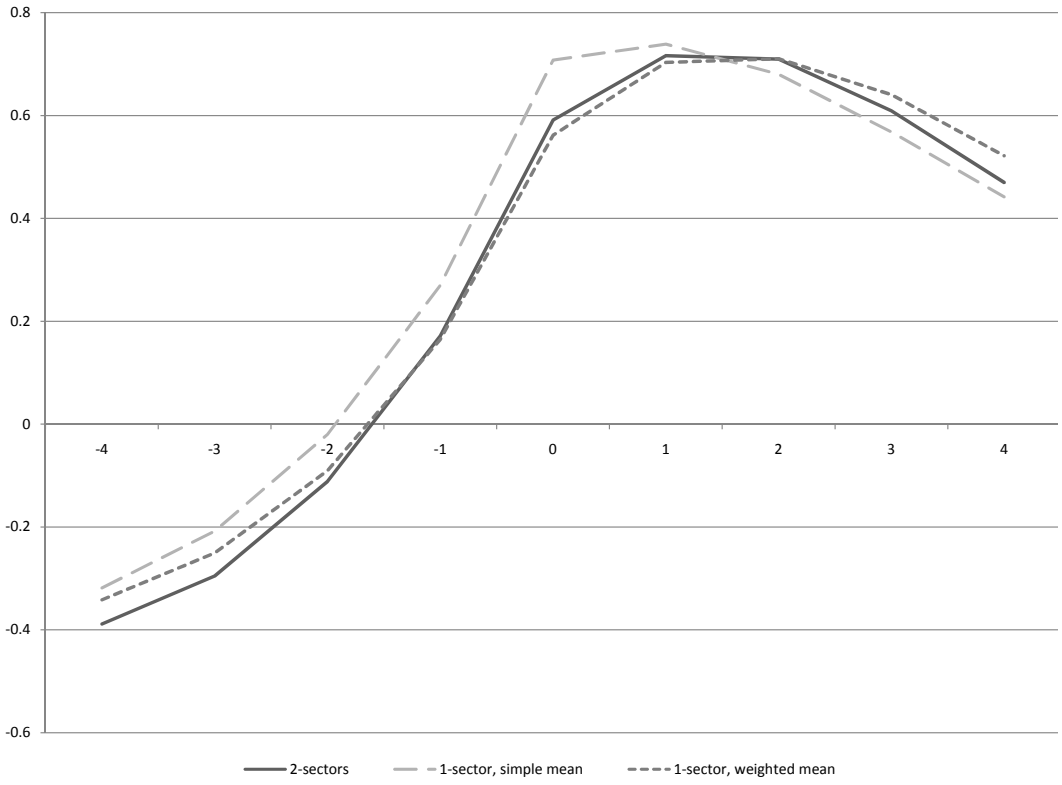
Atkeson and Burstein (2008) propose to explain observed deviations from PPP in a model with trade costs, imperfect competition and variable markups. Variables profit margins require that the elasticity of substitution between the (foreign and domestic) varieties that form a sector be larger than the substitutability across sectors. In their model, markups then depend on market shares, and can therefore fluctuate over time. In their calibration, the elasticity equals 10; deviations from PPP virtually disappear for an alternative value set at 3.

It is perhaps not surprising that models of the real exchange rate should have predictions that depend on the substitutability between domestic and foreign goods. A direct corollary is that the policy consequences of international price differences will also depend crucially on the parameter. Presumably, the relevance of the exchange rate in the monetary policy rule developed in Galí and Monacelli (2005) depends on the substitutability between foreign and domestic varieties. Galí and Monacelli focus on unitary elasticity, so the result is not directly apparent there. But Benigno and De Paoli (2006) introduce a generalization of their model, and their 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 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.

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 σ_{NoFirm} should be preferred to $\bar{\sigma}_{NoFirm}$ in calibrating

one-sector theoretical economies if one wants to capture the fact that σ_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 α_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 σ_k to come closest to the dynamics implied by the multi-sector version.

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



Since the workings of the model are well known, we leave a detailed description of the details to Appendix E. We compare the relative performances of three 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 σ_k. The second model uses the arithmetic average of σ_1 and σ_2 in a conventional one-sector version of BKK. This version is meant to capture a macroeconomic calibration fully ignorant of heterogeneity issues.[14] Finally, we calibrate a one-sector version of BKK using a weighted average of σ_1 and σ_2, $\sum_{k=1,2} n_{kd}\, \sigma_k$, consistent with the allowance for heterogeneity we have argued matters quantitatively. In all three models, we calibrate α_k and w_k^M using our data. Figure 2 reports the J-curves implied by the three 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.

4.4 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 α_k and w_k^M. Second, we investigate the importance of "Common Correlated Effects" in obtaining estimates of σ_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 σ_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 σ_{NoFirm} and η_{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 η_{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 σ_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,

[14] We note the arithmetic average of σ_1 and σ_2 is not necessarily an accurate 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, α_k using STAN sectoral interior demand. Variant 1: w_k^M using imports from BACI and output from STAN, α_k using STAN sectoral interior demand. Variant 2: n_k and w_k^M using imports and output from IO tables, α_k using STAN sectoral interior demand. Variant 3: n_k and w_k^M using imports from BACI and output from STAN, α_k using STAN sectoral interior demand. Variant 4: w_k^M using imports and output from IO tables, α_k using STAN sectoral interior demand (absorption in terms of value added).

Table 3: Estimation without common correlated effects

	Import Elasticity η_{NoFirm}	Substitution Elasticity σ_{NoFirm}
Constrained total elasticity	-2.166^a (.150)	4.442^a (.257)
Constrained partial elasticity	-3.016^a (.225.)	4.016^a (.225)
Unconstrained total elasticity	-4.075^a (.112)	6.584^a (0.145)
Unconstrained partial elasticity	-5.946^a (.209)	6.321^a (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), ^{a} denotes significance at the 1% level.

macroeconomic influences. When estimating the "true", sector-specific substitutability between domestic and foreign varieties, one want 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 η increases slightly, to -2.17, whereas the unconstrained estimate decreases 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 (11) 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 σ_k even within each ISIC 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 α_k and w_k^M to use: we do not observe any direct information on 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 α_k and w_k^M do not matter at the ISIC level. At least, this tests whether the heterogeneity in estimates of σ_k within each ISIC category can be such that our conclusions are altered.

Naturally, the constrained estimates of η_{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 σ_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.

5 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 disagreement. On average, microeconomic data tend to imply substantially higher values than macroeconomic aggregates. The point estimates are also much more heterogeneous. 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 an aggregation bias in elasticity estimates that goes back at least to Orcutt (1950). Such high parameter values change dramatically the conclusions of calibrated models in areas of international economics as varied as the international transmission of shocks, global imbalances, international risk sharing, portfolio choice and optimal monetary policy.

A 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 φ 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 τ, 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.[15]

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 φ, as firms with higher productivity can produce the same amount of output with fewer workers. Firms face a probability δ 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

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

[15] See for instance Segura-Cayuela and Villarubia (2008).

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 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.

$$p^*(\varphi) = \frac{\rho}{\rho - 1} \frac{\tau w^*}{\varphi} \tag{A.2}$$

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 $\tilde{\pi}^*$:

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

Since the threshold firm makes no profit, average profits simplify into

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

Equation (A.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:

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

Equation (A.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

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

where E is the bilateral exchange rate, expressed in terms of the importer's currency. Using

optimal pricing, this implies

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

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

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

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

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

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 (A.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 pc, straightforward algebra implies

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

and

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

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 (A.5) and (A.6), respectively, we obtain

$$\frac{\partial \ln P_{ki}}{\partial \ln w_d} = \frac{\sigma_k - 1}{\sigma_k - \rho_k}\frac{\partial \ln P_k}{\partial \ln w_d} + \frac{1}{\sigma_k - \rho_k}\frac{\partial \ln P}{\partial \ln w_d} \tag{A.7}$$

and

$$\frac{\partial \ln P_{kd}}{\partial \ln w_d} = -\frac{\rho_k}{\sigma_k - \rho_k} + \frac{\sigma_k - 1}{\sigma_k - \rho_k}\frac{\partial \ln P_k}{\partial \ln w_d} + \frac{1}{\sigma_k - \rho_k}\frac{\partial \ln P}{\partial \ln w_d} \tag{A.8}$$

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

$$\frac{\partial \ln P_k}{\partial \ln w_d} = w_k^M \frac{\partial \ln P_{ki}}{\partial \ln w_d} + \left(1 - w_k^M\right)\frac{\partial \ln P_{kd}}{\partial \ln w_d} \tag{A.9}$$

Equations (A.7), (A.8) and (A.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

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

which completes the derivation in the text.

B The Price Elasticity of Imports

The price elasticity of imports is defined as

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

Using equilibrium traded values, this simplifies into

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

Using the definitions of all price indices, we obtain

$$\begin{aligned} \eta_{NoFirm} &= \sum_k n_k (1 - \sigma_k)(1 - w_k^M) - \frac{\partial \ln P}{\partial \ln w_d} \\ &= 1 - \sum_k n_k \ \sigma_k + \sum_k n_k (\sigma_k - 1) w_k^M - \sum_k \alpha_k \left(1 - w_k^M\right) \end{aligned}$$

which is the expression in the text.

C Derivation of the Supply Curve

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

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

Now in the long run we have

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

Substituting M out gives

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

Now recognize that $\bar{\varphi}$ is constant, and define $\exp(\upsilon_{kit}) = (w^*\rho f E)^{\frac{1}{\rho}} \left(\frac{\rho}{\rho-1}w^*\right)^{1-\rho} (\tau E)^{-\rho}$ 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 ω.

D 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 σ_k around its true value:

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

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:

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

Since we control for common correlated effects in the estimation of the σ_ks, $Cov(\hat{\sigma}_k, \hat{\sigma}_{k'})$ 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

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

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

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

E A Two-Sector Version of BKK[16]

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

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

[16]We are grateful to Jean-Olivier Hairault who kindly gave us his codes to the one-sector version of the BKK model.

where $U = \log c + \gamma \frac{(1-n_t)^{1-\eta}}{1-\eta}$ 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 α_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:

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

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.

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

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

The elasticity of substitution between foreign and domestic varieties σ_k is sector-specific. The weights β_k^i 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 δ 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:

$$z_{t+1} = Az_t + \varepsilon^Z_{t+1}$$

where ε^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 $y_{1t} = y^a_{1t} + y^b_{1t}$. The resource contraint equates sectoral GDPs to the sum of (domestic and foreign) consumption and investment:

$$y^k_{1t} = c^{1k}_{1,t} + c^{1k}_{2,t} + i^{1k}_{1,t} + i^{1k}_{2,t}, \quad k = a, b$$

National output is related to expenditure components according to:

$$y_{1t} = c^{1a}_{1t} + c^{1b}_{1t} + i^{1a}_{1t} + i^{1b}_{1t} + P_t(c^{2a}_{1t} + c^{2b}_{1t} + i^{2a}_{1t} + i^{2b}_{1t})$$

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

$$nx_t = \frac{c^{1a}_{2t} + c^{1b}_{2t} + i^{1a}_{2t} + i^{1b}_{2t} - P_t(c^{2a}_{1t} + c^{2b}_{1t} + i^{2a}_{1t} + i^{2b}_{1t})}{y_{1t}}$$

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 E.1: Benchmark Parameter Values taken from BKK (1994)

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

Table E.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 aggregate substitution elasticity. In terms of the model presented in Section 2, α_i, the share of sector a in nominal consumption, is directly related to α_k. In a symmetric steady state with $P = 1$, the β_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:

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

References

Anderson, J. & Van Wincoop, E., 2004, "Trade Costs", *Journal of Economic Literature,* 84(1): Volume XLII, September..

Atkeson, T. & Burstein, A., 2008, "Pricing to Market, Trade Costs, and International Relative Prices", *American Economic Review,* 98(5): 1998-2031.

Backus, D., P. Kehoe & F. Kydland, 1994, "Dynamics of the Trade Balance and the Terms of Trade: The J-Curve?" *American Economic Review*, 84(1): 84-103.

Benigno, G. & B. De Paoli, 2006, "Optimal Monetary and Fiscal Policy for a Small Open Economy", mimeo.

Bernard A., J. Eaton, J.B. Jensen & S. Kortum, 2003, "Plants and Productivity in International Trade", *American Economic Review*, 93(4): 1268-1290.

Blonigen, B. & W. Wilson, 1999, "Explaining Armington: What Determines Substitutability Between Home and Foreign Goods?" , *Canadian Journal of Economics*, 32(1): 1-21.

Broda, C. & D. Weinstein, 2006, "Globalization and the Gains from Variety", *The Quarterly Journal of Economics*, 121(2): 541-585.

Chaney T., forthcoming, "Distorted Gravity: The Intensive and Extensive Margins of International Trade", *American Economic Review.*

Chari, V.V., P. Kehoe & E. McGrattan, 2002, "Can Sticky Price Models Generate Volatile and Persistent Real Exchange Rates?", *Review of Economic Studies*, 69(3): 533-563.

Coeurdacier, N., 2005, "Do trade costs in goods market lead to home bias in equities?", mimeo.

Coeurdacier N., R. Kollmann & P. Martin, 2007, "International Portfolios with Supply, Demand and Redistributive Shocks", *NBER Working Papers*, 13424.

Cole, H. & M. Obstfeld, 1991, "Commodity trade and international risk sharing : How much do financial markets matter?", *Journal of Monetary Economics*, 28(1): 3-24.

Corsetti, G., L. Dedola & S. Leduc, forthcoming, "International Risk Sharing and the transmission of productivity shock", *Review of Economic Studies*.

Feenstra, R., 1994a, "New Product Varieties and the Measurement of International Prices", *American Economic Review*, 84(1): 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. & Monacelli, T., 2005, "Monetary Policy and Exchange Rate Volatility in a Small Open Economy" , *Review of Economic Studies*, 72(3): 707-734.

Goldstein M. & M. Kahn, 1985, "Income and price effect in foreign trade", In: J. Ronald and P. Kennen, Editors, Handbook of International Economics, North-Holland, Amsterdam, 1042–1099.

Hanson, G., 2005, "Market Potential, Increasing Returns and Geographic Concentration", *Journal of International Economics*, 67(1): 1-24.

Harberger, A., 1957, "Some Evidence on the International Price Mechanism", *Journal of Political Economy*, 65(6): 506-521.

Head, K. & J. Ries, 2001, "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade", *American Economic Review*, 91(4): 858-876.

Heathcote, J. & F. Perri, 2008, " The International Diversification Puzzle Is Not As Bad As You Think", mimeo.

Hooper, P., K. Johnson & J. Marquez, 1998, "Trade elasticities for G-7 countries", *International Finance Discussion Papers*, 609.

Houthakker, H. & S. Magee, 1969, "Income and Price Elasticities in World Trade", *The Review of Economics and Statistics*, 51(2): 111-25.

Hummels, D., 2001, "Towards a Geography of Trade Costs", *Mimeo*, Purdue University.

Kahn, M., 1975, "The Structure and Behavior of Imports of Venezuela", *The Review of Economics and Statistics*, 57(2): 221-24.

Kemp, M., 1962, "Error of Measurements and Bias in Estimates of Import Demand Parameters", *Economic Record*, 38: 369-372.

Kose, M. & Yi, K-M., 2006, "Can the standard international business cycle model explain the relation between trade and comovement?", *Journal of International Economics*, 68(2): 267-295.

Krugman, P., 1980, "Scale Economies, Product Differentiation, and the Pattern of Trade", *American Economic Review*, 70(5): 950-59.

Marquez, J., 1990, "Bilateral Trade Elasticities", *The Review of Economics and Statistics*, 72: 70-77.

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

Obstfeld, M. & K. Rogoff, 2005, "Global Current Account Imbalances and Exchange Rate Adjustments", *Brookings Papers on Economics Activity*, 2005(1): 67-146.

Orcutt, G., 1950, "Measurement of Price Elasticities in International Trade", *The Review of Economics and Statistics*, 32(2): 117-132.

Pesaran, H., 2006, "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure", *Econometrica*, 74(4): 967-1012.

Pesaran, H. and R. Smith, 1995, "Estimating Long Run Relationships from Dynamic Heterogeneous Panels", *Journal of Econometrics*, LXVIII: 79-113.

Romalis, J., 2007, "NAFTA's and CUSFTA's Impact on International Trade", *The Review of Economics and Statistics*, 89(3): 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*, 73(2): 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*, 61(2): 306-312.