TRADE THEORY WITH NUMBERS:
QUANTIFYING THE CONSEQUENCES OF GLOBALIZATION

Arnaud Costinot
Andrés Rodríguez-Clare

Trade Theory with Numbers: Quantifying the Consequences of Globalization
Arnaud Costinot and Andrés Rodríguez-Clare
NBER Working Paper No. 18896
March 2013."Tgxkugf"Lwpg"4235
JEL No. F11,F12,F13,F15,F17,F60,F62

ABSTRACT

We review a recent body of theoretical work that aims to put numbers on the consequences of globalization. A unifying theme of our survey is methodological. We rely on gravity models and demonstrate how they can be used for counterfactual analysis. We highlight how various economic considerations—market"structure, firm-level heterogeneity, multiple sectors, intermediate goods, and multiple factors of production—affect"the magnitude of the gains from trade liberalization. We conclude by discussing a number of outstanding"issues in the literature as well as alternative approaches for quantifying the consequences of globalization.

Arnaud Costinot
Department of Economics
MIT, E52-243B
50 Memorial Drive
Cambridge MA 02142-1347
and NBER
costinot@mit.edu

Andrés Rodríguez-Clare
University of California at Berkeley
Department of Economics
Berkeley, CA 94720-3880
and NBER
andres1000@gmail.com

1   Introduction

The theoretical proposition that there are gains from international trade, see Samuelson (1939), is one of the most fundamental result in all of economics. Under perfect competition, opening up to trade acts as an expansion of the production possibility frontier and leads to Pareto superior outcomes. The objective of this chapter is to survey a recent body of theoretical work that aims to put numbers on this and other related comparative static exercises, which we will refer to as globalization.

A unifying theme of our chapter is methodological. Throughout we rely on multi-country gravity models and demonstrate how they can be used for counterfactual analysis. While so-called gravity equations have been estimated since the early sixties, see Tinbergen (1962), the widespread use of structural gravity models in the field of international trade is a fairly recent phenomenon, as also discussed by Head and Mayer (2013) in this volume. The previous handbook of international economics is a case in point. In his opening chapter, Krugman (1995) notes: "the lack of a good analysis of multilateral trade in the presence of trade costs is a major gap in trade theory." This view is echoed by Leamer and Levinsohn (1995) who argue that: "The gravity models are strictly descriptive. They lack a theoretical underpinning so that once the facts are out, it is not clear what to make of them." But the times they are a-changin'.

The last ten years have seen an explosion of alternative microtheoretical foundations underlying gravity equations; see Eaton and Kortum (2002), Anderson and Van Wincoop (2003), Bernard, Eaton, Jensen, and Kortum (2003), Chaney (2008), and Eaton, Kortum, and Kramarz (2011). While new gravity models encompass a large number of market structures—from perfect competition to monopolistic competition with firm-level heterogeneity à la Melitz (2003)—and a wide range of micro-level predictions, they share the same macro-level predictions regarding the structure of bilateral trade flows as a function of bilateral costs. It is this basic macro structure and its quantitative implications for the consequences of globalization that we will be interested in this chapter.

Recent quantitative trade models based on the gravity equation share the same primary focus as older Computational General Equilibrium (CGE) models; see Baldwin and Venables (1995) for an overview in the previous handbook. The main goal is to use theory in order to derive numbers—e.g., explore whether particular economic forces appear to be large or small in the data—rather than pure qualitative insights—e.g., study whether the relationship between two economic variables is monotone or not in theory. There are, however, important differences between old and new quantitative work in international trade that we will try to highlight throughout this chapter. First, new quantitative trade models have more appealing micro-theoretical foundations. One does not need to impose the somewhat ad-hoc assumption that each country is exogenously endowed with a distinct good—the so-called "Armington" assumption—to do quantitative work in international trade. Second, recent quantitative papers offer a tighter connection between theory and data. Instead of relying on off-the-shelf elasticities, today's researchers try to use their own model to estimate the key structural parameters necessary for counterfactual analysis. Estimation and computation go hand in hand. Third, new quantitative trade models put more emphasis on

transparency and less emphasis on realism. The idea is to construct middle-sized models that are rich enough to speak to first-order features of the data, like the role of country size and geography, yet parsimonious enough so that one can credibly identify its key parameters and understand how their magnitude affects counterfactual analysis.

Section 2 starts by studying the simplest gravity model possible, the Armington model. Building on Arkolakis, Costinot, and Rodríguez-Clare (2012), we highlight two basic results. First we show that the changes in welfare associated with globalization, modelled as a change in iceberg trade costs, can be inferred using two variables: (i) changes in the share of expenditure on domestic goods; and (ii) the elasticity of bilateral imports with respect to variable trade costs, which we refer to as the trade elasticity. Second we show how changes in bilateral trade flows, in general, and the share of domestic expenditure, in particular, can be computed using only information about the trade elasticity and easily accessible macroeconomic data. We refer to this approach popularized by Dekle, Eaton, and Kortum (2008) as "exact hat algebra."

Armed with these tools, we illustrate how gravity models can be used to quantify the gains from international trade defined as the (absolute value of) the percentage change in real income that would be associated with moving one country from the current, observed trade equilibrium to a counterfactual equilibrium with no trade, i.e. an equilibrium with infinite iceberg trade costs. Since the share of domestic expenditure on domestic goods under autarky is equal to one, the welfare consequences associated with this counterfactual exercise are easy to compute. Although this is obviously an extreme counterfactual scenario that is (hopefully) not seriously considered by policymakers, we view it as a useful benchmark that can shed light on the quantitative importance of the various channels through which globalization affects the welfare of nations.

Section 3 extends the simple Armington model along several directions. First, we relax the assumption that each country is exogenously endowed with a distinct good and provide alternative assumptions on technology and market structure under which the counterfactual predictions derived in Section 2 remain unchanged. Second, we introduce multiple sectors, intermediate goods, and multiple factors of production and discuss how these considerations affect the consequences of globalization. Third, we briefly discuss other extensions including alternative demand systems—that generate variable markups under monopolistic competition—and multinational production. Although one can still use macro-level data and a small number of elasticities to compute the gains from trade in these richer environments, the results of Section 3 illustrate that some realistic departures from the one-sector benchmark, such as the existence of multiple sectors and tradable intermediate goods tend to increase significantly the magnitude of the gains from trade.

Section 4 focuses on evaluating trade policy. Instead of considering the welfare consequences of a move to autarky, we study counterfactual scenarios in which countries raise their import tariffs, either unilaterally or simultaneously around the world, using the simple Armington model. We then study again how these counterfactual predictions vary across different gravity models. We conclude by discussing how to measure the restrictiveness of trade policy when tariffs are

heterogeneous across sectors.

Section 5 reviews a number of outstanding issues in the literature. Since the main output of quantitative trade models are numbers, a fair question is: Are these numbers that we can believe in? To shed light on this question, we first discuss the sensitivity of the predictions of gravity models to auxiliary assumptions on the nature of trade imbalances and the tradability of capital goods. We then turn to the goodness of fit of gravity models in the cross-section and time series. We conclude by discussing how elasticities, i.e., the main inputs of quantitative trade models, are calibrated.

Sections 6 and 7 discuss other approaches to quantifying the consequences of globalization in the literature. Section 6 focuses on recent empirical studies that have used micro-level data, either at the product or firm-level, to estimate gains from new varieties and productivity gains from trade. We discuss how such empirical evidence, i.e., "micro" numbers, relate to the predictions of gravity models reviewed in this chapter, i.e., "macro" numbers. Section 7 turns to structural approaches to quantifying the consequences of globalization that are not based on gravity models. Due to space constraints, we do not review reduced-form evidence on the gains from openness; see e.g. Frankel and Romer (1999), Feyrer (2009a), and Feyrer (2009b). Readers interested in this important topic are referred to the recent survey by Harrison and Rodríguez-Clare (2010).

Section 8 offers some concluding remarks on the current state of the literature and open questions for future research. Additional information about theoretical results and data can be found in the online Appendix.

2   Getting Started

We start this chapter by describing how to perform counterfactual analysis in the simplest quantitative trade model possible: the Armington model. A central aspect of this model is the gravity equation; see e.g. Anderson (1979) and Anderson and Van Wincoop (2003). As we will see in the next section, there exists a variety of microtheoretical foundations that can give rise to a gravity equation, and in turn, a variety of economic environments in which the simple tools introduced in this section can be applied.

2.1   Armington Model

Consider a world economy comprising $i = 1, ..., n$ countries, each endowed with Q_i units of a distinct good $i = 1, ..., n$.

Preferences.   Each country is populated by a representative agent whose preferences are represented by a Constant Elasticity of Substitution (CES) utility function:

$$C_j = \left(\sum_{i=1}^{n} \psi_{ij}^{(1-\sigma)/\sigma} C_{ij}^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)}, \tag{1}$$

where C_{ij} is the demand for good i in country j; $\psi_{ij} > 0$ is an exogenous preference parameter;

and $\sigma > 1$ is the elasticity of substitution between goods from different countries. The associated consumer price index is given by

$$P_j = \left(\sum_{i=1}^{n} \psi_{ij}^{1-\sigma} P_{ij}^{1-\sigma} \right)^{1/(1-\sigma)}, \tag{2}$$

where P_{ij} is the price of good i in country j.

Trade Costs. International trade between countries is subject to iceberg trade costs. In order to sell one unit of a good in country j, firms from country i must ship $\tau_{ij} \geq 1$ units, with $\tau_{ii} = 1$. For there to be no arbitrage opportunities, the price of good i in country j must be equal to $P_{ij} = \tau_{ij} P_{ii}$. The domestic price P_{ii} of good i, in turn, can be expressed as a function of country i's total income, Y_i, and its endowment: $P_{ii} = Y_i/Q_i$. Combining the two previous expressions we get

$$P_{ij} = Y_i \tau_{ij} / Q_i. \tag{3}$$

Trade Flows. Let X_{ij} denote the total value of country j's imports from country i. Given CES utility, bilateral trade flows satisfy

$$X_{ij} = \left(\frac{\psi_{ij} P_{ij}}{P_j} \right)^{1-\sigma} E_j, \tag{4}$$

where $E_j \equiv \sum_{i=1}^{n} X_{ij}$ is country j's total expenditure. Combining Equations (2)-(4), we obtain

$$X_{ij} = \frac{(Y_i \tau_{ij})^{1-\sigma} \chi_{ij}}{\sum_{l=1}^{n} (Y_l \tau_{lj})^{1-\sigma} \chi_{lj}} E_j,$$

where $\chi_{ij} \equiv (Q_i/\psi_{ij})^{\sigma-1}$. In order to prepare the general analysis of Section 3, we let $\varepsilon \equiv \partial \ln (X_{ij}/X_{jj}) / \partial \ln \tau_{ij}$ denote the the elasticity of imports relative to domestic demand, X_{ij}/X_{jj}, with respect to bilateral trade costs, τ_{ij}, holding income levels fixed. In the rest of this chapter we will refer to ε as the *trade elasticity*. In the Armington model it is simply equal to $\sigma - 1$. Using the previous notation, we can rearrange the expression above as

$$X_{ij} = \frac{(Y_i \tau_{ij})^{-\varepsilon} \chi_{ij}}{\sum_{l=1}^{n} (Y_l \tau_{lj})^{-\varepsilon} \chi_{lj}} E_j. \tag{5}$$

Equation (5) is what we will refer to as the *gravity equation*.

Competitive Equilibrium. In a competitive equilibrium, budget constraint and good market clearing imply $Y_i = E_i$ and $Y_i = \sum_{j=1}^{n} X_{ij}$, respectively, for all countries i. Together with Equation (5), these two conditions imply

$$Y_i = \sum_{j=1}^{n} \frac{(Y_i \tau_{ij})^{-\varepsilon} \chi_{ij}}{\sum_{l=1}^{n} (Y_l \tau_{lj})^{-\varepsilon} \chi_{lj}} Y_j. \tag{6}$$

4

This provides a system of n equations with n unknowns, $\mathbf{Y} \equiv \{Y_i\}$. By Walras' Law, one of these equations is redundant. Thus income levels are only determined up to a constant. Once income levels are known, expenditure levels, $\mathbf{E} \equiv \{E_i\}$, can be computed using budget constraint and bilateral trade flows, $\mathbf{X} \equiv \{X_{ij}\}$, can be computed using the gravity equation. This concludes the description of the Armington model.

2.2 Counterfactual Analysis

We now illustrate how the gravity equation can be used to quantify the welfare consequences of globalization. For simplicity, we focus on a shock to trade costs from $\boldsymbol{\tau} \equiv \{\tau_{ij}\}$ to $\boldsymbol{\tau}' \equiv \{\tau_{ij}'\}$. The same analysis generalizes in a straightforward manner to preference and endowment shocks. To quantify the welfare consequences of a trade shock in a given country j, we proceed in two steps. First, we show how changes in real consumption, $C_j \equiv E_j/P_j$, can be inferred from changes in macro variables, X and Y. Second, we show how to compute changes in macro variables.

Welfare. In this chapter, whenever we refer to welfare changes in country j, we refer to percentage changes in real consumption. Such changes correspond to the equivalent variation associated with a foreign shock (expressed as a share of expenditure before the shock). Namely, percentage changes in real consumption measures the percentage change in income that the representative agent would be willing to accept in lieu of the shock to happen.

The first result that we establish is that changes in real consumption can be inferred using only two statistics: (i) observed changes in the share of expenditure on domestic goods, $\lambda_{jj} \equiv X_{jj}/E_j$; and (ii) the trade elasticity in the gravity equation, ε.

Let us start by considering an infinitesimal change in trade costs from $\boldsymbol{\tau}$ to $\boldsymbol{\tau} + d\boldsymbol{\tau}$. By Shephard's Lemma, we know that

$$d \ln P_j = \sum_{i=1}^{n} \lambda_{ij} d \ln P_{ij},$$

where $\lambda_{ij} \equiv X_{ij}/E_j$ denotes the share of expenditure on goods from country i in country j. Since consumption is chosen to minimize expenditure, changes in consumption levels, C_{ij}, only have second-order effects on the consumer price index in country j. Under the assumption of CES utility, changes in the consumer price index in country j can be rearranged further into changes into domestic and import prices

$$d \ln P_j = \lambda_{jj} d \ln P_{jj} + \left(1 - \lambda_{jj}\right) d \ln P_j^M,$$

where $P_j^M \equiv \left[\sum_{i \neq j} P_{ij}^{1-\sigma} \right]^{1/(1-\sigma)}$ is the component of the price index associated with imports. By differentiating Equation (4), one can also show that

$$d \ln \left(1 - \lambda_{jj}\right) - d \ln \lambda_{jj} = (1 - \sigma) \left(d \ln P_j^M - d \ln P_{jj}\right).$$

Combining the fact that $\lambda_{jj} d \ln \lambda_{jj} = - \left(1 - \lambda_{jj} \right) d \ln \left(1 - \lambda_{jj} \right)$ with the two previous, we get

$$d \ln P_j = d \ln P_{jj} - \left(d \ln \lambda_{jj} / \left(1 - \sigma \right) \right). \tag{7}$$

Thus, changes in real consumption, $C_j \equiv E_j / P_j$, in country j are given by

$$d \ln C_j = \left(d \ln E_j - d \ln P_{jj} \right) + \left(d \ln \lambda_{jj} / \left(1 - \sigma \right) \right). \tag{8}$$

Since there are no domestic trade costs, $\tau_{jj} = \tau'_{jj} = 1$, and trade is balanced, $Y_j = E_j$, Equation (3) implies that the first term is equal to zero. In the simple Armington model, changes in real consumption only depends on the change in the relative price of imported versus domestic goods, P_j^M / P_{jj}, which depends on the share of expenditure on domestic goods, λ_{jj}, and the elasticity of substitution, σ. Using Equations (8) and the definition of the trade elasticity $\varepsilon \equiv \sigma - 1$, we get

$$d \ln C_j = -d \ln \lambda_{jj} / \varepsilon.$$

Since the previous expression holds for any infinitesimal shock, the welfare consequences of large changes from τ to τ' can be inferred by integrating the previous formula:

$$\hat{C}_j = \hat{\lambda}_{jj}^{-1/\varepsilon}, \tag{9}$$

where $\hat{v} \equiv v'/v$ denotes the proportional change in any variable v between the initial and counterfactual equilibria. This establishes that for any change in trade costs, two statistics—the trade elasticity, ε, and the changes in the share of expenditure on domestic goods, λ_{jj}—are sufficient to infer welfare changes.[1]

Macroeconomic variables. At this point, we have shown that conditional on the trade elasticity, ε, changes in real consumption are exclusively determined by changes in λ_{jj}. We now describe how one can use gravity models to predict how trade shocks affect trade flows, in general, and the share of expenditure on domestic goods, λ_{jj}, in particular. The approach that we will describe has been popularized recently by Dekle, Eaton, and Kortum (2008). One can think of this approach as an "exact" version of Jones' hat algebra for reasons that will be clear in a moment.[2]

Let $\lambda_{ij} \equiv X_{ij} / \sum_l X_{lj}$ denote the share of expenditure on goods from country i in country j. Since the gravity equation holds both in the initial and the counterfactual equilibria, we have

$$\hat{\lambda}_{ij} = \frac{\left(\hat{Y}_i \hat{\tau}_{ij} \right)^{-\varepsilon}}{\sum_{l=1}^{n} \lambda_{lj} \left(\hat{Y}_l \hat{\tau}_{lj} \right)^{-\varepsilon}}. \tag{10}$$

---

[1] Since the previous result is based on Shepard's Lemma and the fact that domestic prices are unaffected by the shock, $d \ln P_{jj} = 0$, Equation (9) would also hold in the case of a shock to foreign preferences or endowments.

[2] Application of this approach can also be found in the older CGE literature; see e.g. Rutherford (2002).

In the counterfactual equilibrium, Equation (6) further implies

$$Y_j' = \sum_{i=1}^n \lambda_{ji}' Y_i'.$$

Combining the two previous expressions, we then get

$$\hat{Y}_j Y_j = \sum_{i=1}^n \frac{\lambda_{ji} \left(\hat{Y}_j \hat{\tau}_{ji} \right)^{-\varepsilon} \hat{Y}_i Y_i}{\sum_{l=1}^n \lambda_{li} \left(\hat{Y}_l \hat{\tau}_{li} \right)^{-\varepsilon}}. \tag{11}$$

Although trade costs, endowments, and preference shifters affect bilateral trade flows, as captured by τ_{ij} and χ_{ij} in equation (5), Equation (11) shows that we can compute counterfactual changes in income, $\hat{Y} \equiv \{\hat{Y}_i\}$, as the solution of a system of non-linear equations without having to estimate any of these parameters. All we need to determine changes in income levels (up to normalization) are the initial expenditure shares, λ_{ij}, the initial income levels, Y_i, and the trade elasticity, ε. Given changes in income levels, changes in the shares of expenditure on goods from different countries, $\hat{\lambda}_{ij}$, and changes in real consumption, $\hat{C}_j$, can then be computed using Equations (9) and (10).

2.3 Trade Theory with Numbers: A Preview

In order to illustrate the usefulness of the simple Armington model, we focus on a very particular, but important counterfactual exercise: *moving to autarky*. Formally, we assume that variable trade costs in the new equilibrium are such that $\tau_{ij}' = +\infty$ for any pair of countries $i \neq j$. All other structural parameters are the same as in the initial equilibrium. For this particular shock, we do not need to solve any non-linear system of equations to do counterfactual analysis. Since the share of expenditure on domestic goods must be equal to 1 in the counterfactual equilibrium, $\lambda_{jj}' = 1$, we immediately know that $\hat{\lambda}_{jj} = 1/\lambda_{jj}$.

Throughout this chapter we define the *gains from international trade* in country j, G_j, as the absolute value of the percentage change in real income that would be associated with moving to autarky in country j. Using Equation (9) and the fact that $\hat{\lambda}_{jj} = 1/\lambda_{jj}$, we get

$$G_j = 1 - \lambda_{jj}^{1/\varepsilon}. \tag{12}$$

In order to compute G_j we need measures of the trade elasticity, ε, and the share of expenditure on domestic goods, λ_{jj}. There are many econometric issues associated with estimating ε; see e.g. Hummels and Hillberry (2012). A simple way to estimate the trade elasticity ε is to take the log of the gravity equation (5) and run a cross-sectional regression of the following form

$$\ln X_{ij} = \delta_i^X + \delta_j^M - \varepsilon \ln \tau_{ij} + \delta_{ij}, \tag{13}$$

where the first term $\delta_i^X \equiv \ln \chi_i - \varepsilon \ln Y_i$ is treated as an exporter fixed-effect; the second term $\delta_j^M \equiv \ln Y_j - \ln \left[\sum_{l=1}^n \chi_l \left(Y_l \tau_{lj} \right)^{-\varepsilon} \right]$ is treated as an importer fixed-effect; and the third term δ_{ij} is

treated as measurement error in trade flows that is orthogonal to $\ln \tau_{ij}$. At this point we set $\varepsilon = 5$, which is a typical value used in the literature; see e.g. Anderson and Van Wincoop (2004) and Head and Mayer (2013). We will come back to the sensitivity of our quantitative results to values of the trade elasticity in Section 5.

In order to measure λ_{jj} in the data, recall that $\lambda_{jj} \equiv X_{jj}/E_j = 1 - \sum_{i \neq j} X_{ij} / \sum_{i=1}^{n} X_{ij}$. We can measure $\sum_{i \neq j} X_{ij}$ as total imports by country j, whereas $\sum_i X_{ij}$ is total expenditure by country j. In this exercise as well as all subsequent exercises, we use data from the World Input-Output Database (WIOD) in 2008. The database covers 27 EU countries and 13 other major countries; see Timmer (2012).[3] The first column of Table 1 reports the gains from trade G_j for these countries using Equation (12). According to the simple Armington model, we see that gains from trade are below 2% for three countries: Brazil (1.5%), Japan (1.7%), and the United States (1.8%). Not surprisingly, gains from trade tend to be larger for smaller countries. The largest predicted gains are for Slovakia (7.6%), Ireland (8.0%) and Hungary (8.1%). Given the strong assumptions that have been imposed in Section 2.1, these numbers, of course, should be taken with more than a grain of salt. We now discuss how richer and more realistic models would affect the magnitude of the gains from trade.

3 Beyond Armington

The Armington model is very tractable, which has made it the go-to trade model for quantitative work in policy institutions for more than forty years. This is also a very stylized model, which has lead to quite a bit of skepticism about the robustness of its counterfactual predictions in academic circles for about as many years. Fortunately, one can maintain the tractability of the Armington model, without maintaining the somewhat ad-hoc assumption that each country is exogenously endowed with a distinct good. As discussed below, the gravity equation (5), which is the basis for counterfactual analysis in the Armington model, can be shown to hold under various assumptions about technology and market structure. While each gravity model remains special, in the sense that strong functional form assumptions are required for a gravity equation to hold, the ability of these new models to match a large number of micro-level facts, together with the elegance of their microtheoretical foundations, has lead to an explosion of quantitative work in international trade over the last ten years.

In this section we explore how various features of more complex gravity models—market structure, firm-level heterogeneity, multiple sectors, intermediate goods, and multiple factors of production—affect the gains from trade as defined in Section 2.3. Throughout our analysis we

---

[3]The mapping between the simple Armington model presented here and the data is not trivial for two reasons: (*i*) it assumes the share of expenditures on intermediate goods is zero and (*ii*) it assumes that trade is balanced. This implies that GDP is equal to gross output and that total expenditure is equal to GDP. Neither is true in the data. We will deal with intermediate goods and trade imbalances explicitly in Sections 3.4 and 5.1, respectively. Here, as well as in Section 3, we derive and apply our formulas for gains from trade ignoring trade imbalances. If moving to autarky also implies the closing of trade imbalances, our formulas capture the change in real income rather than the change in real expenditure. See online Appendix for details.

calibrate different models to match the same moments in the macro-data, including bilateral trade flows and trade elasticities. Thus different models may lead to different predictions about the magnitude of the gains from trade because they predict different counterfactual autarky equilibria, not because they predict different trade volumes in the initial equilibrium. In short, trade volumes are taken as data that discipline the behavior of all models, irrespectively of what their particular micro-theoretical foundations may be.[4]

As explained in the Introduction, although a move to autarky is an extreme comparative statics exercise, it should be viewed as a useful benchmark to study the importance, in a well-defined welfare sense, of various economic channels discussed in the literature. We leave the evaluation of trade policy to Section 4 in which we show how to use the exact hat algebra to conduct richer comparative static exercises.

3.1   Many Models, One Equation

The gravity equation (5) has been shown to hold under perfect competition, as in Eaton and Kortum (2002); under Bertrand competition, as in Bernard, Eaton, Jensen, and Kortum (2003); under monopolistic competition with homogeneous firms, as in Krugman (1980); and under monopolistic competition with firm-level heterogeneity, as in Chaney (2008), Arkolakis (2010), Arkolakis, Demidova, Klenow, and Rodríguez-Clare (2008), and Eaton, Kortum, and Kramarz (2011). Our goal in this subsection is not to describe each of these models in detail, but rather highlight the common features that will lead to a gravity equation as well as the key differences that may affect the magnitude of the gains from trade. Detailed discussions of the microfoundations and functional form assumptions leading to gravity equations can be found in Anderson (2010), Arkolakis, Costinot, and Rodríguez-Clare (2012), Head and Mayer (2013), as well as in our online Appendix.

Like the simple Armington model, the alternative gravity models mentioned above assume the existence of a representative agent with CES utility in each country as well as balanced trade, $E_i = Y_i$.[5] The representative agent, however, now has preferences over a continuum of goods or varieties $\omega \in \Omega$:

$$C_j = \left(\int_{\omega \in \Omega} c_j(\omega)^{(\sigma-1)/\sigma} d\omega \right)^{\sigma/(\sigma-1)},$$

with $\sigma > 1$. In equilibrium, each good ω is only imported from one country so that Equation (1) still holds with the aggregate consumption of goods from country i in country j being given by $C_{ij} = \int_{\omega \in \Omega_{ij}} c_j(\omega)^{(\sigma-1)/\sigma} d\omega$, where $\Omega_{ij} \subset \Omega$ denotes the set of goods that country j buys from country i, and $\psi_{ij} = 1$ for all country pairs due to the symmetry across varieties. For the same reason, Equation (2) holds as well with the aggregate price of goods from country i in country j being given by $P_{ij} = \left(\int_{\omega \in \Omega_{ij}} p_j(\omega)^{1-\sigma} d\omega \right)^{1/(1-\sigma)}$ and $\psi_{ij} = 1$ for all i and j.

---

[4]In order to match the same cross-section of trade flows, different gravity models may implictly rely on different values of bilateral trade costs as well as other structural parameters.

[5]In recent work, Arkolakis, Costinot, Donaldson, and Rodriguez-Clare (2012) have developed a gravity model without CES utility. We discuss the implications of such a model in Section 3.6.

A key difference between the simple Armington model and the alternative gravity models mentioned above is that, due to different assumptions on technology and market structure, Ω_{ij} is no longer exogenously given. In these richer models, firms from country i may now decide to stop producing and selling a subset of goods in country j if it is not profitable for them to do so. Hence changes in prices, P_{ij}, may reflect both: (i) changes at the intensive margin, i.e., changes in the price of goods imported in country j, $p_j(\omega)$, and (ii) changes at the extensive margin, i.e., changes in the set of good imported in country j, Ω_{ij}, due either to the selection of a different subset of firms from i in j or the entry of a different set of firms in i. Mathematically, these new economic considerations lead to the following generalization of Equation (3):

$$P_{ij} = \underbrace{\tau_{ij} c_i^p}_{\text{Intensive Margin}} \times \underbrace{\left(\left(\frac{E_j}{c_{ij}^x} \right)^{\frac{\delta}{1-\sigma}} \frac{\tau_{ij} c_i^p}{P_j} \right)^{\eta}}_{\text{Extensive Margin: Selection}} \times \underbrace{\left(\frac{R_i}{c_i^e} \right)^{\frac{\delta}{1-\sigma}}}_{\text{Extensive Margin: Entry}} \times \xi_{ij}, \tag{14}$$

where c_i^p, c_i^e, and c_{ij}^x are endogenous variables that capture how input prices affect variable costs of production, fixed entry costs, and fixed exporting costs, respectively; $E_j \equiv \sum_{i=1}^n X_{ij}$ still denotes total expenditure in country j; $R_i \equiv \sum_j X_{ij}$ denotes total sales or revenues for producers; and $\xi_{ij} > 0$ is a function of structural parameters distinct from variable trade costs, τ_{ij}, such as endowments or fixed exporting costs. The last two parameters, δ and η, will play a central role in our analysis. The parameter δ is a dummy variable that characterizes the market structure: it is equal to one under monopolistic competition with free entry and zero under perfect or Bertrand competition.[6] The parameter $\eta \geq 0$ is related to the extent of heterogeneity across varieties as we discuss more formally below.

In the rest of this chapter we will use Equation (14) to organize the literature and explain how different assumptions about technology and market structure—namely, different assumptions about c_i^p, c_i^e, c_{ij}^x, δ, and η—may lead to different macro-level predictions without getting lost into the algebra through which Equation (14) comes about. At this point, it is therefore important to clarify how each term in Equation (14) relates to previous work in the literature.

The first term, $\tau_{ij} c_i^p$, captures price changes at the intensive margin. This is the only active margin in the Armington model. In that model, $c_i^p = Y_i$, $\delta = 0$, and $\eta = 0$; see Equation (3). This intensive margin will remain active in all models that we study. In most of these models, c_i^p will also remain equal to total income Y_i in country i. This is the case, for instance, if labor is the only factor of production. In this situation, production costs are proportional to wages, which are proportional to countries' total income.

The second term, $\left(\left(E_j/c_{ij}^x \right)^{\frac{\delta}{1-\sigma}} \tau_{ij} c_i^p / P_j \right)^{\eta}$, captures changes at the extensive margin due to selection effects. If $\eta = 0$, then this term is equal to one and there are no selection effects. This occurs in models of monopolistic competition without firm-level heterogeneity and fixed export-

---

[6]For expositional purposes, we often abuse terminology in this chapter and simply refer to the case $\delta = 1$ and $\delta = 0$ as "monopolistic competition" and "perfect competition," respectively.

ing costs, like Krugman (1980), in which all firms always export. The case $\eta > 0$ captures instead situations in which a subset of firms from country i may start or stop exporting when market conditions change in country j. Specifically, in models of perfect competition, like Eaton and Kortum (2002), Bertrand competition, like Bernard, Eaton, Jensen, and Kortum (2003), or monopolistic competition à la Melitz (2003), like Chaney (2008), Arkolakis, Demidova, Klenow, and Rodríguez-Clare (2008), Arkolakis (2010), and Eaton, Kortum, and Kramarz (2011), if firms from country i are less competitive relative to other firms serving market j, i.e., $\tau_{ij} c_i^p / P_j$ is high, then less firms from country i will serve this market, which will lead to a decrease in the number of varieties from i available in j and an increase in P_{ij}. In the previous models, the magnitude of selection effects is formally determined by $\eta \equiv \left(\frac{\theta}{\sigma-1} \right) \left(1 + \frac{1-\sigma}{\theta} \right)$, where $\theta > \sigma - 1$ is the shape parameter of the distribution of productivity draws across varieties. Under perfect and Bertrand competition, this distribution is Fréchet. Under monopolistic competition, it is Pareto. In both cases, θ measures the elasticity of the mass of goods produced domestically with respect to their relative cost; $1/(1-\sigma)$ measures the elasticity of the price index with respect to new goods; and $\left(1 + \frac{1-\sigma}{\theta} \right)$ corrects for the fact that the marginal variety has a higher price than the average variety. In models of monopolistic competition with firm-level heterogeneity à la Melitz (2003) ($\delta = 1, \eta > 0$), selection also depends on the size of market j relative to the fixed costs of exporting from i to j, which is reflected in $\left(E_j / c_{ij}^x \right)^{\frac{\delta}{1-\sigma}}$. The nature of c_{ij}^x depends on where fixed exporting costs are paid. If they are paid in the exporting country, then c_{ij}^x is proportional to total income, Y_i, in country i. If they are paid in the importing country, then c_{ij}^x is proportional to total income, Y_j, in country j.

The third term, $\left(R_i / c_i^e \right)^{\frac{\delta}{1-\sigma}}$, captures changes at the extensive margin due to entry effects. This last channel is specific to models with monopolistic competition and free entry ($\delta = 1$), whether or not they feature firm-level heterogeneity à la Melitz (2003).[7] In such environments, countries in which entry is more profitable, i.e., R_i / c_i^e is high, export more varieties to all countries, which decreases the price index with an elasticity $1/(\sigma - 1)$. If entry costs are paid in terms of labor, as in Krugman (1980) or Melitz (2003), c_i^e is simply proportional to total income Y_i in country i.

To sum up, starting from Equation (14), we can turn off and on the selection effects associated with heterogeneity across varieties by setting η to 0 or not. Similarly, we can turn off and on the scale effects associated with monopolistic competition and free entry by setting δ to 0 or 1. In the next subsections, we study how much these considerations—as well as the introduction of multiple sectors, tradable intermediate goods, and multiple factors of production—affect the overall magnitude of the gains from trade.

---

[7]Without free entry, $\left(R_i / c_i^e \right)^{\frac{\delta}{1-\sigma}}$ would be absent from Equation (14) and one would need, in general, to take into account the effect of trade on profits; see e.g. Ossa (2011a). In the one-sector case reviewed in the next subsection, this distinction turns out to be irrelevant, as shown in Arkolakis, Costinot, and Rodríguez-Clare (2012). With multiple sectors, however, free entry leads to home market effects, with implications for the magnitude of the gains from trade.

3.2 One Sector

It is standard to interpret models with CES utility, such as those presented in Section 3.1, as one-sector models with a continuum of varieties; see e.g. Helpman and Krugman (1985). Here we focus on such models since they are the closest to the Armington model described in Section 2.1. Later we will consider multi-sector extensions of these models as well as incorporate tradable intermediate goods and multiple factors of production.

In line with the existing literature, we assume that, in addition to Equation (14), the two following conditions hold: (*i*) $c_i^p = c_{ii}^m = c_i^e = Y_i$, which reflects the fact that all factors of production are used in the same proportions to produce, export, and develop all varieties; and (*ii*) $R_i = Y_i$, which reflects the fact that trade in goods is balanced.[8] Under these two conditions, Equation (14) simplifies into

$$P_{ij} = \tau_{ij} Y_i \left(\left(\frac{E_j}{c_{ij}^x} \right)^{\frac{\delta}{1-\sigma}} \frac{\tau_{ij} Y_i}{P_j} \right)^{\eta} \xi_{ij}, \tag{15}$$

Note that $R_i = Y_i = c_i^e$ implies that, like in the Armington model, there are no entry effects associated with changes in trade costs, even under monopolistic competition. Since Equation (1) still holds, bilateral trade flows between country *i* and country *j* are still given by Equation (4). Combining this observation with Equation (15), the gravity equation generalizes to

$$X_{ij} = \frac{\left(Y_i \tau_{ij} \right)^{-\varepsilon} \left(c_{ij}^x \right)^{-\delta\eta} \chi_{ij}}{\sum_{l=1}^{n} \left(Y_l \tau_{lj} \right)^{-\varepsilon} \left(c_{lj}^x \right)^{-\delta\eta} \chi_{lj}} E_j. \tag{16}$$

Compared to the Armington model, here $\varepsilon = (1 + \eta)(\sigma - 1)$ and $\chi_{ij} \equiv \xi_{ij}^{1-\sigma}$. Thus if there are selection effects, i.e., if $\eta \neq 0$, the structural interpretation of the trade elasticity is no longer the same as in the Armington model. This reflects the fact that changes in variable trade costs now affect both the price of existing varieties (intensive margin) and the set of varieties sold from country *i* to country *j* (extensive margin). Nevertheless, we can still take the logs of both sides of Equation (4) and estimate the trade elasticity ε as we did in Section 2.3. In other words, the mapping between bilateral trade data, **X**, and the trade elasticity, ε, remains unchanged.[9] Finally, we see that changes in the magnitude of fixed exporting costs, as captured by $\left(c_{ij}^x \right)^{-\delta\eta}$, now affect bilateral trade flows under monopolistic competition with firm-level heterogeneity à la Melitz (2003), i.e., if $\delta = 1$, $\eta > 0$.[10] This extra term—which depends on whether fixed exporting costs

---

[8]Since some of the models that we consider involve fixed exporting costs paid in the importing country, i.e. trade in exporting services, the assumption that overall trade is balanced, $E_i = Y_i$, is different from the assumption that trade in goods is balanced, $R_i = Y_i$. The latter condition corresponds to the macro-level restriction R1 in Arkolakis, Costinot, and Rodríguez-Clare (2012).

[9]This assumes that measures of trade costs are invariant across models. This is a reasonable assumption in the case of import tariffs, but not in the case of price gaps; see Simonovska and Waugh (2012). We come back to the specific issues associated with import tariffs in Section 4.

[10]Equation (16) is a special case of the macro-level restriction R3 in Arkolakis, Costinot, and Rodríguez-Clare (2012). Compared to Arkolakis, Costinot, and Rodríguez-Clare (2012), we no longer need to impose explictly that profits are

are paid in the importing country, the exporting country, or both—opens up the possibility of different predictions across gravity models, as we explain below.

Like in Section 2.2, consider a change in variable trade costs from $\tau \equiv \{\tau_{ij}\}$ to $\tau' \equiv \{\tau'_{ij}\}$, though the same analysis easily extends to shocks to foreign endowments or technology, i.e., changes in $\chi \equiv \{\chi_{ij}\}$. Given CES utility, Equation (8) still holds. The key difference compared to our previous analysis is that $d \ln E_j - d \ln P_{jj}$ is no longer equal to zero. Because of selection effects, a change in variable trade costs may lead to a change in the set of goods produced domestically, Ω_{jj}, and so to a change in the aggregate price index associated with these goods, P_{jj}, relative to total expenditure in country j. Specifically, since $E_j = Y_j$ and $c_{jj}^m = Y_j$, Equation (15) implies

$$d \ln P_{jj} = (1 + \eta) \, d \ln Y_j - \eta d \ln P_j.$$

Using the previous expression with Equation (7), which still holds because of CES utility, and the trade balance condition, $E_j = Y_j$, we then get

$$d \ln E_j - d \ln P_{jj} = - \left(\eta / \left(\eta + 1 \right) \right) \times \left(d \ln \lambda_{jj} / \left(1 - \sigma \right) \right). \tag{17}$$

In models featuring selection effects ($\eta > 0$), a positive terms-of-trade shock, $d \ln \lambda_{jj} / (1 - \sigma) > 0$, is accompanied by a negative shock to real consumption of domestic goods, $d \ln E_j - d \ln P_{jj} < 0$. Intuitively, a positive terms-of-trade shock tends to decrease the profitability of domestic firms on the domestic market, which leads to a decrease in the number of goods produced domestically, and, given the love of varieties embedded in CES utility functions, an increase in the aggregate price index associated with these goods.

Together with $\varepsilon = (1 + \eta) (\sigma - 1)$, Equations (8) and (17) imply $d \ln C_j = -d \ln \lambda_{jj} / \varepsilon$. Like in the Armington model, the previous expression can be integrated to get Equation (9).[11] Thus, as established in Arkolakis, Costinot, and Rodríguez-Clare (2012), changes in the share of expenditure on domestic goods, λ_{jj}, and the trade elasticity, ε, remain two sufficient statistics for welfare analysis. In particular, gains from trade remain given by Equation (12). In short, conditional on observed trade flows and the trade elasticity, selection and scale effects have no impact on the overall magnitude of the gains from trade.

The fact that λ_{jj} remains the only macro variable that matters for welfare is intuitive enough: it measures the magnitude of the terms-of-trade effect and changes in real output themselves are function of changes in the overall price index, which also depends on the magnitude of the terms-of-trade effect. The fact that the only structural parameter that matters can be recovered as the trade elasticity ε from a gravity equation, by contrast, heavily relies on the fact that selection effects, as captured by η, are identical across countries. It is this assumption that simultaneously generates a gravity equation and guarantees that the trade elasticity in that equation is the relevant

---

proportional to revenues. This restriction is implicit in Equation (15).

[11]A more direct, though perhaps less illuminating way of establishing Equation (9) consists in computing P_{jj} from Equation (15) and substituting for it in $\lambda_{jj} = \left(P_{jj} / P_j \right)^{1 - \sigma}$.

elasticity for welfare analysis. It is worth noting that this result does not rely on the fact that new varieties have zero welfare effects. In the models with monopolistic competition considered here, a shock to trade costs, foreign endowments, or foreign technology, may very well increase welfare through its effects at the extensive margin. The point rather is that these effects, no matter how large they are, can always be inferred from changes in aggregate trade flows.

The previous equivalence only establishes that conditional on a change in λ_{jj} and ε, alternative gravity models must predict the same welfare change as the Armington model, but, in principle, they may predict different changes in the share of domestic expenditure for a given trade shock. Under the additional assumption that fixed exporting costs (if any) are paid in the destination country—$c_{ij}^x = c_j^m$, as in Eaton, Kortum, and Kramarz (2011)—one can show a stronger equivalence result. In that case, c_{ij}^x drops out of Equation (16). Thus the counterfactual changes in trade flows and income levels associated with changes in variable trade costs can still be computed using the exact hat algebra of Section 2.2, i.e. Equations (10) and (11). Given observable macro variables in the initial equilibrium, X and Y, and a value of the trade elasticity, ε, this implies that counterfactual predictions for macro variables and welfare are exactly the same as in the Armington model.[12] This stronger equivalence result, however, is very sensitive to the assumption that fixed exporting costs are paid in the importing country. If fixed exporting costs are partly paid in the origin country, then one-sector gravity models with monopolistic competition and firm-level heterogeneity à la Melitz (2003) generally predict different changes in relative factor prices, i.e., relative wages if labor is the only factor of production. This leads to different changes in the share of expenditure on domestic goods and, in turn, different welfare changes for a given change in trade costs (moving to autarky being a notable exception).[13]

3.3 Multiple Sectors

Gravity models can be extended to multiple sectors, $s = 1, ..., S$, by assuming a two-tier utility function in which the upper-level is Cobb-Douglas and the lower-level is CES; see e.g. Anderson and Yotov (2010), Donaldson (2008), Caliendo and Parro (2010), Costinot, Donaldson, and Komunjer (2010), Hsieh and Ossa (2011), Levchenko and Zhang (2011), Ossa (2012), Shikher (forthcominga), and Shikher (forthcomingb). Formally let us assume that the representative agent in

---

[12]In this special case, one can show that the welfare impact of changes in the number of domestic varieties (if any) exactly compensates the welfare impact of changes in the number of foreign varieties (if any), as emphasized in Feenstra (2010). In general, i.e., if fixed costs are partly paid in the origin country, these exact offsetting effects no longer hold, though the welfare change associated with a shock to variable trade costs remains given by Equation (9), as discussed above. Similarly, in the case of shocks to foreign endowments or technology, these exact offsetting effects no longer hold, though welfare changes remain given by Equation (9), as shown in Arkolakis, Costinot, and Rodríguez-Clare (2012). In short, the main equivalence result in Arkolakis, Costinot, and Rodríguez-Clare (2012) does not hinge on these exact offsetting effects.

[13]For finite changes in trade costs, different counterfactual predictions do not arise from selection effects per se, but rather from the fact that fixed exporting costs, if paid at least partly in the origin country, can affect relative demand for factors of production in different countries and so relative factor prices. To see this, note that in a symmetric world economy in which relative factor prices are constant, c_{ij}^x would be constant across countries as well. Thus, in spite of selection effects, c_{ij}^x would again drop out of Equation (16), leading to the same counterfactual predictions as in the Armington model.

each country aims to maximize

$$C_j = \prod_{s=1}^{S} C_{j,s}^{\beta_{j,s}}, \tag{18}$$

where $\beta_{j,s} \geq 0$ are exogenous preference parameters satisfying $\sum_{s=1}^{S} \beta_{j,s} = 1$ and $C_{j,s}$ is total consumption of the composite good s in country j,

$$C_{j,s} = \left(\int_{\omega \in \Omega} c_{j,s}(\omega)^{(\sigma_s - 1)/\sigma_s} d\omega \right)^{\sigma_s/(\sigma_s - 1)}, \tag{19}$$

where $\sigma_s > 1$ is the elasticity of substitution between different varieties, which is allowed to vary across sectors. In multi-sector gravity models, each variety remains sourced from only one country so that a sector level version of (1) still holds, $C_{j,s} = \left(\sum_{i=1}^{n} C_{ij,s}^{(\sigma_s - 1)/\sigma_s} \right)^{\sigma_s/(\sigma_s - 1)}$, with $C_{ij,s} = \int_{\omega \in \Omega_{ij,s}} c_{j,s}(\omega)^{(\sigma_s - 1)/\sigma_s} d\omega$. The associated consumer price index is $P_j = \prod_{s=1}^{S} P_{j,s}^{\beta_{j,s}}$, with sector-specific price indices given by a sector level version of (2), $P_{j,s} = \left(\sum_{i=1}^{n} P_{ij,s}^{1-\sigma_s} \right)^{1/(1-\sigma_s)}$, with $P_{ij,s} = \left(\int_{\omega \in \Omega_{ij,s}} p_{j,s}(\omega)^{(1-\sigma_s)} d\omega \right)^{1/(1-\sigma_s)}$.

In line with the literature, we assume that our reduced-form assumption on price indices, Equation (14), now holds sector-by-sector; that factors of production are used in the same way across all activities in all sectors, so that $c_{i,s}^{p} = c_{ii,s}^{m} = c_{i,s}^{e} = Y_i$; and that trade in goods is balanced, $R_i = Y_i$.[14] Combining these assumptions, we get

$$P_{ij,s} = \tau_{ij,s} Y_i \left[\left(e_{j,s} \frac{E_j}{c_{ij,s}^{x}} \right)^{\frac{\delta_s}{1-\sigma_s}} \frac{\tau_{ij,s} Y_i}{P_{j,s}} \right]^{\eta_s} r_{i,s}^{\frac{\delta_s}{1-\sigma_s}} \xi_{ij,s}. \tag{20}$$

where $e_{j,s} \equiv E_{j,s}/E_j$ denotes the share of total expenditure in country j allocated to sector s and $r_{i,s} = R_{i,s}/R_i$ denote the share of total revenues in country i generated from sector s. All other variables have the same interpretation as in Section 3.1, except that they are now free to vary across sectors. Compared to Equation (15), Equation (20) allows for scale effects in monopolistically competitive sectors both through selection, $\left(e_{j,s} E_j/c_{ij,s}^{x} \right)^{\frac{\delta_s}{1-\sigma_s}}$, and entry, $r_{i,s}^{\frac{\delta_s}{1-\sigma_s}}$. In the one-sector case, the latter effect is necessarily absent because $r_{i,s} = 1$. Here an expansion of production in a monopolistically competitive sector, i.e. a higher value of $r_{i,s}$, leads to entry and gains from new varieties, i.e. a lower value of $P_{ij,s}$, with the standard "love of variety" elasticity of $1/(\sigma_s - 1)$.[15]

We now focus on the gains from international trade, as defined in Section 2.3, and discuss how they are affected by the introduction of multiple sectors under the assumption that Equation (20) holds. Bilateral trade flows at the sector-level satisfy $X_{ij,s} = \left(P_{ij,s}/P_{j,s} \right)^{1-\sigma_s} e_{j,s} E_j$. Together with

---

[14]The assumption, $R_i = Y_i$, is stronger than in the one-sector case. In Section 3.2, it holds in models with monopolistic competition and firm-level heterogeneity independently of where fixed exporting costs are paid, as long as productivity distributions are Pareto. In a multi-sector environment, this may no longer be true even under Pareto if fixed exporting costs are partly paid in the importing country and sectors differ in the share of revenues associated with fixed exporting costs. Such considerations lead to a slight change in the present analysis, which we come back to below.

[15]Similar scale effects can be introduced in the one sector case by allowing factor supply to be elastic; see Balistreri, Hillberry, and Rutherford (2009). The magnitude of such effects then crucially depends on the elasticity of factor supply.

Equation (20), this implies the sector-level gravity equation:

$$X_{ij,s} = \frac{(\tau_{ij,s}Y_i)^{-\varepsilon_s} \left(c_{ij}^x\right)^{-\delta_s\eta_s} r_{i,s}^{\delta_s} X_{ij,s}}{\sum_l (\tau_{lj,s}Y_l)^{-\varepsilon_s} \left(c_{lj,s}^x\right)^{-\delta_s\eta_s} r_{l,s}^{\delta_s} X_{lj,s}} e_{j,s} E_j,$$ (21)

where $\varepsilon_s = (1 + \eta_s)(\sigma_s - 1)$ and $\chi_{ij,s} \equiv \xi_{ij,s}^{1-\sigma_s}$. As in Section 3.1, one can combine Equation (20), Equation (21), and the fact that $\lambda_{jj,s} = \left(P_{jj,s}/P_{j,s}\right)^{1-\sigma_s}$ to show that changes in real consumption associated with a trade shock are now given by:

$$\hat{C}_j = \prod_{s=1}^S \left(\hat{\lambda}_{jj,s} \left(\hat{e}_{j,s}^{\eta_s}\hat{r}_{j,s}\right)^{-\delta_s}\right)^{-\beta_{j,s}/\varepsilon_s}.$$ (22)

Under Cobb-Douglas preferences, we know that $e_{j,s} = e'_{j,s} = \beta_{j,s}$. Thus the previous expression can be simplified further into $\hat{C}_j = \prod_{s=1}^S \left(\hat{\lambda}_{jj,s}\hat{r}_{j,s}^{-\delta_s}\right)^{-\beta_{j,s}/\varepsilon_s}$.[16] To compute the gains from trade, we only need to solve for $\hat{r}_{j,s}$ when the counterfactual entails autarky. Since $r'_{j,s} = e'_{j,s}$ under autarky and $e_{j,s} = e'_{j,s} = \beta_{j,s}$, we must have $\hat{r}_{j,s} = e_{j,s}/r_{j,s}$. Using the fact that $\hat{\lambda}_{jj,s} = 1/\lambda_{jj,s}$ for all s, we then get

$$G_j = 1 - \prod_{s=1}^S \left(\lambda_{jj,s} \left(\frac{e_{j,s}}{r_{j,s}}\right)^{\delta_s}\right)^{\beta_{j,s}/\varepsilon_s}.$$ (23)

Since δ_s appears in Equation (23), gains from trade predicted by multi-sector gravity models with monopolistic competition differ from those predicted by models with perfect competition because of scale effects, as discussed in Arkolakis, Costinot, and Rodríguez-Clare (2012). In contrast, selection effects still have no impact on the overall magnitude of the gains from trade. Since η_s does not appear in Equation (23), conditional on observed trade flows and the trade elasticity, the gains from trade predicted by monopolistically competitive gravity models with and without firm-level heterogeneity are the same, even with multiple sectors.

To compute the gains from trade using Equation (23), we need measures of $\lambda_{jj,s}$, $e_{j,s}$, $\beta_{j,s}$ and $r_{j,s}$ as well as sector-level trade elasticities ε_s for $s = 1, ..., S$. To compute $\lambda_{jj,s}$, $e_{j,s}$, $\beta_{j,s}$ and $r_{j,s}$, we use data on 35 sectors from the WIOD in 2008.[17] Sector-level trade elasticities are from Caliendo and Parro (2010).For the purposes of this chapter, the main advantage of the estimation procedure in Caliendo and Parro (2010) is that it is consistent with all quantitative trade models satisfying the sector-level gravity equation (21).[18]

---

[16]Without balanced trade in goods, $R_j \neq Y_j$, this would generalize to $\hat{C}_j = \prod_{s=1}^S \left(\hat{\lambda}_{jj,s} \left(\hat{r}_{j,s}\hat{R}_j/\hat{Y}_j\right)^{-\delta_s}\right)^{-\beta_{j,s}/\varepsilon_s}$.

[17]In theory, since the models that we consider do not feature intermediate goods, we have $e_{j,s} = \beta_{j,s}$. In the data, however, gross expenditure shares, $e_{j,s}$, differ from final demand shares, $\beta_{j,s}$. In the analysis that follows we let $\beta_{j,s}$ be different from $e_{j,s}$ when computing gains from trade using Equation (23).

[18]Caliendo and Parro (2010) use COMTRADE data from 1993. They assume that iceberg trade costs can be decomposed into $\tau_{ij,s} = t_{ij,s}d_{ij,s}\mu_{ij,s}$ where $t_{ij,s}$ is one plus the ad-valorem tariff applied by country j on good s imported from i and d_{ni} is a symmetric component of the iceberg-trade cost, i.e., $d_{ij,s} = d_{ji,s}$. Taking a triple log-difference of

Columns 2 and 3 in Table 1 reports the gains from trade G_j for the same set of countries as in Section 2.3, but using Equation (23) rather than Equation (12). In Column 2, all sectors are assumed to be perfectly competitive, $\delta_s = 0$ for all s, while in column 3, all sectors are assumed to be monopolistically competitive, $\delta_s = 1$ for all s. Two features of these results stand out. First, even in an environment with active entry effects, such as the one considered here, there are no systematic differences between the gains from trade predicted by multi-sector models with perfect competition, Column 2, and those predicted by models with monopolistic competition, Column 3. For some countries the gains under monopolistic competition are larger than under perfect competition (e.g., gains in Germany increase from 12.7% to 17.6%), while for other countries the opposite holds (e.g., gains for Greece decrease from 16.3% to 4.7%). Second, the gains from trade predicted by multi-sector models under both market structures, Columns 2 and 3, are significantly larger than those predicted by one-sector models, Column 1. For example, moving from Column 1 to Column 2 increases G_j for Belgium from 7.5% to 32.7%, while it increases G_j for Canada from 3.8% to 17.4%. The average among all the countries in Table 1 more than triples, increasing from 4.4% (Column 1) to 15.3% (Column 2), a point also emphasized in Ossa (2012).

The fact that models with perfect and monopolistic competition predict, on average, similar gains reflect the opposite consequences of entry effects—$\prod_{s=1}^{S}\left(e_{j,s}/r_{j,s}\right)^{\beta_{j,s}\delta_s/\varepsilon_s}$ in Equation (23)—on countries with a comparative advantage and a comparative disadvantage in sectors with strong scale effects. If $e_{j,s}/r_{j,s}$ is negatively correlated with δ_s/ε_s, then country j tends to be a large exporter of goods with strong scale effects. For such a country, the gains from trade tend to be larger than in the absence of scale effects since trade allows specialization in the sectors characterized by strong returns to scale. The converse is true, however, for a country in which $e_{j,s}/r_{j,s}$ is positively correlated with δ_s/ε_s, i.e., a country with a comparative disadvantage in sectors with strong scale effects. In theory, such a country may even lose from opening up to trade. This idea is the basis of Frank Graham's argument for protection; see Ethier (1982a) and Helpman and Krugman (1985) for a general discussion.

Why do multi-sector gravity models predict much larger gains than their one-sector counterparts? Part of the answer is: Cobb-Douglas preferences. This assumption implies that if the price of a single good gets arbitrarily large as a country moves to autarky—because it cannot produce that good—then gains from trade are infinite. According to Equation (23), this will happen either if the share of expenditure on domestic goods, $\lambda_{jj,s}$, is close to zero—which implies arbitrarily large costs of production for that good at home—or if the trade elasticity, ε_s, is close to zero—which implies that foreign varieties are essential. We come back to this issue when discussing the more general case of nested CES utility functions in Section 5.3.

Equation (21) with $\delta_s = 0$ yields

$$\ln \frac{X_{ij,s}X_{jl,s}X_{li,s}}{X_{il,s}X_{lj,s}X_{ji,s}} = \varepsilon_s \ln \frac{t_{ij,s}t_{jl,s}t_{li,s}}{t_{il,s}t_{lj,s}t_{ji,s}} + \mu_{ijl,s}.$$

where the error-term $\mu_{ijl,s} = \ln \mu_{ij,s}^{\varepsilon_s}\chi_{ij,s} + \ln \mu_{jl,s}^{\varepsilon_s}\chi_{jl,s} + \ln \mu_{li,s}^{\varepsilon_s}\chi_{li,s} - \ln \mu_{il,s}^{\varepsilon_s}\chi_{il,s} - \ln \mu_{lj,s}^{\varepsilon_s}\chi_{lj,s} - \ln \mu_{ji,s}^{\varepsilon_s}\chi_{ji,s}$. We discuss how we map WIOD and COMTRADE data in the online Appendix.

3.4 Tradable Intermediate Goods

We now enrich the supply-side of gravity models by introducing tradable intermediate goods and input-output linkages as in the early work of Krugman and Venables (1995), Eaton and Kortum (2002), Alvarez and Lucas (2007), and more recently Di Giovanni and Levchenko (2009), Caliendo and Parro (2010), and Balistreri, Hillberry, and Rutherford (2011). Formally, we maintain the same preference structure as in the previous section and introduce intermediate goods parsimoniously by assuming that, in each sector s, they are produced in the exact same way as composite goods for final consumption:

$$I_{j,s} = \left(\int_{\omega \in \Omega} i_{j,s}(\omega)^{(\sigma_s - 1)/\sigma_s} d\omega \right)^{\sigma_s/(\sigma_s - 1)}, \tag{24}$$

where $i_{j,s}(\omega)$ denote the amount of variety ω used in the production of intermediate goods in country j and sector s. Accordingly, the sector-level price index $P_{j,s}$ defined in Section 3.3 now measures the aggregate price of sector s goods in country j for both final consumption and production.

As we did in the previous section, and in line with the existing literature, we assume that sector-level price indices satisfy Equation (14) and that trade in goods is balanced so that, together with the overall trade balance, we have total expenditure equals total producer revenues, $E_i = R_i$, in each country. Compared to Section 3.3, we allow $c_{i,s}^p$ to vary across sectors to reflect the differential effect of intermediate goods on unit costs of production in different sectors:

$$c_{i,s}^p = Y_i^{1-\alpha_{i,s}} \prod_{k=1}^S P_{i,k}^{\alpha_{i,ks}}, \tag{25}$$

where $\alpha_{i,ks} \geq 0$ are exogenous technology parameters such that $\alpha_{i,s} \equiv \sum \alpha_{i,ks} \in [0,1]$. As in the previous section, we assume that all sectors use primary factors in the same way, hence the term Y_i in Equation (25). In line with the existing literature, we also assume that entry and exporting activities also use intermediate goods in the same proportion as production, $c_{i,s}^e = c_{ii,s}^x = c_{i,s}^p$.[19]

Under the previous assumptions, Equation (14) now implies

$$P_{ij,s} = \tau_{ij,s} c_{i,s} \left[\left(\frac{e_{j,s}}{v_j} \frac{Y_j}{c_{ij,s}^x} \right)^{\frac{\delta_s}{1-\sigma_s}} \frac{\tau_{ij,s} c_{i,s}}{P_{j,s}} \right]^{\eta_s} \left(\frac{r_{i,s}}{v_i} \frac{Y_i}{c_{i,s}} \right)^{\frac{\delta_s}{1-\sigma_s}} \xi_{ij,s}, \tag{26}$$

where $c_{i,s} = c_{i,s}^p$ as given by Equation (25); $v_i \equiv Y_i / R_i$ is the ratio of total income to total revenues in country i; and $e_{i,s}$ and $r_{i,s}$ still denote expenditure and revenue shares, the difference being that expenditure and revenue are now "gross," as they include the purchases and sales of both consumption and intermediate goods. In the absence of intermediate goods we have $\alpha_{i,s} = 0$ and $v_i = 1$ for all i and s, so the previous equation reduces to (20).

---

[19]Alternatively one could assume that entry and exporting activities only use primary factors of production: $c_{i,s}^e = c_{ii,s}^x = Y_i$. This assumption, however, immediately creates inconsistencies between the predictions of models of monopolistic competition and our dataset. Given our estimates of the trade elasticities ε_s, the factor costs associated with fixed entry and exporting activities is sometimes higher than the total factor costs observed in the data.

Bilateral trade flows now include trade in consumption and intermediate goods. But since both are combined using the same CES aggregator, Equations (19) and (24), trade flows still satisfy $X_{ij,s} = \left(P_{ij,s}/P_{j,s}\right)^{1-\sigma_s} e_{j,s} E_j$. Together with Equation (26), this implies the following sector-level gravity equation:

$$X_{ij,s} = \frac{\left(\tau_{ij,s} c_{i,s}\right)^{-\varepsilon_s} \left(c_{ij,s}^x\right)^{-\delta_s \eta_s} \left(\frac{r_{i,s}}{v_i} \frac{Y_i}{c_{i,s}}\right)^{\delta_s} \chi_{ij,s}}{\sum_{l=1}^n \left(\tau_{lj,s} c_{l,s}\right)^{-\varepsilon_s} \left(c_{lj,s}^x\right)^{-\delta_s \eta_s} \left(\frac{r_{l,s}}{v_l} \frac{Y_l}{c_{l,s}}\right)^{\delta_s} \chi_{lj,s}} e_{j,s} E_j, \tag{27}$$

where again $\chi_{ij,s} \equiv \xi_{ij,s}^{1-\sigma_s}$. One can now follow a similar strategy as in previous sections to show that welfare changes associated with a foreign shock are given by

$$\hat{C}_j = \prod_{s,k=1}^S \left(\hat{\lambda}_{jj,k} \left(\left(\frac{\hat{e}_{j,k}}{\hat{v}_j}\right)^{\eta_k} \left(\frac{\hat{r}_{j,k}}{\hat{v}_j}\right)\right)^{-\delta_k}\right)^{-\beta_{j,s} \widetilde{a}_{j,sk}/\varepsilon_k}, \tag{28}$$

where $\widetilde{a}_{j,sk}$ is the elasticity of the price index in sector s with respect to changes in the price index in sector k. These price elasticities are given by the elements of the "adjusted Leontief inverse" of the input-output matrix, i.e., the $(S \times S)$ matrix $\left(Id - \widetilde{A}_j\right)^{-1}$, with the elements of the $\widetilde{A}_j$ matrix given by the adjusted technology parameters $\widetilde{\alpha}_{j,sk} \equiv \alpha_{j,sk} \left(1 + \delta_k \left(1 + \eta_k\right)/\varepsilon_k\right)$. Under perfect competition, $\left(Id - \widetilde{A}_j\right)^{-1}$ is the standard Leontief inverse matrix, i.e., $\left(Id - \widetilde{A}_j\right)^{-1} = \left(Id - A_j\right)^{-1}$ where A_j is the matrix with typical element $\alpha_{j,sk}$. Under monopolistic competition with intermediate goods used in entry and exporting activities, however, we need to adjust the technology parameters $\alpha_{j,sk}$ by $1 + \left(1 + \eta_k\right)/\varepsilon_k = 1 + 1/\left(\sigma_k - 1\right)$ to take into account that a decline in the price index of sector s not only decreases the price index of sector k through the standard input-output channels, but also by lowering fixed entry and exporting costs, thereby increasing the number of available varieties.

In addition, intermediate goods affects the results derived in Section 3.3 in two important ways. First, expenditure shares, $e_{j,s}$, are no longer equal to exogenous consumption shares, $\beta_{j,s}$. Hence, $\hat{e}_{j,s}$ may be different from one, which implies that the use of the formula given in Equation (28) requires either observing $\hat{e}_{j,s}$, after the shock, or predicting it, before the shock. Second, the scale effect term, $\left(\hat{e}_{j,k}/\hat{v}_j\right)^{\eta_k} \left(\hat{r}_{j,k}/\hat{v}_j\right)$, now depends on the change in the ratio of value added to gross output, $\hat{v}_j$. Intuitively, under monopolistic competition, welfare depends on entry, which itself depends on the ratio of revenues to factor prices, v_j.

The associated formula for the gains from international trade is obtained from (28) by setting $\hat{\lambda}_{jj,k} = 1/\lambda_{jj,k}$ and by solving for $\left(\frac{\hat{e}_{j,k}}{\hat{v}_j}\right)^{\eta_k} \left(\frac{\hat{r}_{j,k}}{\hat{v}_j}\right)$ as we move to autarky. Since $e_{j,k}$ and $r_{j,k}$ are data while $e_{j,k}^A = r_{j,k}^A$ in autarky, then we just need to solve for $e_{j,k}^A/v_j^A$. In the online Appendix we show that this is given by $e_{j,k}^A/v_j^A = \sum_{l=1}^S \beta_{j,l} a_{j,kl}$, where $a_{j,kl}$ are now the elements of the Leontief inverse

19

$\left(Id - A_j \right)^{-1}$. The gains from trade are then

$$G_j = 1 - \prod_{s,k=1}^{S} \left(\lambda_{jj,k} \left(\left(\frac{e_{j,k}}{b_{j,k}} \right)^{\eta_s} \frac{r_{j,k}}{b_{j,k}} \right)^{-\delta_k} \right)^{\beta_{j,s} \tilde{a}_{j,sk} / \varepsilon_k}, \tag{29}$$

where $b_{j,k} \equiv v_j \left(\sum_{l=1}^{S} \beta_{j,l} a_{j,kl} \right)$ summarizes how intermediate goods affect the magnitude of scale effects in models with monopolistic competition.

To implement the previous formula using the WIOD data, we compute $\lambda_{jj,s}$, $R_{j,s}$, $E_{j,s}$, $r_{j,s}$, $e_{j,s}$, and $\beta_{j,s}$ in the exact same way as in Section 3.3. In the raw data, we also observe purchases, $X_{ij,ks}$, of intermediate goods from sector k and country i in sector s and country j. Using those, we can then compute shares of intermediate purchases $\alpha_{j,ks} = \sum_i X_{ij,ks} / R_{j,s}$ and value added by sector as $Y_{j,s} = R_{j,s} - \sum_k \sum_i X_{ij,ks}$. In a number of simulations below, we also follow Balistreri, Hillberry, and Rutherford (2011) and use an alternative measure of shares of intermediate goods, $\alpha_{j,ks}^* = \left(\sum_k \sum_i X_{ij,ks} / R_{j,s} \right) \times \left(E_{j,k} / E_j \right)$, when computing gains from trade using Equation (29). Compared to the true share, $\alpha_{j,ks}$, this alternative measure, $\alpha_{j,ks}^*$, counterfactually assumes that firms allocate expenditure on intermediate goods from different sectors in the same proportions, $E_{j,k} / E_j$, though some sectors may have higher shares of intermediate goods, $\sum_k \sum_i X_{ij,ks} / R_{j,s}$. We come back to the benefit of this simplification in a moment.

Columns 4-7 of Table 1 report the gains from trade G_j under different market structures using Equation (29). Column 4 corresponds to gains from trade under perfect competition, $\delta_s = 0$ for all s, using the true intermediate good shares, $\alpha_{j,ks} = \sum_i X_{ij,ks} / R_{j,s}$. We see that predicted gains from trade are much higher than those predicted by the same models without intermediate goods (Column 2). For example, the gains from trade for the United States and Spain in Column 4 are twice as high as those in Column 2, while for Japan the gains increase by a factor of five. One can think about these results in two ways. First, trade in intermediates leads to a decline in the price of domestic goods, which implies additional welfare gains. If domestic goods are used as inputs in domestic production, this triggers additional rounds of productivity gains, leading to even larger gains; this is the input-output loop often mentioned in the literature.[20] Second, for given data on the share of *expenditure* on domestic goods, $\lambda_{jj,s}$, models featuring intermediate goods necessarily predicts more trade relative to total *income*. So, perhaps, it should not be too surprising that the same models predict that real income increases by more because of trade.

Ideally, one would like to study the predictions of models of monopolistic competition using the same data on intermediate good shares, $\alpha_{j,ks}$. Unfortunately, this is not possible in the context of our dataset since it would lead some of the elements of the adjusted Leontief matrix to become negative, in which case gains from trade G_j are not well-defined. To understand why this issue arises, consider a simpler model with monopolistic competition, $\delta_s = 1$, intermediate goods, $\alpha_s >$

---

[20] A simple way to illustrate this mechanism is to return to the one-sector model. In the case of perfect competition, the formula above becomes $G_j = 1 - \left(\lambda_{jj} \right)^{1/\varepsilon(1-\alpha)}$; see Eaton and Kortum (2002) and Alvarez and Lucas (2007). Thus a higher share of intermediate goods, α, leads to higher gains from trade.

0, but only one sector, $S = 1$. In that environment, Equation (29) simplifies to $G_j = 1 - \lambda_{jj}^{\tilde{a}_j/\varepsilon}$, where $\tilde{a}_j = \left(1 - \alpha_j \left(1 + \frac{1}{\sigma-1}\right)\right)^{-1}$. As $\alpha_j \left(1 + \frac{1}{\sigma-1}\right)$ gets close to one, $\tilde{a}_j$ goes to infinity and real consumption in autarky goes to zero. If $\alpha_j \geq (\sigma - 1)/\sigma$ then the price index, real consumption and of course G_j are not well-defined. Intuitively, a given increase in the number of varieties leads to a decline in the price index, which triggers a decline in the cost of entry, which, in turn, leads to a further increase in the number of varieties. We obtain infinite amplification whenever the share of intermediates in production, α_j, is high relative to the love of variety, $1/(\sigma - 1)$. This is precisely what happens when any of the elements of the adjusted Leontief matrix becomes negative.

To get around this issue, Columns 5-7 report gains from trade using the alternative measure of intermediate good shares, $\alpha_{j,ks}^*$, as in Balistreri, Hillberry, and Rutherford (2011). To make sure that the difference between models of perfect and monopolistic competition is not being driven by a different treatment of intermediate goods, Column 5 again reports the gains from trade under perfect competition. The results are very similar whether we use true shares (Column 4) or alternative shares (Column 5). Column 6 reports gains from trade under monopolistic competition without firm-level heterogeneity, $\delta_s = 1$ and $\eta_s = 0$, whereas Column 7 reports gains under monopolistic competition with firm-level heterogeneity, $\delta_s = 1$ and $\eta_s > 0$. Following Balistreri, Hillberry, and Rutherford (2011), we set $\eta_s = 0.65$ for all s.[21] In general, gains from trade are slightly higher with monopolistic than perfect competition. For example, the gains for the US increase from 8% to 8.6% as we move from Column 5 to Column 6 in Table 1. Across all countries, the average gains increase from 27.1% to 32.3%. The intuition is simple. When entry activities use intermediate goods, trade leads to a decline in the cost of entry and hence to an expansion in the variety of goods produced domestically, bringing about additional welfare gains.

As shown by the results in Column 7, the gains from trade are even higher when we allow for firm-level heterogeneity. To see why, recall that $\varepsilon_s = (\sigma_s - 1)(1 + \eta_s)$, hence if $\varepsilon_s = 3.2$ (as in the Chemicals sector), then $\eta_s = 0.65$ implies $\sigma_s = 2.9$, whereas under the assumption $\eta_s = 0$ we would have concluded $\sigma_s = 4.2$. The difference in the implied elasticity of substitution σ_s between models with and without firm-level heterogeneity leads to large differences in the magnitude of the scale effects arising from love of variety. We come back to this issue in more detail in Section 5.3. For now, we merely want to point out that: (*i*) welfare calculations are highly sensitive to the value of this parameter; and (*ii*) the reason behind this sensitivity is that conditional on the value of the trade elasticity ε_s, the value of η_s pins down σ_s and, in turn, the magnitude of scale effects.

In summary, the introduction of tradable intermediate goods dramatically increases the magnitude of the gains from trade, both under perfect and monopolistic competition. Under the latter market structure, the scale effects associated with decreases in the price of intermediate goods are so large that if one were to use the true shares of intermediate purchases across countries and sectors, $\alpha_{j,ks}$, rather than a made-up average, $\alpha_{j,ks}^*$, one would conclude that for all countries in our

---

[21]Balistreri, Hillberry, and Rutherford (2011) use non-linear least squares to estimate the trade elasticity for manufacturing as a whole (as well as other parameters). Their preferred estimate is $\varepsilon = 4.58$. Following Bernard, Eaton, Jensen, and Kortum (2003), they set $\sigma = 3.8$. Since $\varepsilon = (\sigma - 1)(1 + \eta)$, these two values imply that $\eta = 0.65$. In Section 5.3 we discuss more direct ways to estimate η_s across sectors using firm-level data.

dataset, gains from trade cannot be finite.

3.5 Multiple Factors of Production

So far, we have restricted ourselves to gravity models featuring only one factor of production, or equivalently multiple factors of production that are used in the same proportions in all sectors. This assumption is formally reflected in the fact that producer prices are proportional to GDP, Y_i, in Equations (20) and (25). In this section we introduce differences in factor intensity across sectors, as in the extensions of Eaton and Kortum (2002), Bernard, Eaton, Jensen, and Kortum (2003), and Melitz (2003) considered by Chor (2010), Burstein and Vogel (2010), and Bernard, Redding, and Schott (2007), respectively.

For expositional purposes, we restrict ourselves to an economic environment in which there are only two factors of production, skilled labor and unskilled labor, and no intermediate goods. Throughout this subsection, we assume that aggregate production functions are CES in all countries and sectors and given by

$$Q_{j,s} = \left[\mu_s^H \left(H_{j,s} \right)^{(\rho-1)/\rho} + \mu_s^L \left(L_{j,s} \right)^{(\rho-1)/\rho} \right]^{\rho/(\rho-1)},$$

where $H_{j,s}$ and $L_{j,s}$ denote total employment of skilled and unskilled workers, respectively, in country j and sector s; $\rho > 0$ is the elasticity of substitution between skilled and unskilled labor; and $\mu_j^f > 0$ determines the intensity of factor $f = H, L$ in sector s, with $\mu_{j,s}^H + \mu_{j,s}^F = 1$. In line with the existing literature, we also assume that factors of production have the same share in variable costs of production as in entry and exporting costs (if any): $c_{i,s}^e = c_{ii,s}^m = c_{i,s}^p \equiv c_{i,s}$. Assuming as above that trade in goods is balanced, $Y_i = R_i$, Equation (20) then generalizes to

$$P_{ij,s} = \tau_{ij,s} c_{i,s} \left[\left(e_{j,s} \frac{E_j}{c_{ij,s}^x} \right)^{\frac{\delta_s}{1-\sigma_s}} \frac{\tau_{ij,s} c_{i,s}}{P_{j,s}} \right]^{\eta_s} \left(r_{i,s} \frac{Y_i}{c_{i,s}} \right)^{\frac{\delta_s}{1-\sigma_s}} \zeta_{ij,s}. \tag{30}$$

All variables have the same interpretation as in previous sections, except for the fact that unit costs are now proportional to

$$c_{i,s} = \left[\left(\mu_s^H \right)^\rho \left(w_i^H \right)^{1-\rho} + \left(\mu_s^L \right)^\rho \left(w_i^L \right)^{1-\rho} \right]^{1/(1-\rho)}, \tag{31}$$

where w_i^H and w_i^L are the wages of skilled and unskilled workers, respectively, in country i.

As in previous sections, if lower-level utility functions are CES, then gravity holds in this environment. Combining Equations (19) and (30), we get

$$X_{ij,s} = \frac{\left(\tau_{ij,s} c_{i,s} \right)^{-\varepsilon_s} \left(c_{ij,s}^x \right)^{-\delta_s \eta_s} \left(r_{i,s} \frac{Y_i}{c_{i,s}} \right)^{\delta_s} \chi_{ij,s}}{\sum_{l=1}^n \left(\tau_{lj,s} c_{l,s} \right)^{-\varepsilon_s} \left(c_{lj,s}^x \right)^{-\delta_s \eta_s} \left(r_{l,s} \frac{Y_l}{c_{l,s}} \right)^{\delta_s} \chi_{lj,s}} e_{j,s} E_j.$$

22

For any trade shock, we can normalize factor prices such that $Y_j = Y'_j$. Using this normalization, we can then express welfare changes as

$$\hat{C}_j = \Pi_{s=1}^{S} \left(\hat{c}_{j,s}\right)^{-\beta_{j,s}} \left(\hat{\lambda}_{jj,s} \left(\left(\hat{c}_{j,s}\right)^{-(1+\eta_s)} \hat{e}_{j,s}^{\eta_s} \hat{r}_{j,s}\right)^{-\delta_s}\right)^{-\beta_{j,s}/\varepsilon_s}.$$

Gains from trade, in turn, are given by

$$G_j = 1 - \Pi_{s=1}^{S} \left(\hat{c}_{j,s}^A\right)^{-\beta_{j,s}} \left(\lambda_{jj,s} \left(\left(\hat{c}_{j,s}^A\right)^{-(1+\eta_s)} \frac{e_{j,s}}{r_{j,s}}\right)^{\delta_s}\right)^{\beta_{j,s}/\varepsilon_s}, \tag{32}$$

where $\hat{c}_{j,s}^A$ denote the change in production costs between the initial equilibrium and autarky. Compared to one-factor models, changes in real consumption now also depend on changes in relative factor prices, which affects production costs across sectors, as reflected in $\hat{c}_{j,s}$ and $\hat{c}_{j,s}^A$. Under perfect competition, such changes only affect variable costs of production, whereas under monopolistic competition, they also affect the fixed costs of exporting and entry and so, the number of available varieties in country j and sector s, as reflected in the extra terms $\left(\hat{c}_{j,s}\right)^{-(1+\eta_s)}$ and $\left(\hat{c}_{j,s}^A\right)^{-(1+\eta_s)}$.

In order to compute changes in production costs, one can again use the exact hat algebra introduced in Section 2.2. We illustrate here how this can be done as we move from the initial equilibrium to autarky, though the same methodology can be applied to any shock. Equation (31) implies

$$\hat{c}_{j,s}^A = \left[\varphi_{j,s}^H \left(\hat{w}_i^{A,H}\right)^{1-\rho} + \varphi_{j,s}^L \left(\hat{w}_i^{A,L}\right)^{1-\rho}\right]^{1/(1-\rho)}, \tag{33}$$

where $\varphi_{j,s}^f \equiv \left(\mu_s^f\right)^{\rho} \left(w_j^f\right)^{1-\rho} / c_{j,s}^{1-\rho}$ is the share of total factor spending going to factor f in country j and sector s in the initial equilibrium. Changes in factor prices, in turn, can be computed by manipulating the two factor-market clearing conditions:

$$\hat{w}_j^{A,H} = \Sigma_{s=1}^{S} h_{j,s} \left(\frac{e_{j,s}}{y_{j,s}}\right) \left(\frac{\left(\hat{w}_j^{A,H}\right)^{1-\rho}}{\varphi_{j,s}^H \left(\hat{w}_j^{A,H}\right)^{1-\rho} + \varphi_{j,s}^L \left(\hat{w}_j^{A,L}\right)^{1-\rho}}\right), \tag{34}$$

$$\hat{w}_j^{A,L} = \sum_{k=1}^{S} l_{j,s} \left(\frac{e_{j,s}}{y_{j,s}}\right) \left(\frac{\left(\hat{w}_j^{A,L}\right)^{1-\rho}}{\varphi_{j,s}^H \left(\hat{w}_j^{A,H}\right)^{1-\rho} + \varphi_{j,s}^L \left(\hat{w}_j^{A,L}\right)^{1-\rho}}\right), \tag{35}$$

where $h_{j,s} \equiv H_{j,s}/H$ and $l_{j,s} \equiv L_{j,s}/L$ denote the share of skilled and unskilled workers, respectively, employed in sector s in country j in the initial equilibrium, and $y_{j,s} \equiv Y_{j,s}/Y_j$ denotes the share of total income earned in sector s in country j. Combining Equations (32)-(35), we can compute the gains from trade in the multi-factor case.

To implement this new formula, we need additional data on the elasticity of substitution be-

tween skilled and unskilled workers, ρ, the share of employment of skilled and unskilled workers across sectors, $h_{j,s}$ and $l_{j,s}$, and the factor cost shares, $\varphi_{j,s}^f$. For simplicity, we assume Cobb-Douglas technologies, i.e., $\rho = 1$, and common cost shares across countries, i.e., $\varphi_{j,s}^f = \varphi_s^f$.[22] We compute factor cost shares from the NBER manufacturing database that contains information about employment and average wages for both production and non-production workers in the United States between 1987 and 2005. Following Berman, Bound, and Griliches (1994), we treat skilled workers in the model as non-production workers in the data and unskilled workers in the model as production workers in the data. Given cost shares, employment shares are computed as $h_{j,s} = \frac{\varphi_s^H y_{j,s}}{\sum_k \varphi_k^H y_{j,k}}$ and $l_{j,s} = \frac{\varphi_s^L y_{j,s}}{\sum_k \varphi_k^L y_{j,k}}$. Since we do not have data on cost shares for sectors outside of manufacturing, we aggregate all non-manufacturing sectors into a single sector which we assume is non-tradable and which has factor cost shares equal to the overall factor cost shares in manufacturing.

When computing the gains from trade using Equation (32) under the assumption of perfection competition, $\delta_s = 0$ for all s, we find gains from trade that are virtually the same as those presented in Column 2. This reflects the fact that the factor content of trade for skilled and unskilled labor is basically zero in our dataset. In particular, one can check that under the assumption that production functions are Cobb-Douglas ($\rho = 1$), we have $\hat{w}_j^{A,f} = \frac{\sum_s e_{j,s} \mu_s^f}{\sum_s y_{j,s} \mu_s^f} \approx 1$. At this point, thus, it does not appear that, conditional on observed trade flows, allowing for standard Heckscher-Ohlin forces has large effects on the magnitude of the gains from trade.

An attractive feature of multi-factor gravity models is that they provide a theoretical framework to explore quantitatively the distributional consequences of globalization. Using Equations (34) and (35), one could easily compute the change in the skill premium associated with international trade. One caveat, however, is that multi-factor gravity models considered here implicitly rule out differences in factor intensity across firms within the same sector. As discussed in Burstein and Vogel (2010), more productive firms tend to be more skill intensive in practice, which opens up a new channel through which trade liberalization may contribute to an increase in inequality by leading to the exit of the least efficient firms. Burstein and Vogel (2010) find that this "skill-biased" mechanism is quantitatively more important than the Heckscher-Ohlin mechanism. Another caveat is that models in this section abstract from trade in capital goods. Given the existence of capital-skill complementarity, this is another channel through which trade may affect inequality. This issue is explored quantitatively in Burstein, Cravino, and Vogel (2011) and Parro (2012).

3.6 Other Extensions

Non-CES Utility. Although gravity models reviewed so far differ in terms of their supply-side assumptions, they all share the same demand structure. Starting with the Armington model presented in Section 2, all models features representative agents with (nested) CES utility. A non-

---

[22]We have explored the sensitivity of our results to the assumption of Cobb-Douglas technologies for the United States. Following Katz and Murphy (1992), we have set $\rho = 1.4$ rather than $\rho = 1$ and recomputed G_j using Equation (32). The results are basically unchanged.

trivial cost of this assumption is that implies constant markups across firms under monopolistic competition. Hence, the "pro-competitive" effects of trade under this market structure, namely the idea that welfare gains from trade may be larger because of their effects on firm-level markups, are de facto ruled out. In this final subsection, we briefly discuss recent work that has tried to incorporate these considerations into otherwise standard quantitative trade models with monopolistic competition.

The three main alternatives to CES utility in trade models are: (*i*) separable, but non-CES utility functions, as in the pioneering work of Krugman (1979); (*ii*) a quadratic, but non-separable utility function, as in Melitz and Ottaviano (2008); and (*iii*) a translog expenditure function, as in Feenstra (2003). In Arkolakis, Costinot, Donaldson, and Rodriguez-Clare (2012), the authors start from a general demand system that encompasses all three alternatives.

Under the assumptions that there exists a finite reservation price for all varieties; that there are no fixed exporting costs; and that the distribution of firm-level productivity is Pareto, they show that, in spite of variable markups, trade flows satisfy the same gravity equation (16) as in models with CES utility. In this situation, one can use the exact hat algebra to show that the macro-level predictions of non-CES models regarding the consequences of trade liberalization—namely, the predictions regarding the effects of changes in trade costs on wages and trade flows—are the same as in quantitative trade models with CES utility functions discussed in Section 3.1.

Although the predictions regarding trade flows and wages of CES and non-CES model considered in Arkolakis, Costinot, Donaldson, and Rodriguez-Clare (2012) are equivalent, their welfare predictions are not. The gains from trade liberalization, in particular, may be larger or smaller depending on whether they trigger factor reallocations towards goods exhibiting larger or smaller markups, respectively. In the data, the authors find slightly smaller gains from trade liberalization using models with variable markups.

A potentially fruitful alternative to studying the pro-competitive effects of trade consists in departing from monopolistic competition, while maintaining CES. As already discussed in Section 3.1, the model with Bertrand competition developed by Bernard, Eaton, Jensen, and Kortum (2003) leads to the exact same predictions as one-sector gravity models with perfect and monopolistic competition, though alternative assumptions about the distribution of firm-level productivity may lead to different results; see de Blas and Russ (2010) and Holmes, Hsu, and Lee (2010). In recent work, Edmond, Midrigan, and Xu (2011) have used the model with Cournot competition developed by Atkeson and Burstein (2008) to study the magnitude of the gains from trade in economies with variable markups. When calibrating this model using data on manufacturing Taiwanese firms, they find large gains from trade, though the numbers vary depending on assumptions made on the correlation of productivity between Taiwanese and non-Taiwanese firms.

Multinational Production. In all the models reviewed so far, firms serve foreign markets exclusively through exports. In practice, however, a large share of foreign sales are done by foreign affiliates of multinational firms.

Ramondo and Rodríguez-Clare (2013) extend the perfectly competitive framework of Eaton

and Kortum (2002) so that the technological know-how from a given country can be used to produce elsewhere, albeit perhaps at a cost. The cost of using local technologies abroad limits the extent of multinational production in the same way as iceberg trade costs limits the extent of international trade in gravity models. They explore how the substitutability and complementarity forces between trade and multinational production affect the magnitude of the gains from trade and the gains from openness more broadly defined. Their calibrated model implies that the gains from trade can be twice as high as the gains calculated in trade-only models.[23]

Arkolakis, Ramondo, Rodriguez-Clare, and Yeaple (2012) extend the monopolistically competitive framework of Melitz (2003) to allow firms to serve foreign markets by exporting from their home country or by setting up foreign affiliates. In this environment, a gravity equation akin to Equation (5) still holds for the sales by firms from a particular origin, i.e., for the sales in country j of goods produced in country i by affiliates of firms from country l. Given data on bilateral trade and multinational production flows (i.e., the value of production by firms from i in country l, independently of where the goods are sold), they show how their model can be used to conduct counterfactual experiments using the exact hat algebra of Section 2.2.

4   Evaluating Trade Policy

In this section we go from changes in iceberg trade costs to changes in trade policy, namely import tariffs. For expositional purposes, we first go back to the simple Armington model presented in Section 2 and describe the welfare implications of tariff changes in that environment. Like in Section 3, we then study how these predictions vary across different gravity models. We conclude by discussing issues of aggregation in the presence of heterogeneous tariffs across sectors. Throughout this section we modify our WIOD data such that, in line with the static models that we study, overall trade is balanced country-by-country; see online Appendix for details. We tackle the issue of trade imbalances explicitly in Section 5.1.

4.1   Back to Armington

Consider a world economy similar to the one presented in Section 2.1. Compared to our earlier analysis, trade flows may now be subject to import tariffs so that Equation (3) generalizes to

$$P_{ij} = Y_i \tau_{ij} \left(1 + t_{ij} \right) / Q_i, \tag{36}$$

where Y_i denotes factor income in country i, i.e. GDP net of tariff revenues, and $t_{ij} \geq 0$ denotes the ad-valorem tariff imposed by country j on goods from country i. In the rest of this section, we let $\phi_{ij} \equiv \tau_{ij} \left(1 + t_{ij} \right)$ denote the total trade costs between country i and j. Given CES utility, the value

---

[23]In related work, Ramondo (2012) adapts the framework of Eaton and Kortum (2002) to environments with multinational production but no trade. She derives a gravity equation for multinational production flows and computes the associated gains. Burstein and Monge-Naranjo (2009), Prescott and McGrattan (2010), and Garetto (2012) also quantify the gains from multinational production, albeit in models without a gravity equation for multinational production.

of bilateral trade flows (inclusive of tariffs) are thus given by the following gravity equation:

$$X_{ij} = \frac{\left(Y_i \phi_{ij}\right)^{-\varepsilon} \chi_{ij}}{\sum_{l=1}^{n} \left(Y_l \phi_{lj}\right)^{-\varepsilon} \chi_{lj}} E_j. \tag{37}$$

In the presence of import tariffs, budget balance now requires $E_j = Y_j + T_j$, where $T_j \equiv \sum_{i=1}^{n} \frac{t_{ij}}{1+t_{ij}} X_{ij}$ denotes total tariff revenues in country j, whereas the good market clearing condition requires $Y_i = \sum_{j=1}^{n} \frac{1}{1+t_{ij}} X_{ij}$. Together with Equation (37), these two conditions lead to the following generalization of Equation (6):

$$Y_i = \sum_{j=1}^{n} \frac{\left(Y_i \phi_{ij}\right)^{-\varepsilon} \chi_{ij}}{\sum_{l=1}^{n} \left(Y_l \phi_{lj}\right)^{-\varepsilon} \chi_{lj}} \frac{Y_j}{1 - \pi_j}, \tag{38}$$

where $\pi_j \equiv \sum_{i=1}^{n} \frac{t_{ij}}{1+t_{ij}} \frac{\left(Y_i \phi_{ij}\right)^{-\varepsilon} \chi_{ij}}{\sum_{l=1}^{n} \left(Y_l \phi_{lj}\right)^{-\varepsilon} \chi_{lj}} \in (0,1)$ denotes the share of tariff revenues in country j's total expenditure. This completes the description of a competitive equilibrium with import tariffs.

Now consider an arbitrary change in import tariffs from $\mathbf{t} \equiv \{t_{ij}\}$ to $\mathbf{t'} \equiv \{t'_{ij}\}$. To compute proportional changes in factor income, $\hat{\mathbf{Y}} \equiv \{\hat{Y}_i\}$, we can again use the exact hat algebra. Following the same steps as in Section 2.2, we get

$$\hat{Y}_j Y_j = \sum_{i=1}^{n} \frac{\lambda_{ji} \left(\hat{Y}_j \hat{\phi}_{ji}\right)^{-\varepsilon}}{\sum_{l=1}^{n} \lambda_{li} \left(\hat{Y}_l \hat{\phi}_{li}\right)^{-\varepsilon}} \frac{\hat{Y}_i Y_i}{1 - \pi'_i}, \tag{39}$$

where the share of tariff revenues in the counterfactual equilibrium is itself given by

$$\pi'_i = \sum_{j=1}^{n} \frac{t'_{ji}}{1 + t'_{ji}} \frac{\lambda_{ji} \left(\hat{Y}_j \hat{\phi}_{ji}\right)^{-\varepsilon}}{\sum_{l=1}^{n} \lambda_{li} \left(\hat{Y}_l \hat{\phi}_{li}\right)^{-\varepsilon}}. \tag{40}$$

Combining the two previous expressions we can solve for $\hat{\mathbf{Y}} \equiv \{\hat{Y}_i\}$ (up to a normalization). Although the previous system of equations is not quite as compact as Equation (11), it still does not depend directly on preference shifters, endowments, or trade costs. All we need to determine changes in factor income levels, $\hat{Y}_i$, are the initial expenditure shares, λ_{ij}, the initial factor income levels, Y_i, and the trade elasticity, ε. Like in Section 2.2, once changes in tariff income are known, we can compute changes in expenditure shares using the gravity equation (37),

$$\hat{\lambda}_{ij} = \frac{\left(\hat{Y}_i \hat{\phi}_{ij}\right)^{-\varepsilon}}{\sum_{l=1}^{n} \lambda_{lj} \left(\hat{Y}_l \hat{\phi}_{lj}\right)^{-\varepsilon}}.$$

Finally, to compute the welfare change caused by the change in import tariff, we can start from

27

Equation (9). Integrating and taking into account the fact that $E_j = Y_j / (1 - \pi_j)$, we obtain

$$\hat{C}_j = \left(\frac{1 - \pi_j}{1 - \pi'_j} \right) \hat{\lambda}_{jj}^{-1/\varepsilon}. \tag{41}$$

where the share of tariff revenues in the initial and counterfactual equilibria are given by $\pi_j = \sum_{i=1}^{n} \frac{t_{ij}}{1+t_{ij}} \lambda_{ij}$ and $\pi'_j = \sum_{i=1}^{n} \frac{t'_{ij}}{1+t'_{ij}} \lambda_{ij} \hat{\lambda}_{ij}$, respectively.[24] Like in Section 2.2, welfare changes can be computed using only a few sufficient statistics. One does not need to estimate all structural parameters of the model to estimate the welfare effect of an arbitrary tariff change.

We start by considering the welfare effects of a change in tariffs imposed by a single country, i.e., a unilateral change in tariffs. For simplicity, we assume that import tariffs are equal to zero in the initial equilibrium, $t_{ij} = 0$ for all i and j. This is not a bad approximation for most OECD countries today, where tariff revenues are a negligible share of GDP.[25] We consider a counterfactual equilibrium in which a single country j imposes a uniform tariff on all its trading partners, i.e. $t'_{ij} = t > 0$ for all $i \neq j$. All other structural parameters are held fixed between the initial equilibrium and the counterfactual equilibrium.

In Figure 1, we plot the welfare change associated with a unilateral tariff of t% for $j =$United States, France, Portugal, and Ireland under the assumption that the trade elasticity ε is equal to 5. For each counterfactual exercise, we first compute proportional income changes using Equation (39) and (40) and then, for each country j, welfare changes using Equation (41).

A few features of Figure 1 stand out. First, we see that for all countries, the optimal tariff is around 20%. This should not be surprising. In the two-country case, we know from the work of Gros (1987)—see also Helpman and Krugman (1989), Chapter 7—that the optimal tariff in country j is equal to $t_j^* = 1/ \left(\varepsilon \lambda_{jj}^* \right)$, where λ_{jj}^* denotes the share of expenditure in the rest of the world on goods produced in the rest of the world. For all countries, one should expect λ_{jj}^* to be very close to one in practice. So $1/\varepsilon = 20\%$ provides a good approximation of what the optimal tariff should be. Second, we see that the potential gains from trade protection are modest, but nontrivial. At the optimal tariff, they range from 0.3% for the United States to 1.3% for Ireland. To put these numbers in perspective, note that the overall gains from trade predicted by the Armington model for these two countries are 1.8% and 8%, respectively; see Table 1 Column 1. Third, smaller countries that gain more from trade also gain more from optimal trade protection. Here, although smaller countries have less ability to manipulate their terms-of-trade, they benefit disproportionately more from an improvement in their terms-of-trade.[26] Finally, we see that the range over

---

[24]Equation (41) implies that, with tariffs, the gains from trade are given by $G_j = 1 - \left(1 - \pi_j \right) \lambda_{jj}^{1/\varepsilon}$. This generalizes Equation (12) to environments in which the share of tariff revenues in total expenditure, π_j, is non-zero. See Felbermayr, Jung, and Larch (2013) for a discussion of the role of tariff revenues on the gains from trade.

[25]For example, in 2005 revenues as a share of GDP were 0.25%, 0.20%, and 0.36% for Canada, the United States and Mexico, respectively. Developed countries in general have tariff revenue shares of 0.2% or less. Tariff revenues are more important for developing countries. For example, they were 1.8% of GDP in India in 2004. Other prominent examples of high tariff revenues are Russia (3.16%) and China (0.84%), both in the year 2000 – before they joined the WTO.

[26]In general, one would expect gains from trade protection to have an inverted-U shape as a function of country size, since gains from applying an optimal tariff must go to zero as countries become either infinitesimally small or infinitely
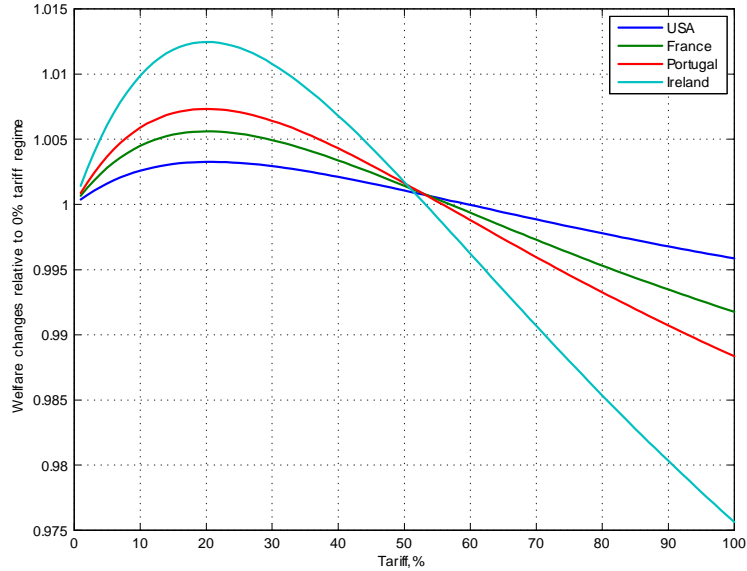
28

Figure 1: Welfare changes associated with a unilateral tariff in the country imposing the tariff. Trade elasticity $\varepsilon = 5$. Data are from WIOD in 2008.

which trade protection is welfare-improving is large. For all four countries, import tariffs up to 50% are found to increase real consumption relative to free trade.

Up to this point, we have only focused on welfare changes in the country imposing the import tariff. We now turn to the effects on the rest of the world. We restrict ourselves to a counterfactual scenario in which the United States unilaterally imposes an import tariff of 40% on all its trading partners, $t'_{iUS} = 0.4$ for all $i \neq$ United States. This is close to the tariff level observed in the United States in the late 19th and early 20th century; see e.g. Irwin (1998). The welfare changes caused by the 40% tariff in each trading partner of the United States are reported in Column 1 of Table 2. The striking feature of these numbers is that in spite of the fact that the United States is the largest country in the world (in terms of GDP) and that 40% is a large tariff by historical standards, the impact on the rest of the world is small. On average, real consumption goes down by 0.2%. The only two exceptions are Canada and Mexico, which incur loses of 1.2% and 1.1%, respectively.

The previous counterfactual exercise may mask the true cost of trade protection if other countries retaliate by increasing their own import tariffs. A very crude way to assess the quantitative importance of trade wars is to consider a counterfactual scenario in which all countries, rather than one, impose a 40% import tariff on all their trading partners. The associated welfare changes are reported in Column 5 of Table 2. On average, countries experience losses of 2.3%. The biggest loser is Ireland (-4.4%) while the country that loses the least is the United States (-0.8%). In this particular counterfactual scenario, no country wins the trade war, though the largest country in the world is the one that loses the least.

While the previous numbers are illustrative, a more satisfactory way to study the consequences

---

large. Figure 1 suggests that countries in our dataset are sufficiently large that they are on the declining segment of this curve.

of trade wars is to solve for the Nash equilibrium of a non-cooperative tariff game. As demonstrated in Ossa (2011b), one can still do so using the exact hat algebra. The only difference between this analysis and the one presented in this section is that proportional changes in trade costs, $\hat{\phi}_{ij}$, are no longer exogenously given, but correspond to the best-response of each country to the vector of import tariffs chosen by its trading partner. Using a multi-sector gravity model with monopolistic competition and estimates of trade elasticities from Broda and Weinstein (2006), Ossa (2011b) finds Nash tariffs around 60% around the world. Ossa finds that the associated welfare loss associated with going from the current, observed equilibrium to the counterfactual Nash equilibrium is equal to 3.5% on average.

4.2 Alternative Gravity Models Revisited

In the previous subsection we have focused on the predictions of a simple Armington model. We now go back the main gravity models considered in Sections 3.2, 3.3, and 3.4 to see how multiple sectors, intermediate goods, and monopolistic competition affect these predictions. For expositional purposes, we restrict ourselves to counterfactual scenarios in which tariffs are common across sectors. We discuss the specific issues associated with heterogeneity in tariffs in the next subsection.

We start by allowing for multiple sectors and intermediate goods, in the exact same way as in Sections 3.3 and 3.4, while retaining the assumption of perfect competition, $\delta_s = 0$ for all s. As in Section 4.1, we assume that all tariffs are zero in the initial equilibrium. Columns 2 and 4 in Table 2 report the welfare effects of a unilateral 40% tariff imposed by the United States on all its trading partner in all sectors. Column 2 corresponds to the predictions of the multi-sector model without intermediate goods, $\alpha_{j,s} = 0$, whereas Column 4 corresponds to the predictions with intermediate goods, $\alpha_{j,s} > 0$.[27] Columns 6 and 7 show the analogous results when the 40% tariff is imposed uniformly by all countries around the world. For both counterfactual scenarios, welfare changes are computed using the exact-hat algebra as we did in the case of the simple Armington model.

Moving from a one-sector to a multi-sector model has ambiguous effects on the welfare losses from a tariff; for many countries we actually see smaller losses. For example, the welfare losses in China from the 40% worldwide tariff fall from 1.6% in the one-sector model (Column 5) to 0.4% in the multi-sector model (Column 6). On average, welfare losses go from 2.3% (Column 5) to 1.4% (Column 6). Although we know from the results of Section 3.3 that a prohibitive tariff should lead to higher welfare losses when there are multiple-sectors, we see that even for tariffs as high as 40% worldwide, this is not the case. Thus one should be careful when extrapolating from the autarky exercises of Section 3 to richer comparative static exercises. Models that point towards larger gains from trade liberalization for one counterfactual scenario may very well lead to smaller gains from trade liberalization for another.

In contrast, allowing for intermediate goods significantly amplifies the losses generated by

---

[27]In the latter case, shares of intermediate goods are computed as $\alpha_{j,ks}^* = \left(\sum_k \sum_i X_{ij,ks} / R_{j,s} \right) \times \left(E_{j,k} / E_j \right)$.

tariffs for all countries, echoing the results of Section 3.4. Canada now experiences a welfare loss of 2.2% (Column 4) rather than 1.2% (Column 2) as a consequence of the 40% tariff imposed by the United States, while average losses are now 0.33% (Column 4) rather than 0.14% (Column 2). A worldwide tariff of 40% now leads to losses in Belgium of 6.8% (Column 7) rather than 2.63% (Column 6). On average, allowing for intermediate goods leads to welfare losses of 3.5% (Column 7) rather than 1.4% (Column 6).

We next study how market structure affects the welfare effects of a worldwide tariffs of 40%. As explained in Section 3, moving from perfect to monopolistic competition introduces scale effects, which complicates the computation of counterfactual equilibria with tariffs. To ease the computational burden, we use a coarser version of our dataset in all the exercises that follow. Specifically, we aggregate the set of countries in the WIOD into 10 regions and 16 sectors (15 tradable sectors and 1 non-tradable sector). Using this coarser dataset, we compare the predictions of gravity models under the assumption of: (*i*) perfect competition; (*ii*) monopolistic competition without firm-level heterogeneity, as in Krugman (1980); and (*iii*) monopolistic competition with firm-level heterogeneity, as in Melitz (2003).[28] In the case of models with monopolistic competition with firm-level heterogeneity, counterfactual predictions depend on whether fixed exporting costs are paid in the importing country or exporting country and whether import tariffs are imposed before or after firm-level markups. In what follows we assume that fixed exporting costs are paid in the exporting country, $c_{ij}^x = Y_i$, and that tariffs are imposed before markups; see online Appendix for details.[29]

Columns 1 to 3 of Table 3 present results for the model without intermediates, as in Section 3.3, while Columns 4 to 6 present results for the model with intermediates, as in Section 3.4. Under perfect competition without intermediate goods (Column 1 in Table 3), the welfare losses associated with a 40% worldwide tariff are lower than before (Column 6 in Table 2). This reflects the fact that, by construction, tariffs are now forced to be zero within regions. Column 2 of Table 3 shows the analogous results under monopolistic competition without firm-level heterogeneity. As in Section 3.3, the differences between perfect and monopolistic competition arise because of sector-level scale effects triggered by trade protection. All regions now have larger losses from the tariffs, except Eastern Europe and the Rest of the World. Finally, Column 3 of Table 3 shows the results under monopolistic competition with firm-level heterogeneity. We see that all regions have slightly lower welfare losses than in the absence of firm-level heterogeneity. One potential explanation is that the introduction of tariffs leads to an increase in the ratio of expenditures—which include tariff revenues—to factor prices. This tends to increase the set of varieties available

---

[28]For computational reasons, we assume that the petroleum sector ($s = 8$) is perfectly competitve in all three cases. Formally, in case (*ii*) we set $\eta_s = 0$ for all s and $\delta_s = 1$ for all $s \neq 8$ and $\delta_s = 0$ for $s = 8$, whereas in case (*iii*) we set $\eta_s = 0.65$ for all s and $\delta_s = 1$ for all $s \neq 8$ and $\delta_s = 0$ for $s = 8$.

[29]When tariffs are imposed before markups, they act as "cost-shifters." Under this assumption, the elasticity of bilateral trade flows with respect to tariffs and iceberg trade costs coincide. When tariffs are imposed after markups, they act as "demand-shifters." Under this assumption, the elasticity with respect to tariffs is equal to $\varepsilon + \eta$ rather than ε; see online Appendix and Felbermayr, Jung, and Larch (2013). Since there is little empirical evidence to help us discriminate between these two cases, we prefer working under the assumption that there is only one elasticity.

to consumers when a tariff is imposed in models with firm-level heterogeneity through selection effects. Since consumers love variety, this tends to reduce the welfare losses associated with a tariff in such models.

How are the previous results affected by the introduction of intermediate goods? We again find that tariffs lead to larger welfare losses in models with monopolistic competition. Moving from Column 4 to Column 5 in Table 3, the losses from the tariff for Western Europe and North America increase from 1.7% to 10.0% and from 1.2% to 1.7%, respectively. Allowing for firm-level heterogeneity (Column 6 in Table 3) leads to even larger welfare losses. The effect is particularly pronounced for China, whose losses almost triple from 14.5% to 40%. The average welfare losses without China increase from 6.3% to 7.9%. Like in Section 3.4, this reflects stronger scale effects under models of monopolistic competition with firm-level heterogeneity.

In recent work Balistreri, Hillberry, and Rutherford (2011) have compared the predictions of a multi-sector Armington model with intermediate goods to those of a multi-sector Melitz (2003) model with intermediate goods. To compare the two models, they first estimate the structural parameters of the Melitz (2003) model using non-linear least squares and investigate the welfare consequences of a 50% reduction in observed tariffs. They then ask what the welfare consequences of the same tariff change would have been in an Armington model calibrated using the same elasticity of substitution between goods (but different trade elasticities) and the same variable trade costs. They find gains from 50% reduction in tariffs that are four times larger according to the Melitz (2003) model and conclude that it is important to add firm-level heterogeneity to standard quantitative trade models.

The similarity between the quantitative results of Balistreri, Hillberry, and Rutherford (2011) and those presented in Table 3 notwithstanding, there are some important methodological differences. In order to demonstrate the importance of firm-level heterogeneity, Balistreri, Hillberry, and Rutherford (2011) compare Armington to Melitz (2003). Here we compare Krugman (1980) to Melitz (2003). To us the latter comparison is preferable to the former since comparing Armington and Melitz (2003) makes it impossible to separate the role of monopolistic competition from the role of firm-level heterogeneity. The results of Table 3 illustrate the importance of this distinction both qualitatively and quantitatively. When comparing Columns 1 and 3, one would conclude that gains from trade liberalization are larger in models with firm-level heterogeneity. Yet comparing Columns 2 and 3, we see that the average welfare loss would be even higher in the absence of firm-level heterogeneity. Similarly, when comparing Columns 4 and 6, one would conclude that average welfare losses are four times as large. Yet the relevant comparison, Column 5 to Column 6, suggests less dramatic differences, with average welfare losses going from 7.1% to 11.1%.

The more substantial methodological difference comes from how we calibrate the models whose predictions we want to compare. In Balistreri, Hillberry, and Rutherford (2011), different models do not match the same set of moments, including the trade elasticity. In this chapter, they do. Our view is that whether or not one wants to "recalibrate" models depends on why one is interested in comparing the predictions of different models in the first place.

The thought experiment that motivates us can be sketched as follows. Consider two trade economists trying to predict the welfare effects of the same tariff change. Both economists have access to the same trade data in 2008, but the first-one has been trained in the 1980s and thinks that Krugman (1980) is a useful approximation of the world economy, whereas the second has been trained in the 2000s and thinks that Melitz (2003) is. We want to know whether *conditional on the same trade data*, the forecasts of the first and second economists would differ. When comparing one model to another, we therefore recalibrate all structural parameters so that each models fits the same cross-section of bilateral trade flows and trade elasticity.

The thought experiment behind the results of Balistreri, Hillberry, and Rutherford (2011) is implicitly different. The idea is to start with one economist, trained in the 2000s, ask for her forecast of the effects of trade policy in 2008 and then ask whether she would have made a different forecast *in a counterfactual world* without firm-level heterogeneity. This is why Balistreri, Hillberry, and Rutherford (2011) first structurally estimate a Melitz (2003) model and then "turn off" heterogeneity across firms (among other things).

So which of these two thought experiments is more useful? If the question is how the Melitz (2003) model has changed the evaluation of trade policy, we think that the first thought experiment is the appropriate one and that comparing Column 2 to Column 3 or Column 5 to Column 6 in Table 3 is the right way to go. In contrast, if one is interested in decomposing the results of Columns 3 or 6 so that one can quantify the role of heterogeneity for welfare within the Melitz (2003) model, then the second thought experiment is more useful.

4.3 Heterogeneous Tariffs

In practice, tariffs are not uniform across sectors within the same country. This raises a number of questions: Does heterogeneity in tariffs raise the cost of trade protection in the country imposing the tariffs? If so, is this also true in the countries facing the tariffs? In theory, what is the "right" way to measure the restrictiveness of trade policy when trade policy is a multi-dimensional object?

Let us start by focusing on the first two questions. We use the simplest model in which these questions can be addressed: a multi-sector gravity model without intermediate goods and with perfect competition. Thus the assumptions are the same as in Section 3.3 with $\delta_s = 0$ for all s. We again consider the counterfactual scenario in which only one country, the United States, imposes the same tariffs, $t'_{iUS,s} = t_s$, on all its trading partners. But we now allow tariffs to vary across sectors s. Namely, we assume that $t_s = t + \Delta_s$, where $\Delta_s \in \{-a, a\}$ is randomly drawn across sectors. Draws of a and $-a$ have equal probabilities so that the expected value of the tariff is t in each sector. In the simulations below we draw the vector $\Delta \equiv (\Delta_1, \Delta_2, ..., \Delta_S)$ 1,000 times and report the average results for $\hat{C}_j$ across all these draws.

In Columns 2 and 3 of Table 2 we report the welfare change associated with a counterfactual scenario in which $t = 0.4$ and $a = 0$ and $a = 0.2$, respectively. Thus, the expected tariff is 40% in both cases, but there is dispersion in the U.S. tariff across sectors in Column 3 and no dispersion in Column 2. Specifically, in Column 3, some sectors face a 20% import tariffs, whereas others face
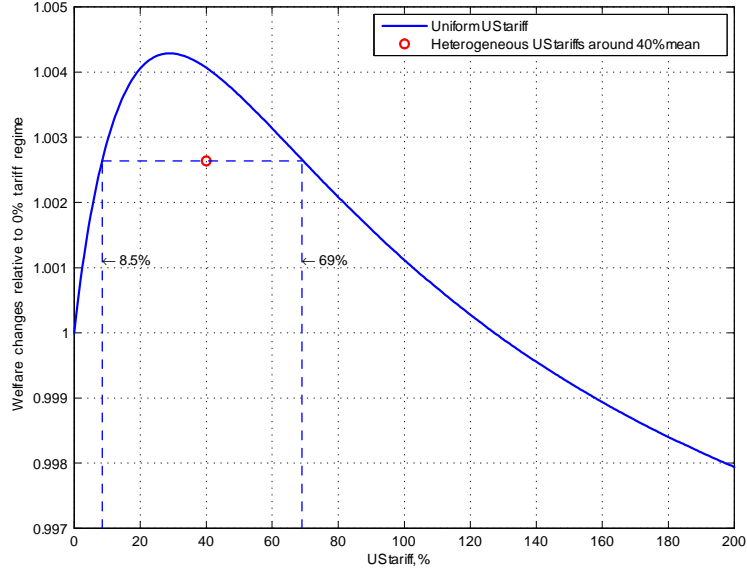
Figure 2: Welfare changes in the United States associated with a unilateral U.S. tariff with and without dispersion across sectors. $\Delta = 0.4$ in the case with dispersion. Sector-level trade elasticities are from Caliendo and Parro (2010). Data are from WIOD in 2008.

a 60% tariff. Comparing Columns 2 and 3 we see that introducing dispersion leads to a decline in the welfare gain from imposing tariffs in the United States from 0.4% to 0.3%. Although there is no systematic relationship between tariff dispersion and welfare, as discussed in Anderson and Neary (2005), these two numbers resonate well with a simple partial equilibrium intuition. If the welfare distortions associated with a tariff are given by the "triangle" below the import demand curve, then holding the average import tariff fixed, an increase in dispersion should tend to lower welfare, which is what happens in this example. By comparing Columns 2 and 3 in Table 2, we also see that other countries are actually better off when the there is dispersion in the US tariff. For example, Canada's welfare loss from the U.S. tariff falls from 1.2% to 1%, while Mexico's losses fall from 0.9% to 0.7%. Intuitively, expenditure functions are concave in prices. So if dispersion in U.S. tariffs also increases dispersion in the price of U.S. goods abroad, this should tend to reduce expenditure and, in turn, increase welfare abroad.

To conclude, let us go back to the issue of how to measure the restrictiveness of trade policy when trade policy is a multi-dimensional object. Anderson and Neary (2005) propose the following two-step approach. First, ask: What is the welfare loss associated with a vector of heterogeneous tariffs in the country imposing those tariffs? Second, compute the uniform tariff across sectors that would lead to the same welfare loss. This number is what Anderson and Neary (2005) refer to as the Trade Restrictiveness Index (TRI). One way to compute this index is to estimate import demand elasticities and use linear approximations; see e.g. Feenstra (1995) and Kee, Nicita, and Olarreaga (2008). An alternative is to use the full structure of the model. We now do so again for the case above with $t = 0.4$ and $a = 0.2$. From Table 2 Column 2, we know that the welfare change in the United States is 0.27%. The Trade Restrictiveness Index can then be read off Figure

34

2, which is the counterpart of Figure 1 in the multi-sector case. In this particular example, there are two values of the TRI: 8.5% and 69%. The non-uniqueness of the TRI reflects the importance of general equilibrium effects in gravity models.

5 Numbers We Can Believe In?

5.1 Sensitivity to Auxiliary Assumptions

In Sections 3 and 4 we have studied how different assumptions about technology and market structure may affect the welfare consequences of trade liberalization. This exercise already gives us a sense of the sensitivity of the counterfactual predictions of gravity models to a set of key economic assumptions. Here we go one step further by exploring how apparently innocuous auxiliary assumptions may also have large implications for counterfactual analysis.

Nature of Trade Imbalances. All models reviewed in this chapter are static models that aim to capture the long-run consequences of trade liberalization. Given their static nature, all these models predict that trade should be balanced country-by-country. In practice, it is not. Volumes of intertemporal trade are substantial: some countries, like the United States, run large current-account deficits, whereas others, like China, run large current-account surpluses. To deal with trade imbalances in the context of a static model, the standard approach consists in modelling trade deficits and surpluses as lump-sum transfers between countries; see e.g. Dekle, Eaton, and Kortum (2008). Short of a dynamic theory of trade imbalances, in which today's trade deficits are paid back by future trade surpluses, the researcher interested in the effects of a change in trade policy has to assume that these lump-sum transfers remain "unchanged" between the initial and the counterfactual equilibrium.

An obvious question when adopting the previous approach is: "unchanged" compared to what? When thinking about the U.S. deficit, for instance, should one hold its value fixed relative to U.S. GDP or World GDP to mention only two of the most natural alternatives? To get a sense of the importance of these considerations, let us go back to the Armington model with import tariffs considered in Section 4.1. In the presence of trade imbalances across countries, the equilibrium condition (38) generalizes to

$$Y_i = \sum_{j=1}^n \frac{\left(Y_i \phi_{ij}\right)^{-\varepsilon} \chi_{ij}}{\sum_{l=1}^n \left(Y_l \phi_{lj}\right)^{-\varepsilon} \chi_{lj}} \frac{Y_j + D_j}{\kappa_j}, \tag{42}$$

where $D_j \equiv E_j - (Y_j + T_j)$ denotes the trade deficit in country j. The two assumptions mentioned above about the nature of trade imbalances correspond to $D_j = \varkappa_j Y_j$ and $D_j = \varkappa_j \sum_{i=1}^n Y_i$, respectively, where $\varkappa_j$ is treated as an exogenous structural parameter that determines the magnitude of lump-sum transfers across countries.

Starting from Equation (42) and some assumption about how trade deficits relate to factor
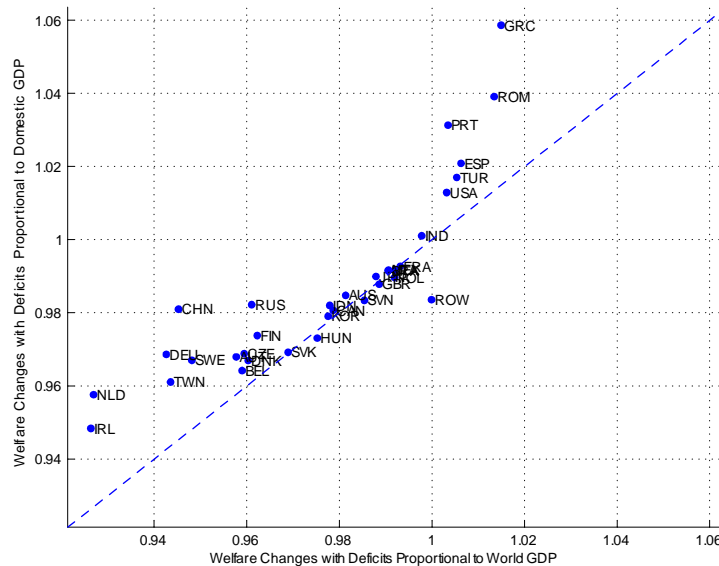
Figure 3: Welfare changes associated with a 40% tariff around the world when deficits are proportional to World GDP (x-axis) and domestic GDP (y-axis). Trade elasticity $\varepsilon = 5$. Data are from WIOD in 2008.

incomes, either $D_j = \varkappa_j Y_j$ or $D_j = \varkappa_j \sum_{i=1}^{n} Y_i$, one can conduct counterfactual analysis using the exact hat algebra, as we did in Section 4.1. In Figure 3, we plot the welfare changes associated with 40% import tariffs imposed around the world under the assumption that $D_j = \varkappa_j Y_j$ (x-axis) and $D_j = \varkappa_j \sum_{i=1}^{n} Y_i$ (y-axis). The correlation between the two measures is equal to 0.86. Not surprisingly, the biggest discrepancies between the two numbers come from countries running large trade deficits, like Greece or Romania, or large trade surpluses, like Ireland and the Netherlands.[30]

It is worth pointing out, however, that both sets of results differ from the numbers presented in Section 4.1 in which we had removed trade imbalances before analyzing the consequences of a worldwide tariff. The correlation between either one of the two measures presented here and the results of Section 4.1 is 0.57. This discrepancy is caused by larger terms-of-trade improvements for countries that maintain large trade deficits. Intuitively, as tariffs increase, consumers have incentives to substitute away from foreign goods towards domestic goods. Thus, in order to maintain trade deficits "unchanged" in the counterfactual equilibrium, domestic prices must increase relative to the price of foreign goods.

Nature of Physical Capital. In earlier sections we have implicitly treated physical capital as a primary factor of production whose endowment is fixed and unaffected by trade. Alternatively, one could think of physical capital as capital goods that, like other intermediate goods, are traded internationally; see e.g. Eaton and Kortum (2001). To get a sense of how important this assumption

---

[30]While this example illustrates the potential sensitivity of quantitative predictions to ad-hoc assumptions about the nature of trade imbalances, it should be clear that the importance of these considerations depends, in general, on the variable of interest, the underlying model, and the nature of the counterfactual exercise. For instance, when the counterfactual exercise consists in moving to autarky under perfect competition, changes in real factor income are unaffected by the size and nature of transfers across countries.

on the nature of physical capital could be, we come back to the model of Section 3.4 with multiple sectors, tradable intermediate goods, and perfect competition. Compared to the analysis of Section 3.4, the only difference is that we adjust value-added shares in the data so that the (presumed) one third share of capital in value added is distributed proportionally among intermediate goods from all industries. Formally, we adjust shares by setting $\alpha'_{i,ks} = \alpha_{i,ks} \left(1 + \frac{1}{3} \frac{1 - \alpha_{i,s}}{\alpha_{i,s}} \right)$. By construction, the new value added share is therefore two thirds of the one reported in the data: $1 - \alpha'_{i,s} = \frac{2}{3} \left(1 - \alpha_{i,s} \right)$. When physical capital is treated as another tradable intermediate goods, the amplification effect from the input-output loop discussed in Section 3.4 is now stronger, and the gains are larger. Using again WIOD data from 2008, we find that average gains from trade are equal to 34.6% compared to 26.9% when physical capital is assumed to be immobile across countries. For example, gains from trade in Canada are now 36.9% rather than 30.2%, and in the United States they are 11.3% rather than 8.3%.

5.2 Goodness of Fit

One way to try to build confidence in the predictions of gravity models is to assess whether they can match cross-sectional or time-series evidence on bilateral trade flows.

Cross-sectional Evidence. A standard practice in the literature consists in assuming first that iceberg trade costs τ are deterministic functions of various observables such as physical distance between country i and country j, whether country i and j share a common border or language, or whether they are part of a regional trade agreement or a monetary union. Once trade cost functions—together with other structural parameters of the model—have been estimated or calibrated, one can assess the goodness of fit of the gravity model by comparing trade flows predicted by the model to those that are observed in the data; see e.g. Di Giovanni and Levchenko (2009), Waugh (2010), and Fieler (2011).

The empirical fit of gravity models tends to be very good; see Anderson (2010). It is not clear, however, whether this should make us very confident in the counterfactual predictions of these models. If iceberg trade costs τ are used as free parameters, then gravity models can replicate any cross-section of bilateral trade flows, X. Thus, one way to interpret the good fit of the gravity models under particular restrictions on the structure of iceberg trade costs is that these restricted trade costs are good proxies of the "wedges" that are necessary for the model to match the data.

An alternative strategy, in the spirit of the test of Heckscher-Ohlin-Theorem in Trefler (1993), is to infer from the trade data the iceberg trade costs such that gravity models fit perfectly and then check whether the inferred iceberg trade costs match observable measures of trade costs such as tariffs and freight costs or price gaps (which should reflect trade costs under perfect competition). An issue with this approach is that one cannot separately identify iceberg trade costs from country-specific biases in preferences in the model. If "wedges" only match observed measures of trade costs very imperfectly, should we conclude that there is a problem with the model or that biases in preferences are important?

In recent work, Anderson and Yotov (2012) propose to test gravity models by focusing on the restrictions that it imposes on the relationship between exporter- and importer-fixed effects in equation (13), whatever the structure of iceberg trade costs may be. Another possible strategy to assess the goodness of fit of quantitative trade models consists in using other, i.e. non-trade, data sources. For instance, since gravity models offer a mapping between trade data and consumer price indices, one could compare the prices predicted by the model to those observed in the data. Waugh (2010) and Fitzgerald (forthcoming) offer two examples of papers exploring the implications of gravity models for prices.

Time Series Evidence. Perhaps the most direct way to test the predictions of a gravity model is to calibrate the model in a baseline year t, consider a change in trade costs between the baseline year t and some other year t', and report the "distance" between the predictions of the model and the data at date t'.

A natural candidate for large changes in trade costs are regional trade agreements like the North American Free Trade Agreement (NAFTA). Caliendo and Parro (2010) use a multi-sector version of Eaton and Kortum (2002) with intermediate goods calibrated to match trade data from 1993 to study the effects of NAFTA. This model is a special case of the models considered in Section 3.4 in that it does not feature any scale effects: $\delta_s = 0$ for all s. After feeding tariff changes caused by NAFTA between 1993 and 2005, they find that the simulated changes can account for 93% of the observed change in Mexico's total trade over GDP, 58% for Canada, and 55% for the United States.

The relatively good fit of their model contrasts with the fairly poor performance of earlier CGE models reviewed in Kehoe (2005). He evaluates the performances of three of the most prominent multisectoral applied general equilibrium models used to predict the impact of NAFTA: the Brown-Deardorff-Stern model of all three North American economies (Brown, Deardorff, and Stern (1992)), the Cox-Harris model of Canada (Cox and Harris (1992)) and the Sobarzo model of Mexico (Sobarzo (1992)). He finds that each of these models drastically underestimated the impact of NAFTA on trade volumes. For instance, the Brown-Deardorff-Stern model predicts changes in exports relative to GDP between 1988 and 1999 equal to 50.8% for Mexico, 4.3% for Canada, and 2.9% for the United States. In the data, these changes are equal 140.6%, 52.9%, and 19.1%, respectively.

It is not entirely clear, however, what accounts for the difference in the performance between more recent and older CGE models. Caliendo and Parro (2010) emphasize the importance of input-output linkages for their results, but such linkages are already present in older CGE models. One economically meaningful difference is that the previous CGE models feature imperfect competition and scale effects, whereas Caliendo and Parro (2010) does not. But there are many other potential sources of discrepancies, from the calibration of the trade elasticities (more on that below) to the choice of the baseline year. Indeed, although the predictions of the models reviewed in Kehoe (2005) are compared with changes in the data over the period 1988-1999, the Brown-Deardorff model was calibrated to a 1976 input-output matrix for Canada, a 1980 input-output

matrix for Mexico, and a 1977 input-output matrix for the United States. The good fit of gravity models in the time-series is also called into question by the results of Lai and Trefler (2002) who find that changes in tariffs between 1972 and 1992 fed into a simple model à la Krugman (1980) can explain little of the variation in bilateral trade flows over this time period.

A poor fit of a quantitative trade model may always reflect the fact that the trade shock under study is small compared to other unobserved shocks—e.g. technology, factor endowments—that may have been occurring over the same time period. This generates a tension. One would want to focus on shorter time periods around a particular episode of trade liberalization to exclude other shocks, but quantitative trade models are static models that are better equipped to capture the long-run consequences of trade liberalization than short-term dynamics. In the case of NAFTA, for instance, would quantitative trade models perform better when the final year is 1995, 2000, or 2005?

Donaldson (2008) offers an innovative way to test the predictions of gravity models. He focuses on one of history's great transportation projects: the network of railroads built in colonial India from 1870 to 1930. He first estimates the impact of the railroad on trade costs using information about price gaps across Indian districts. He then feeds the observed changes in trade costs into a multi-sector version of Eaton and Kortum (2002)—a special case of the models considered in Section 3.3 without scale effects: $\delta_s = 0$ for all s—to predict the changes in real income across Indian districts. The test of the model comes from comparing changes in real income predicted by the model to the reduced-form estimates of changes in real income caused by the railroad. He finds that 86% of the total impact of the railroads on real income in an average district can be explained by the model.

5.3 Calibrating Elasticities

The key parameters for counterfactual analysis using gravity models, like other CGE models, are elasticities. We conclude this section by discussing some of the issues arising when calibrating elasticities.

Trade Elasticities. As can already be seen from the counterfactual analysis carried in the context of the Armington model in Section 2, the single most important structural parameter in gravity models is the trade elasticity. Conditional on observed trade shares, it determines both the response of bilateral trade flows and real consumption. In more general environments such as those considered in Section 3, it remains one of the key statistics required to estimate the gains from international trade; see Equations (23), (27), and (32). So, how large are trade elasticities?

This question is an old one. It is as important for recent gravity models reviewed in this chapter as for earlier CGE models. The broad consensus in the CGE literature is, in the words of Dawkins, Srinivasan, and Whalley (2001), that "the quantity and quality of literature-based elasticity parameters for use in calibrated models is another Achilles' heel of calibration." As John Whalley notes "It is quite extraordinary how little we know about numerical values of elasticities [...] In the international trade area researchers commonly use import price elasticities in the neighborhood of

unity, even for small economies, even though elasticity estimates as high as nine appear in the literature." The same pessimism can be found in the review of trade elasticities by McDaniel and Balistreri (2003): "The estimates from the literature provide a wide range of point estimates, and little guidance on correct estimates to apply to a given commodity in a given model for a given regional aggregation. Most of the controversy surrounding the [trade] elasticities reduces to a general structural inconsistency between the econometric models used to measure the response and the simulation models used to evaluate policy."

Part of the success of new quantitative work in international trade lies in the tight connection between the structural estimation of the trade elasticity and the underlying economic model. In their seminal work, Eaton and Kortum (2002) offer multiple ways to estimate structurally the trade elasticity using a gravity equation akin to Equation (16). Their estimates range from $\varepsilon = 3.60$ to $\varepsilon = 12.86$ with a preferred value of $\varepsilon = 8.28$ when using price gaps as a measure of trade costs. While the range is wide, it remains in the range of elasticities used in earlier CGE models; see Hertel (1997). In recent work, Simonovska and Waugh (2011) refine Eaton and Kortum (2002)'s preferred estimation strategy to take into account the fact that price gaps are only lower-bounds on trade costs. Simonovska and Waugh (2011) propose a simulated method of moments to correct for the fact that trade elasticities using price gaps tend to overestimate the sensitivity of trade flows to trade costs. Their preferred estimate of ε is 4.12.

The merits of a tight connection between theory and data notwithstanding, the state of affairs remains far from ideal. An issue with simulation using earlier CGE models, such as those used in the evaluation of NAFTA, is that they require a very large number of elasticities. Many recent papers, following Eaton and Kortum (2002), side-step this issue by assuming that all goods enter utility functions through a unique CES aggregator. But empirically, there is ample evidence of significant variation in the trade elasticity across sectors; see e.g. Feenstra (1994), Broda and Weinstein (2006), and Hummels and Hillberry (2012). If so, why should we be more confident in the counterfactual predictions of simpler gravity models that abstract from this heterogeneity? Given the heterogeneity in the trade elasticity across sectors, is the trade elasticity estimated from an aggregate gravity equation like (16) the "right" trade elasticity to calibrate a one-sector model?[31]

A natural way to address the previous concerns is to write down multiple-sector models, such as those considered in Section 3.3, and incorporate formally the heterogeneity in trade elasticities across sectors. But this raises new issues. As the number of elasticities that needs to be estimated increases, the precision with which each of those elasticities is estimated tends to decrease. Accordingly, results become much more sensitive to the presence of outliers. To take a concrete example, the sector-level elasticities from Caliendo and Parro (2010) used in Section 3.3 are around 8 on average. But for some sectors, like automobiles, trade elasticities are not statistically different

---

[31]Imbs and Mejean (2011) argue that such aggregate trade elasticities suffers from heterogeneity bias and that one should instead use a properly weighted average of sector-level trade elasticities to calibrate one-sector models. Yet if the true model is indeed a multi-sector model, it is not a priori obvious that there exists a single weighted average— independently of how weights are defined—that can be used to predict both changes in trade flows and welfare.

from zero. An elasticity of zero would imply infinite gains from trade.[32]

Upper-level Elasticities (I): Substitution Across Sectors. Going from a one-sector to a multi-sector model raises another question: How large is the elasticity of substitution between sectors? All papers referenced in Section 3.3 follow what Dawkins, Srinivasan, and Whalley (2001) refer to as the "idiot's law of elasticities": all elasticities are equal to one until shown to be otherwise. How important is the assumption that upper-level utility functions are Cobb-Douglas for the predictions of multi-sector quantitative trade models?

To shed light on this question, let us make the same assumptions as in Section 3.3, expect for the fact the upper-level utility function is now given by

$$C_j = \left(\sum_{s=1}^{S} \beta_{j,s} C_{j,s}^{(\gamma-1)/\gamma} \right)^{\gamma/(\gamma-1)}, \tag{43}$$

where $\gamma > 0$ denotes the upper-level elasticity of substitution between goods from different sectors; $\beta_{j,s} \geq 0$ are exogenous preference parameters, which we normalize such that $\sum_{s=1}^{S} \beta_{j,s} = 1$ for all j; and $C_{j,s}$ still denotes total consumption of the composite good s in country j. The Cobb-Douglas case studied in Section 3.3 corresponds to $\gamma = 1$.

For simplicity let us focus on the case of perfect competition, $\delta_s = 0$ all s. Following a procedure similar to that in Section 3.1, one can show that if sector-level price indices are given by Equation (20), then the welfare impact of a shock generalizes to

$$\hat{C}_j = \left(\sum_{s=1}^{S} e_{j,s} \left(\hat{\lambda}_{jj,s}^{-1} \right)^{(\gamma-1)/\varepsilon_s} \right)^{1/(\gamma-1)}. \tag{44}$$

Gains from trade are thus given by

$$G_j = 1 - \left(\sum_{s=1}^{S} e_{j,s} \left(\lambda_{jj,s} \right)^{(\gamma-1)/\varepsilon_s} \right)^{1/(\gamma-1)}. \tag{45}$$

In Section 3.3, we have pointed out that multi-sector models with Cobb-Douglas preferences predict significantly larger gains than one-sector models. Using Equation (45), we can now quantify the importance of the Cobb-Douglas assumption for this prediction. In Figure 4 we plot G_j as a function of γ using (45) for several countries—the United States, Canada, France, Germany, and Japan—as well as the average G_j for all countries considered in Table 1. We see that the value of the upper-level elasticity γ—for which the existing empirical literature provides little guidance—has large effects on the magnitude of the gains from trade. As we go from the Leontief case, $\gamma = 0$, to the Cobb-Douglas case, $\gamma = 1$, to an upper-level elasticity equal to the average of lower-level elasticities, $\gamma = 8$, average gains from trade decrease from 45% to 15% to 3%.[33] The intuition is simple.

---

[32]The existing literature rarely reports confidence intervals. Two notable exceptions are Lai and Trefler (2002) and Shapiro (2012) who use the standard errors of the trade elasticities that they have estimated to compute, and report, the standard errors for the gains from trade liberalization.

[33]In the particular case in which all lower-level trade elasticities are equal to the upper-level trade elasticity, $\varepsilon_s = \gamma - 1$ for all s, the discrepancy between the predictions of multi-sector and one-sector models disappears: Equation (45)
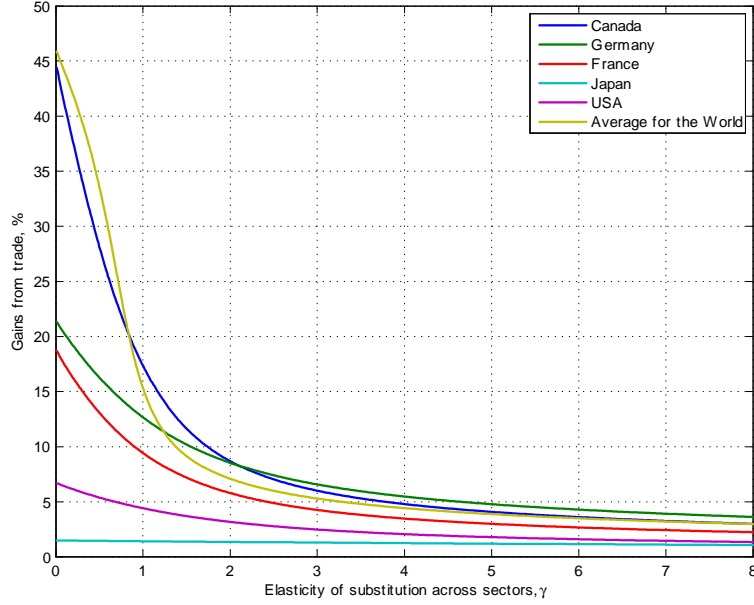
Figure 4: Gains from Trade computed according to Equation (45) for different levels of γ. Sector-level trade elasticities are from Caliendo and Parro (2010). Data are from WIOD in 2008.

If the elasticity of substitution between sectors is high, then the consequences of autarky are mitigated by consumers' ability to substitute consumption away from the most affected sectors,i.e., those with lowest values of $\lambda_{jj,s}^{1/\varepsilon_s}$, towards the least affected sectors, i.e., those with highest values of $\lambda_{jj,s}^{1/\varepsilon_s}$. By the same token, however, the gains from further trade liberalization would tend to be higher with a higher γ, since consumers could more easily reallocate their consumption towards goods that experience larger price declines.[34]

Upper-level Elasticities (II): Domestic versus Foreign. Another, and perhaps deeper issue regarding gravity estimates of the trade elasticity is that they capture the elasticity of substitution between foreign sources of imports. Yet, the elasticity that one needs, for instance, for measuring the gains from trade is the elasticity of substitution between home and import goods. To see this formally, let us go back to the simple Armington model presented in Section 2.1, but let us

reduces to Equation (12).

[34]The value of the upper-level elasticity of substitution also has more subtle implications for quantitative trade models with monopolistic competition. Away from Cobb-Douglas preferences, expenditure shares $e_{j,s}$ are no longer constant, which may lead to multiple equilibria. Intuitively, a higher expenditure share in a given sector lead to more entry, which reduces the price index in that sector, thereby leading to a higher expenditure share. Thus, one may obtain different estimates of the gains from trade—or the effect of trade policy—depending on the counterfactual equilibrium one focuses on. Another issue concerns the difference between the predictions of models with firm-level heterogeneity, like Melitz (2003), and those without, like Krugman (1980). In Section 3.3, we have shown that in the Cobb-Douglas case, entry effects lead to a distinction between the predictions of models with perfect and monopolistic competition, but not between models with and without firm-level heterogeneity. This is no longer true in the case of general CES utility functions.

generalize Equation (1) so that

$$C_j = \left[(C_{jj})^{(\gamma-1)/\gamma} + \left(C_j^M \right)^{(\gamma-1)/\gamma} \right]^{\gamma/(\gamma-1)},$$

where C_j^M measures total consumption of imported goods,

$$C_j^M = \left(\Sigma_{i \neq j} \, \psi_{ij}^{\sigma/(\sigma-1)} C_{ij}^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)}.$$

The upper-level elasticity $\gamma > 1$ now represents the elasticity of substitution between the domestic good and the composite of the foreign goods, whereas the lower-level elasticity $\sigma > 0$ still represents the elasticity of substitution between foreign goods. The simple Armington model corresponds to the special case, $\gamma = \sigma$. Under this new demand system, bilateral trade flows still satisfy a gravity-like equation:

$$X_{ij} = \left(\frac{P_j^M}{P_j} \right)^{1-\gamma} \left(\frac{P_{ij}}{P_j^M} \right)^{1-\sigma} E_j, \text{ for all } i \neq j, \tag{46}$$

where $P_{ij} = \tau_{ij} P_{ii}$ is the price of goods from country i in country j; $P_j^M \equiv \left(\Sigma_{i \neq j} \, \psi_{ij}^{1-\sigma} P_{ij}^{1-\sigma} \right)^{1/(1-\sigma)}$ is the import price index; and $P_j = \left((P_{jj})^{1-\gamma} + \left(P_j^M \right)^{1-\gamma} \right)^{1/(1-\gamma)}$ is the consumer price index in country j. In this more general environment, one can still rearrange bilateral trade flows as we did in Equation (13) and use the cross-sectional variation in trade flows and trade costs to estimate $1 - \sigma$.

Now, like in Section 2.2, consider a small change in trade costs that affects country j. The change in the consumer price index is still given by

$$d \ln P_j = \lambda_{jj} d \ln P_{jj} + \left(1 - \lambda_{jj} \right) d \ln P_j^M.$$

But our new demand system now implies

$$d \ln \left(1 - \lambda_{jj} \right) - d \ln \lambda_{jj} = (1 - \gamma) \left(d \ln P_j^M - d \ln P_{jj} \right).$$

Following the same strategy as in Section 2.2, one can therefore show that

$$d \ln C_j = d \ln \lambda_{jj} / \left(1 - \gamma \right). \tag{47}$$

While gravity estimates can uncover the lower-level elasticity, σ, Equation (47) shows that the upper-level elasticity, γ, i.e., the elasticity of substitution between domestic and foreign goods, is the relevant elasticity for welfare analysis.[35]

---

[35] In more general gravity models, this upper-level elasticity γ would remain a key determinant for welfare analysis,

In standard gravity models, it is only the assumption of symmetric CES utility that allow researchers to go from the commonly estimated elasticity, σ, to the welfare-relevant elasticity, γ. When estimated, does the elasticity of substitution between home and import goods, γ, look similar to the elasticity of substitution between foreign goods, σ? Head and Ries (2001) suggest that the answer is yes. They measure the average of the elasticity of demand for Canadian goods in Canada relative to U.S. goods and the elasticity of demand for U.S. goods in the United States relative to Canadian goods. If all trade was U.S.-Canada trade, their estimate would therefore provide an estimate of γ. They find an average elasticity equal to 7.8, quite in line with previous gravity estimates of σ. Likewise, using the methodology of Feenstra (1994) to estimate both γ and σ, Feenstra, Obstfeld, and Russ (2013) cannot reject the null that γ and σ are equal.

Factor Supply Elasticities. The multi-sector gravity models that we have reviewed assume a perfectly-elastic factor supply to each sector. Thus, except for the case with sector-level differences in factor intensities considered in Section 3.5, the aggregate production possibilities frontier (PPF) is linear. In practice one may expect factors to be imperfect substitutes across sectors. For instance some workers may have a comparative advantage in particular sectors, as in a Roy-type model, or some natural resources may be critical inputs to production in some sectors and not others. Such considerations would lead to more "curvature" in the PPF and, conditional on observed trade flows, larger gains from trade.

To take an extreme example, consider the petroleum sector. The trade elasticity, ε_s, for this sector estimated by Caliendo and Parro (2010) is around 70. The formula presented in Section 3.3 would therefore predict very small gains from trade in this sector. Yet, of course, one would expect many oil importing countries to face enormous losses from moving to autarky. One simple way to capture such considerations would be to go back to the multi-factor model presented in Section 3.5 and assume factors employed in the petroleum sector are specific to that sector, effectively making petroleum an endowment. To explore the quantitative importance of this type of considerations, we have recomputed the gains from trade in the multi-factor model under the assumption that all sectors are perfectly competitive and all factors are sector-specific.[36] In line with the previous discussion, we find larger gains from trade under the assumption that factors are sector-specific, with the cross-country average for G_j going from 15.3% to 17.2%.

Other Elasticities: Love of Variety and Extensive Margin. We conclude by discussing the calibration of the elasticity of substitution, σ, and the extensive margin elasticity, η, introduced in Equation (15) and its sector-level counterparts, Equations (20), (26), and (30). Given estimates of the trade elasticity, ε, these two elasticities are irrelevant for welfare analysis under perfect competition. Under monopolistic competition, however, we have seen that in the presence of intermediate goods, as in Section 3.4, or in multi-sector models with general CES preferences, as discussed above, the values of σ and η do matter above and beyond the value of the trade elasticity, ε. In

___
though not the only one.

[36]Formally, this amounts to using the formula for gains from trade given by Equation(32) under the assumption that $\delta_s = 0$ and $\hat{c}_{j,s}^A = e_{j,s}/r_{j,s}$ for all s.

these richer environments, the predictions of models with and without firm-level heterogeneity are different and the magnitude of the difference crucially depend on how σ and η are calibrated.

As shown in Section 3.2, the three elasticities ε, σ, and η are not independent of one another. In gravity models, ε determines the overall response of trade flows to changes in trade costs, whereas $\sigma - 1$ determine their responses at the intensive margin and $\eta = \left(\frac{\varepsilon - (\sigma - 1)}{\sigma - 1}\right)$ determines their response at the extensive margin, $\varepsilon - (\sigma - 1)$, weighted by the love of variety, $\sigma - 1$. Given an estimate of the trade elasticity, ε, one therefore only needs an estimate of σ to compute η and vice versa. The most direct way to estimate σ or η is to use firm-level trade data; see e.g. Crozet and Koenig (2010) and Eaton, Kortum, and Kramarz (2011). When available, they offer a simple way to estimate the intensive margin elasticity, i.e., by how much the sales of a given set of firms respond to changes in trade costs, and the extensive margin elasticity, i.e., by how much the number of firms responds to changes in trade costs.[37] Eaton, Kortum, and Kramarz (2011) estimate a value of 1.5 for η in a one sector model, whereas Crozet and Koenig (2010) obtain estimates of ε_s and η_s for several sectors. Interestingly, the average η_s across s estimated by Crozet and Koenig (2010) is also 1.5, with little variation across sectors.[38]

Balistreri and Rutherford (2012) nicely illustrates the issues inherent to the calibration of models of monopolistic competition. The authors compare the predictions of a model without firm-heterogeneity, like Krugman (1980), to those of a model with firm-heterogeneity, like Melitz (2003), in a model with identical countries, three sectors, nested CES preferences, but no inter-industry trade, $e_{j,s} = r_{j,s}$. Changes in real consumption are given by Equation (22). Using the fact that $\eta_s = \left(\frac{\varepsilon_s - (\sigma_s - 1)}{\sigma_s - 1}\right)$, one can show that the overall scale effects, $\left(\hat{e}_{j,s}^{\eta_s} \hat{r}_{j,s}\right)^{\delta_s/\varepsilon_s}$, are equal to $\left(\hat{r}_{j,s}\right)^{1/(\sigma_s - 1)}$ in the two models. Thus the only difference between the predictions of the two models comes from the different values of σ_s used in the calibration of the two models. In their calibration, Balistreri and Rutherford (2012) assume that ε_s is equal to 4.6 in both models, but that η_s is equal to 0 in the Krugman-version and 0.65 in the Melitz-version. This implies calibrated values of $\sigma_s = \varepsilon_s / (1 + \eta_s) + 1$ equal to 5.6 in the Krugman-version and 3.8 in the Melitz-version. This leads to stronger love of variety and, in turn, larger entry effects and gains from trade liberalization in the latter model. The question, of course, is whether one should take seriously that love of variety is much stronger than previously thought because the intensive margin is only one particular margin of adjustment of trade flows, $\sigma_s - 1 < \varepsilon_s$, in models with firm-level heterogeneity.

---

[37] A related, though distinct way to estimate the relative importance of the intensive margin, $\sigma - 1$, consists in using the prevalence of zeros in aggregate trade data; see Helpman, Melitz, and Rubinstein (2008). The obvious drawback of this approach is that gravity models, in general, do not allow for zeros, therefore creating a tension between the structural model used for counterfactual analysis and the way its parameters are estimated; see Eaton, Kortum, and Sotelo (2012) for recent work on this topic. Alternatively, one can rely more heavily on the overall structure of the model and use other firm-level measures, like sales and markups, to estimate ε and σ; see e.g. Di Giovanni and Levchenko (2009) and Hsieh and Ossa (2011). Once ε and σ are known, η can then be computed as a residual.

[38] This value is much larger than the value of $\eta_s = 0.65$ used in Section 3.4. If one were to use $\eta_s = 1.5$ rather than 0.65 when computing the gains from trade using Equation (29), one would conclude that gains from trade predicted by models of monopolistic competition with firm-level heterogeneity cannot be finite for any country in our dataset.

6 "Micro" versus "Macro" Numbers

Our chapter so far has described what could be referred to as a "macro" approach to quantifying the consequences of globalization. The distinguishing features of this approach can be summarized as follows. First, it aims to study the overall welfare gains from trade. Second, it is based on commonly available aggregate trade data and a few key structural parameters. Third, it tries to remain agnostic about micro-theoretical foundations. These foundations only matter to the extent that they affect "macro" aggregates, as illustrated, for instance, in Equations (15) and (16).

A common alternative in the literature consists in following a "micro" approach, more closely related to demand and production function estimation in Industrial Organization. Unlike the "macro" approach described in this chapter, the "micro" approach typically focuses on one particular source of gains from trade using micro-level data, either at the product or firm-level. The two most prominent examples of "new" sources of gains from trade emphasized in this literature are: (*i*) gains from new varieties and (*ii*) productivity gains. The broad objective of this section is to discuss how empirical evidence related to (*i*) and (*ii*), i.e., "micro" numbers, relate to the predictions of quantitative trade models reviewed in this chapter, i.e., "macro" numbers.

6.1 Gains From New Varieties

One of the hallmarks of the "new trade theory" pioneered by Paul Krugman in the early 80s is that consumers may gain from trade through access to new varieties. These may be direct gains coming from trade in final goods, as in Krugman (1979) and Krugman (1980), or indirect gains coming from trade in intermediate goods, as in Ethier (1982b). An important contribution of the empirical trade literature in the last decade is to put a number on the "new-goods" channel; see e.g. Klenow and Rodríguez-Clare (1997), Broda and Weinstein (2006), Arkolakis, Demidova, Klenow, and Rodríguez-Clare (2008), Feenstra and Weinstein (2010), and Goldberg, Khandelwal, Pavcnik, and Topalova (2009).

In order to illustrate the similarities and differences with the "macro" approach of this chapter, we will focus on the most influential paper in the "micro" approach: Broda and Weinstein (2006). The starting point of this paper is the observation that statistical agencies typically ignore gains from the creation of new varieties (and destruction of old varieties), when computing various components of the Consumer Price Index (CPI), including the Import Price Index. The authors then go on to ask: How different would measured changes in the Import Price Index of the United States be if one were to take into account the creation and destruction of varieties over time?

To answer this question, the authors first need to define what the counterpart of a variety in theory is in the data and then to specify a demand system to evaluate the welfare impact of new varieties. Broda and Weinstein (2006) define a variety as a 10-digit HTS product from a given country, e.g. "Umbrellas and Sun Umbrellas Having a Telescopic Shaft from China," and assume nested CES utility functions like in the present chapter. Given these assumptions, the contribution of appearing and disappearing varieties to the import component of the CES price

index, i.e., the expenditure function of a representative US consumer, can then be computed using an estimate of the elasticity of substitution across foreign varieties and measures of the changes in the share of expenditures on continuing varieties as in Feenstra (1994). The main finding of Broda and Weinstein (2006) is that, between 1972 and 2001, the decrease in the Import Price Index associated with the expansion of imported varieties, and omitted by statistical agencies, has lead to an additional 2.6% decrease in the U.S. CPI.

It should be clear that in spite of its use of micro-level data, the empirical strategy of Broda and Weinstein (2006) has a lot in common with the approach presented in earlier sections of this chapter. First, the focus is squarely on welfare, i.e. money metric indirect utility function that are constructed by means of the expenditure function. Second, the analysis heavily relies on functional form assumptions, namely CES utility. In particular, the correction term introduced by Feenstra (1994) is closely related to our welfare formula in the simple Armington model; see Equation (9). The basic idea is that under CES utility, changes in relative prices can be inferred from changes in relative expenditure modulo the elasticity of substitution. This is true regardless of whether changes in relative prices are due to movements in the price of existing goods or the creation of new goods, a fact that we have exploited more generally in Section 3.1.

There is, however, a number of notable differences. A benefit of the "micro" approach taken by Broda and Weinstein (2006) is that it only requires assumptions on the demand system, which together with data on prices and quantities, are sufficient to compute changes in the Import Price Index of a representative consumer with nested CES utility. These "micro" numbers are independent of technological considerations and market structure. In contrast, the "macro" numbers presented in this chapter do heavily depend on functional form assumptions on the supply-side of the economy, e.g. Fréchet distributions of productivity under perfect competition or Pareto distributions under monopolistic competition.

The previous observation does not imply that technological considerations and market structure are irrelevant for such empirical micro studies. This is best illustrated by the following thought experiment. Suppose first, as Broda and Weinstein (2006) do in the last part of their analysis, that Krugman (1980) is an accurate description of the world. Then changes in the Import Price Index due to new varieties should be reflected one-for-one in the overall CPI. Under this assumption, "micro" numbers indeed measure additional welfare gains. Now suppose instead that Eaton and Kortum (2002) or some other Ricardian model is an accurate description of the world. There is a priori nothing in the trade data that Broda and Weinstein (2006) look at that would contradict this view of the world. Yet, under this alternative assumption on market structure and technology, changes in the Import Price Index due to the creation of new varieties should be exactly offset by changes in the domestic component of the CPI due to the destruction of old varieties. In this neoclassical world, the set of goods available to consumers is fixed. So the "new-goods" channel is rightly ignored by statistical agencies: in net, it should have no welfare effect. To put it bluntly, "micro" numbers are unaffected by assumptions on technology and market structure, but whether

or not one should care about these numbers crucially depend on these assumptions.[39]

Another important difference between the "micro" approach taken by Broda and Weinstein (2006) and the "macro" approach presented in this paper is that, at its core, the "micro" approach is a measurement exercise, not a counterfactual exercise. The goal is to document what actually happened in the data, e.g. in the United States between 1972 and 2001, not what could have happened if China had grown at a different rate or if the United States had chosen different trade policies. To address such counterfactual questions, one again needs the full structure of the model; specifying and estimating a demand system is not enough.

6.2 Productivity Gains

In the canonical model of trade with firm-level heterogeneity developed by Melitz (2003), trade liberalization causes a reallocation of resources from the least to the most productive firms. This theoretical prediction has received considerable support in the data, with many empirical studies finding very large effects of trade on measured productivity. For instance, Trefler (2004) finds that the reallocation of market shares across firms following the Canada-US free trade agreement has caused productivity, measured as real value added per hour worked, to go up by 8.4% in Canadian manufacturing. Using the methodology of Olley and Pakes (1996) to measure productivity at the firm-level, Pavcnik (2002) documents productivity gains of the same order of magnitude following the trade liberalization of the manufacturing sector in Chile.

The goal of this subsection is to clarify the relationship between productivity studies and the body of work reviewed in this chapter. Given space constraints, we will make no attempt at systematically reviewing this large empirical literature; see e.g. Tybout (2001), Melitz and Trefler (2012), as well as Melitz and Redding (2013) in this volume for excellent overviews. Here we will focus on the following conceptual issues: What is the difference between productivity and welfare? If welfare and measured productivity simultaneously go up, should we think of productivity gains as a new source of gains from trade? At a qualitative level, should we always expect welfare and measured productivity to go in the same direction? At a quantitative level, can we directly compare the "micro" numbers from productivity studies with the "macro" numbers presented earlier in this chapter?

In theory, productivity measures how much "output" one can produce with a given amount of "input," whereas welfare measures how much "consumption" one can afford. At the macro level, the two concepts are therefore closely related. In a closed economy, the more productive a country is, the more it can produce, and so, the more it consumes. This is why real GDP computed on the revenue and expenditure sides coincide. In an open economy, the two concepts differ because of terms-of-trade effects. Absent any shock to a country's productivity, a country may be able to consume more if it can purchase more imports with the same amount of exports, i.e., if its

---

[39]A similar issue arises if one believes that Melitz (2003) rather than Krugman (1980) is an accurate description of the world. In this case, however, changes in the Import Price Index and the domestic component of the CPI due to the creation and destruction of varieties only exactly cancel out under special circumstances, as discussed in footnote 12.

terms-of-trade improve.

An obvious, but important difference between the results in this chapter and productivity studies therefore lies in the treatment of terms-of-trade effects. They are central to our analysis. In the simple Armington model, like in any neoclassical model, terms-of-trade effects are the only channel through which foreign shocks may affect welfare in another country. One of the main goals of quantitative trade models therefore is to predict the terms-of-trade changes associated with particular shocks. In contrast, terms-of-trade effects play no direct role in productivity studies, whose main goal is to identify the consequences of trade liberalization on productivity at the firm- and industry-level.

For the rest of our discussion, let us set aside the issue of terms-of-trade effects. The exact relationship between measured productivity and welfare depends, of course, on how one measures productivity as well as the model used to compute welfare; see Burstein and Cravino (2011) for a general discussion. For our purposes here, let us simply assume that all firm-level data are consistent with Melitz (2003) and that industry-level measures of productivity (in the data) increases whenever the distribution of firm-level productivity (in the model) shifts to the right.[40] Even under such ideal circumstances, changes in measured productivity caused by reallocations across firms and welfare may be quite different, as we now illustrate through a series of short examples.

Consider first a small decrease in variable trade costs. The decrease in variable trade costs, which can be thought of as an outward shift in the production possibility frontier, will lead to a direct first-order increase in welfare. In addition, there will be a reallocation of resources from the least to the most efficient firms, and this will lead to an increase in measured productivity at the industry level. Since the initial equilibrium is efficient under monopolistic competition with CES utility, see e.g. Dhingra and Morrow (2012), this reallocation of resources is associated with positive, albeit second-order, welfare effects. At this point, one might therefore be tempted to conclude that the increase in measured productivity is an additional source of welfare gains. But it should be clear that it is the reallocation rather than the change in productivity per se that has caused the additional welfare gain.

To see this, consider the opposite shock: a small increase in variable trade costs. By the exact same logic, this shock would decrease measured productivity at the industry level through the expansion of the least efficient firms. But as before, the reallocation of resources would remain associated with positive, albeit second-order, welfare effects. That is to say, the first-order welfare loss caused by the increase in trade costs is not as large as it would have been if workers had not been reallocated and measured productivity had remained higher.

A careful reader may have noted that in the two previous examples, there nevertheless exists a perfect correlation between measured productivity and welfare. When trade costs go down, as in our first example, both measured productivity and welfare go up. When trade costs go up, as in

---

[40]This assumption is less innocuous than it may appear. For instance, as Burstein and Cravino (2011) demonstrate, if industry productivity is measured as total value added divided by total employment and the underlying model is Melitz with a Pareto distribution of productivity, then measured productivity would be unaffected by the reallocation of resources induced by trade liberalization.

our second example, the opposite is true. This is fairly intuitive. Any shock that tends to cause an increase in real wages, i.e., welfare, should also tend to cause the exit of the least profitable firms, and in turn, an increase in measured productivity. This suggests that if one were only interested in qualitative questions—did welfare go up or down?—changes in measured productivity may still be used as a useful proxy for welfare changes. As intuitive as it may sound, the monotonic relationship between measured productivity and welfare does hinge on the nature of the shock being considered. This is illustrated by the results of Demidova and Rodríguez-Clare (2009) and Felbermayr, Jung, and Larch (2011) who study the effects of import tariffs and export subsidies in the Melitz (2003) model. Starting from free trade, they show that an import tariff increases welfare, whereas an export subsidy decreases welfare. Yet, an import tariff leads to labor reallocations from high- to low-productivity firms, whereas an export subsidy leads to a reallocation from low- to high-productivity firms. In this situation, changes in measured productivity at the industry-level and welfare move in opposite directions. The reason is that trade policy not only affects output per variety, but also the set of available varieties, which consumers care about.

The previous discussion illustrates that changes in welfare and changes in measured productivity are, in general, qualitatively different. For the purposes of the present chapter, however, the key question is not a qualitative one—"Are we getting the sign right when focusing on measured productivity rather than welfare?"—but a quantitative one—"Even when changes in welfare and measured productivity are of the same sign, are we getting the order of magnitude right?". The answer to the latter question depends on why firms of different productivity coexist in the economy. To take an extreme example, suppose that firms produce the same homogeneous good under constant returns to scale using labor as their only factor production, but that because of taxes or other distortions, the most efficient firms are prevented from taking over the market. In this hypothetical world, changes in measured productivity are trivially equal to changes in consumption, and hence welfare. In this hypothetical world, "micro" and "macro" numbers should coincide.[41] But this hypothetical world is very different from canonical models of trade with firm-level heterogeneity in which love of variety plays a central role. If one believes such models are a reasonable description of the world we live in, one cannot simply celebrate the death of small firms because they raise measured productivity within an industry. These firms may be less productive, but it is efficient for them to produce differentiated varieties in equilibrium. According to such models, "micro" and "macro" numbers are bound to be different and, according to empirical studies reviewed in Section 6.1, the difference between the two may be quite large.

To sum up: changes in measured productivity do not capture changes in welfare, either qualitatively or quantitatively, except in settings that are very different from canonical trade models.[42]

---

[41]This equivalence is subject to the caveat that empirical studies typically identify the productivity changes in industries that experience changes in tariffs relative to those that do not; see e.g.Trefler (2004). If industries that do not experience tariff changes are affected by trade liberalization through general equilibrium effects, "micro" numbers estimated using a difference-in-difference approach and "macro" numbers derived using the full structure of the model would still differ.

[42]Although we have focused on Melitz (2003), which is, by far, the most influential trade model featuring firm-level heterogeneity, the caveats illustrated by our simple examples apply more generally. Consider, for instance, the model

Simply put, "micro" studies estimating the effect of trade shocks on measured productivity and "macro" studies quantifying the effect of trade shocks on welfare are getting at different things. To us, there are costs and benefits in both kinds of studies. Empirical micro studies such as Pavcnik (2002) and Trefler (2004) have the great virtue of letting the data speak. Measured productivity gains are what they are, regardless of what the underlying economic environment may be. However, whether such "gains" are a useful metric for policy evaluation remains open for discussion. In contrast, quantitative macro studies such as those reviewed in this chapter aim to speak directly to welfare considerations, albeit at the cost of stronger assumptions.

7 Life Without Gravity

We have described how gravity models can be used for quantitative analysis. We have shown that even richer versions of these models can be calibrated using only a few elasticities. There is, however, a price to parsimony. All gravity models discussed in this chapter rely on strong functional form assumptions. It is those functional form assumptions that allow researchers to estimate trade elasticities from the cross-section—since elasticities are assumed to be constant across countries— and use those to do counterfactual analysis—since elasticities are also assumed to be constant across equilibria. We conclude by discussing alternative structural approaches to quantifying the consequences of globalization that weaken the need for extrapolation by functional form.

7.1 Sufficient-Statistic Approach

A recent literature in public finance has popularized "the sufficient-statistic approach" to counterfactual and welfare analysis. In the words of Chetty (2009), "the central concept of the sufficient statistic approach is to derive formulas for the welfare consequences of policies that are functions of high-level elasticities rather than deep primitives. Even though there are multiple combinations of primitives that are consistent with the inputs to the formulas, all such combinations have the same welfare implications." While there are some similarities between the approach described in this chapter and the sufficient-statistic approach in the public finance literature—namely, the focus on high-level elasticities rather than the multiple combinations of primitives that are consistent with those—there are also some key differences, which are best illustrated through a simple example.

Consider a small, uniform change in tariffs, $d \ln (1 + t_j)$, in some country j. The sufficient-statistic approach in the sense of Chetty (2009) consists in using envelope conditions from optimization to reduce the set of parameters that need to be identified. In the case of a small open economy j, the assumptions of (i) utility maximization and (ii) profit maximization under perfect

---

of monopolistic competition developed by Melitz and Ottaviano (2008). In this model productivity varies across firms because markups vary. Nevertheless, terms-of-trade effects as well as labor reallocations in and out of the outside sector that affect the set of differentiated varieties available to consumers would still introduce a discrepancy between measured productivity and welfare; see Arkolakis, Costinot, Donaldson, and Rodriguez-Clare (2012).

competition imply, without any functional form assumptions, that the welfare changes associated with a change in tariff are given by:

$$d \ln C_j = \left(1 - \lambda_{jj} \right) \times \varepsilon^{SS} \times d \ln \left(1 + t_j \right), \tag{48}$$

where $\varepsilon^{SS} \equiv d \ln M_j / d \ln \left(1 + t_j \right)$ denotes the total elasticity of imports, M_j, with respect to the tariff change. Since country j is a small open-economy, $d \ln C_j$ simply represents the deadweight loss associated with the tariff, i.e., Harberger's triangle. In the international trade literature, the previous formula is the basis of numerous papers evaluating the effects of trade policy; see e.g. Johnson (1960) and Feenstra (1995).

Like in the analysis of the previous sections, Equation (48) emphasizes a unique upper-level elasticity. It should be clear, however, that ε^{SS} plays a very different role than the trade elasticity ε in our previous analysis. Here, ε^{SS} is necessary to compute the area of the "triangle" below the demand curve. In Sections 2 and 3, we are interested in "rectangles": $d \ln C_j = \left(1 - \lambda_{jj} \right) d \ln \tau_j$ in the case of changes in iceberg trade costs. In this context, ε is necessary instead to map changes in prices into changes in shares of expenditures. Such changes can then be integrated to compute the gains from trade in a straightforward manner.

The second difference between the two approaches comes from the nature of the upper-level elasticity. ε^{SS} is a total elasticity in the sense that it captures all general equilibrium effects associated with the tariff change. In contrast, ε is a partial elasticity: it captures the impact of changes in trade costs or tariffs on trade flows holding all general equilibrium effects fixed; see Lai and Trefler (2002) and Arkolakis, Costinot, and Rodríguez-Clare (2012) for discussions of the distinction between the two. While the recent public finance literature has been very successful in exploiting credible reduced-form estimates of the elasticity of taxable income, credible estimates of the total import elasticity, ε^{SS}, are much more scarce; see Feenstra (1995). In contrast, estimating the trade elasticity ε using a gravity equation is fairly straightforward.

The third difference comes from the importance of price effects. In a closed economy, such as the one considered in Harberger (1964), price changes are a transfer from one group of agents to another. Thus they can be ignored in welfare computations. In an open economy, in contrast, terms-of-trade effects do affect welfare. In the case of a large economy imposing a tariff, Equation (48) generalizes to

$$d \ln C_j = \left(1 - \lambda_{jj} \right) \times \left(\varepsilon^{SS} \times d \ln \left(1 + t_j \right) + d \ln P_j^M \right),$$

where $d \ln P_j^M$ is the change in the relative price of imports, i.e. the term-of-trade of country j. To assess the effect of a change in tariff in country j, one therefore either needs to observe price changes after the shock or a model to predict $d \ln P_j^M$ before the shock, as we did in earlier sections. Alternatively, one can ignore terms-of-trade effects by focusing on welfare changes at the world level; see Atkeson and Burstein (2010), Burstein and Cravino (2011), and Fan, Lai, and Qi (2013).

In this situation, we are back to the closed economy analysis of Harberger (1964).[43]

7.2 Other Structural Approaches

Another strategy, closely related to the sufficient-statistic approach presented above, consists in using optimality conditions to derive upper-bounds on the gains from trade. This is the approach followed by Bernhofen and Brown (2005). They exploit a unique quasi-natural experiment: the opening up of Japan in 1858. Under perfect competition, one can show that the gains from trade are weakly lower than the difference between the vector of consumption in the trade equilibrium and the vector of consumption in the autarky equilibrium, both evaluated at autarky prices, which is itself lower than the opposite of the vector of net exports in the trade equilibrium, also evaluated at autarky prices. Since Bernhofen and Brown (2005) have direct access to autarky prices in Japan in 1858 as well as net exports in subsequent years, they can compute an upper-bound of the gains from trade by assuming perfect competition, and nothing more. They find upper-bounds between 5.4% and 9.1% of Japan's GDP depending on the trade equilibrium they focus on. Unfortunately, autarky prices are rarely available in practice.[44]

Costinot and Donaldson (2011) develop a structural approach that tries instead to weaken the need for strong functional form assumptions by focusing on a very particular sector: agriculture. The basic idea is that agriculture is a sector of the economy in which scientific knowledge of how essential inputs such as water, soil and climatic conditions map into outputs is uniquely well understood. As a result, when studying the consequence of restricting trade, one does not need strong functional form assumptions to extrapolate from "factual" productivity levels, that are currently observed, to "counterfactual" productivity levels, that would be observed in economic activities in which production does not currently take place, but would take place in the absence of trade. Agronomists can provide direct information about "counterfactual" productivity levels. Using very detailed micro-level data from the Food and Agriculture Organization's (FAO) Global Agro-Ecological Zones (GAEZ) project, Costinot and Donaldson (2011) are able to construct the Production Possibility Frontiers (PPF) associated with 1,500 U.S. counties and explore the gains from economic integration between these local markets from 1880 to 2002.

8 Concluding Remarks

The first goal of our chapter was methodological. We have described how to use gravity models to perform welfare and counterfactual analysis. An appealing feature of this approach, which we have referred to as the exact hat algebra, is that the impact of various counterfactual scenarios can be computed without estimating the full structure of the model. All the relevant information about

---

[43]Atkeson and Burstein (2010) and Burstein and Cravino (2011) both feature models with monopolistic competition and CES utility. In this case, envelope conditions from optimization can still be invoked since the solution of the planning problem and the decentralized equilibrium coincide; see Dhingra and Morrow (2012).

[44]Irwin (2005) is another exception. He focuses on the Jeffersonian embargo from 1807 to 1809. He finds the cost of autarky, i.e. the gains from trade for the United States over that time period, to be around 5% of GDP.

preferences, technology, and trade costs can be inferred directly from the cross-section of bilateral trade flows and estimates of the trade elasticity. This is a flexible approach that requires only commonly available data, that can be applied to answer a wide range of counterfactual questions, and that can be generalized to study a rich set of economic environments.

The second goal of our chapter was to use this methodology to explore quantitatively how various economic considerations—market structure, firm-level heterogeneity, multiple sectors, intermediate goods, and multiple factors of production—affect the magnitude of the gains from trade as well as the effects of trade policy. What are the main lessons?

Out of the various economic channels that we study, multiple sectors and tradable intermediate goods appear to have the largest effects on the magnitude of the gains from trade. For example, the gains from trade for the United States increase from 1.8% to 8.3% simply by moving from the one-sector model to the model with perfect competition with multiple sectors and intermediate goods. Market structure and firm-level heterogeneity also matter, though their implications are more subtle. For instance, in the presence of multiple sectors, gains from trade may be higher or lower under monopolist competition depending on whether countries specialize in sectors characterized by larger or lower scale effects. Conditional on the same observed trade flows, firm-level heterogeneity may also lead to much larger gains through its indirect implications for the magnitude of scale effects in the presence of intermediate goods. Given the sensitivity of the predictions of gravity models with monopolistic competition to scale effects, more direct evidence on their magnitude would be very valuable.

Regarding the effects of trade policy, gravity models predict welfare gains from unilateral import tariffs over a surprisingly large range. In the one-sector Armington model, the unilaterally optimal tariff is well approximated by the inverse of the trade elasticity. Thus, a trade elasticity of 5 implies optimal tariffs of 20% around the world. In turn, it takes import tariffs to be as high as 50% to get back to the welfare levels observed under free trade. This range gets even larger once we move from a one-sector to a multi-sector model. In addition, the overall welfare effects of large unilateral tariffs on other countries appear to be minimal. Are these numbers we can believe in? Is there something that is present in the data and absent from a baseline gravity model that would dramatically affect these numbers?

The third goal of our chapter was to discuss the costs and benefits of using gravity models to quantify the consequences of globalization. The main benefit of gravity models is parsimony. They rely on a few key structural parameters, first and foremost the trade elasticity, that one can try to estimate credibly in order to make counterfactual and welfare predictions. Parsimony, however, comes at the cost of strong functional form assumptions. How restrictive are those? One would like to see more research regarding the fit of gravity models. The fact these models can perfectly fit any cross-section of bilateral trade flows certainly does not imply that out-of-sample their fit will be perfect as well. If so, what are the dimensions in which gravity models succeed and those in which they fail? How do their predictions compare to reduced-form evidence on the gains from openness? Would alternative models—or alternative ways of estimating the same models—

lead to different conclusions? In the short-run, can we think of the consequences of globalization without recognizing that there may be important frictions in factor markets? In the long-run, can we think of the consequences globalization without knowing how it affects growth? These are all exciting opportunities for future research.

References

ALVAREZ, F., AND R. E. LUCAS (2007): "General Equilibrium Analysis of the Eaton-Kortum Model of International Trade," *Journal of Monetary Economics*, 54(6), 1726–1768.

ANDERSON, J. E. (1979): "A Theoretical Foundation for the Gravity Equation," *The American Economic Review*, 69(1), 106–116.

——— (2010): "The Gravity Model," *forth. Annual Review of Economics*.

ANDERSON, J. E., AND J. P. NEARY (2005): *Measuring the Restrictiveness of International Trade Policy*. MIT Press.

ANDERSON, J. E., AND E. VAN WINCOOP (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *The American Economic Review*, 93(1), 170–192.

——— (2004): "Trade Costs," *Journal of Economic Literature*, 42(3), 691–751.

ANDERSON, J. E., AND Y. V. YOTOV (2010): "The Changing Incidence of Geography," *American Economic Review*, 100(5), 2157–2186.

——— (2012): "Gold Standard Gravity," *NBER Working Paper*.

ARKOLAKIS, C. (2010): "Market Penetration Costs and the New Consumers Margin in International Trade," *Journal of Political Economy*, 118(6), 1151–1199.

ARKOLAKIS, C., A. COSTINOT, D. DONALDSON, AND A. RODRIGUEZ-CLARE (2012): "The Elusive Pro-Competitive Effects of Trade," *mimeo*.

ARKOLAKIS, C., A. COSTINOT, AND A. RODRÍGUEZ-CLARE (2012): "New Trade Models, Same Old Gains?," *American Economic Review*, 102(1), 94–130.

ARKOLAKIS, C., S. DEMIDOVA, P. J. KLENOW, AND A. RODRÍGUEZ-CLARE (2008): "Endogenous Variety and the Gains from Trade," *American Economic Review, Papers and Proceedings*, 98(4), 444–450.

ARKOLAKIS, C., N. RAMONDO, A. RODRIGUEZ-CLARE, AND S. YEAPLE (2012): "Innovation and Production in the Global Economy," *mimeo*.

ATKESON, A., AND A. BURSTEIN (2008): "Pricing to Market, Trade Costs, and International Relative Prices," *American Economic Review*, 98(5), 1998–2031.

——— (2010): "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, 118(3), 433–489.

BALDWIN, R. E., AND A. J. VENABLES (1995): *Regional economic integration*, vol. 3 of *Handbook of International Economics*. Elsevier.

BALISTRERI, E., R. HILLBERRY, AND T. RUTHERFORD (2011): "Structural Estimation and Solution of International Trade Model with Heterogeneous Firms," *Journal of International Economics*, 83(1), 95–108.

BALISTRERI, E., AND T. RUTHERFORD (2012): *Computing General Equilibrium Theories of Monopolistic Competition and Heterogeneous Firms*, Handbook of Computable General Equilibrium Modeling. Elsevier.

BALISTRERI, E. J., R. H. HILLBERRY, AND T. F. RUTHERFORD (2009): "Trade and Welfare: Does Industrial Organization Matter," *Economics Letters*, 109(2), 85–87, Manuscript, University of Melbourne.

BERMAN, E., J. BOUND, AND Z. GRILICHES (1994): "Changes in Demand for Skilled Labor within US Manufacturing: Evidence from the Annual Survey of Manufactures," *Quarterly Journal of Economics*, 109(2), 367–397.

BERNARD, A. B., J. EATON, J. B. JENSEN, AND S. KORTUM (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93(4), 1268–1290.

BERNARD, A. B., S. J. REDDING, AND P. SCHOTT (2007): "Comparative Advantage and Heterogeneous Firms," *Review of Economic Studies*, 74(1), 31–66.

BERNHOFEN, D. M., AND J. C. BROWN (2005): "An Empirical Assessment of the Comparative Advantage Gains from Trade: Evidence from Japan," *American Economic Review*, 95(1), 208–225.

BRODA, C., AND D. WEINSTEIN (2006): "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 121(2), 541–585.

BROWN, D., A. DEARDORFF, AND R. STERN (1992): "A North American Free Trade Agreement: Analytic Issues and a Computational Assessment," *The World Economy*, 15, 15–29.

BURSTEIN, A., AND J. CRAVINO (2011): "Measured Aggregate Gains from International Trade," *mimeo*.

BURSTEIN, A., J. CRAVINO, AND J. VOGEL (2011): "Importing Skill-Biased Technology," *mimeo*.

BURSTEIN, A., AND A. MONGE-NARANJO (2009): "Foreign Know-How, Firm Control, and the Income of Developing Countries," *Quarterly Journal of Economics*, 124(1), 149–195.

BURSTEIN, A., AND J. VOGEL (2010): "Globalization, Technology, and the Skill Premium," Manuscript, Columbia University and UCLA.

CALIENDO, L., AND F. PARRO (2010): "Estimates of the Trade and Welfare Effects of NAFTA," Manuscript, University of Chicago and Yale University.

CHANEY, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *The American Economic Review*, 98(4), 1707–1721.

CHETTY, R. (2009): "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 1, 451–488.

CHOR, D. (2010): "Unpacking Sources of Comparative Advantage: A Quantitative Approach," *forth. Journal of International Economics*.

COSTINOT, A., AND D. DONALDSON (2011): "How Large Are the Gains from Economic Integration? Theory and Evidence from U.S. Agriculture, 1880-2002," *mimeo MIT*.

COSTINOT, A., D. DONALDSON, AND I. KOMUNJER (2010): "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," Forthcoming, Review of Economic Studies.

COX, D., AND R. HARRIS (1992): "North American Free Trade and Its Implications for Canada: Results from a CGE Model of North American Trade," *The World Economy*, 15, 31–44.

CROZET, M., AND P. KOENIG (2010): "Structural gravity equation with extensive and intensive margins," *Canadian Journal of Economics*, 43(1).

DAWKINS, C., T. SRINIVASAN, AND J. WHALLEY (2001): *Calibration*, vol. 5 of *Handbook of Econometrics*. Elsevier.

DE BLAS, B., AND K. N. RUSS (2010): "Teams of Rivals: Endogenous Markups in a Ricardian World," *NBER Working Paper*, 16587.

DEKLE, R., J. EATON, AND S. KORTUM (2008): "Global Rebalancing with Gravity: Measuring the Burden of Adjustment," *IMF Staff Papers*, 55(3), 511–540.

DEMIDOVA, S., AND A. RODRÍGUEZ-CLARE (2009): "Trade policy under firm-level heterogeneity in a small economy," *Journal of International Economics*, 78(1), 100–112.

DHINGRA, S., AND J. MORROW (2012): "The Impact of Integration on Productivity and Welfare Distortions Under Monopolistic Competition," *mimeo, LSE*.

DI GIOVANNI, J., AND A. A. LEVCHENKO (2009): "Firm Entry, Trade, and Welfare in Zipf's World," Manuscript, University of Michigan.

DONALDSON, D. (2008): "Railroads of the Raj: Estimating the Economic Impact of Transportation Infrastructure," Manuscript, London School of Economics.

EATON, J., AND S. KORTUM (2001): "Trade in Capital Goods," *European Economic Review*, 45(7), 1195–1235.

——— (2002): "Technology, Geography and Trade," *Econometrica*, 70(5), 1741–1779.

EATON, J., S. KORTUM, AND F. KRAMARZ (2011): "An Anatomy of International Trade: Evidence from French Firms," *Econometrica*, 79(5), 1453–1498.

EATON, J., S. KORTUM, AND S. SOTELO (2012): "International Trade: Linking Micro and Macro," *NBER working paper 17864*.

EDMOND, C., V. MIDRIGAN, AND D. XU (2011): "Competition, Markups and the Gains from Trade," *mimeo, Duke and NYU*.

ETHIER, W. J. (1982a): "Decreasing Costs in International Trade and Frank Graham's Argument for Protection," *Econometrica*, 50(5), 1243–1268.

——— (1982b): "National and International Returns to Scale in the Modern Theory of International Trade," *American Economic Review*, 72(3), 389–405.

FAN, H., E. LAI, AND H. QI (2013): "Global Gains from Reduction of Trade Costs," *Unpublished*.

FEENSTRA, R. C. (1994): "New Product Varieties and the Measurement of International Prices," *The American Economic Review*, 84(1), 157–177.

——— (1995): *Estimating the Effects of Trade Policy*, vol. 3 of *Handbook of International Economics*. Elsevier.

——— (2003): "A Homothetic Utility Function for Monopolistic Competition Models, Without Constant Price Elasticity," *Economics Letters*, 78(1), 79–86.

——— (2010): "Measuring the Gains from Trade under Monopolistic Competition," *Canadian Journal of Economics*, pp. 1–28, Manuscript, University of California, Davis.

FEENSTRA, R. C., M. OBSTFELD, AND K. N. RUSS (2013): "In Search of the Armington Elasticity," *mimeo*.

FEENSTRA, R. C., AND D. WEINSTEIN (2010): "Globalization, Markups, and the U.S. Price Level," *NBER working paper*, 15749.

FELBERMAYR, G., B. JUNG, AND M. LARCH (2013): "Icebergs versus Tariffs: A Quantitative Perspective on the Gains from Trade," *Unpublished*.

FELBERMAYR, G. J., B. JUNG, AND M. LARCH (2011): "Optimal Tariffs, Retaliation and the Welfare Loss from Tariff Wars in the Melitz Model," *CESifo Working Paper Series*.

FEYRER, J. (2009a): "Distance, Trade, and Income – The 1967 to 1975 Closing of the Suez Canal as a Natural Experiment," *NBER Working Paper*.

——— (2009b): "Trade and Income – Exploiting Time Series in Geography," *NBER Working Paper*, 14910.

FIELER, A. C. (2011): "Nonhomotheticity and Bilateral Trade: Evidence and a Quantitative Explanation," *Econometrica*, 79(4), 1069–1101.

FITZGERALD, D. (forthcoming): "Trade Costs, Asset Market Frictions, and Risk Sharing," *American Economic Review*.

FRANKEL, J. A., AND D. ROMER (1999): "Does Trade Cause Growth," *The American Economic Review*, 89(3), 379–399.

GARETTO, S. (2012): "Input Sourcing and Multinational Production," *mimeo Boston University*.

GOLDBERG, P., A. KHANDELWAL, N. PAVCNIK, AND P. TOPALOVA (2009): "Trade Liberalization and New Imported Inputs," *American Economic Review, Papers and Proceedings*, 99(2), 494–500.

GROS, D. (1987): "A Note on the Optimal Tariff, Retaliation and the Welfare Loss from Tariff Wars in a Framework with Intra-Industry Trade," *Journal of International Economics*, 23, 357–367.

HARBERGER, A. C. (1964): "The Measurement of Waste," *American Economic Review*, 54(3), 58–76.

HARRISON, A. E., AND A. RODRÍGUEZ-CLARE (2010): *Trade, Foreign Investment, and Industrial Policy for Developing Countries*, vol. 5 of *Handbook of Development Economics*. Elsevier.

HEAD, K., AND T. MAYER (2013): *Gravity Equations: Toolkit, Cookbook, Workhorse*, vol. 4 of *Handbook of International Economics*. Elsevier.

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

HELPMAN, E., AND P. KRUGMAN (1985): *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy*. MIT Press, Cambridge, Massachusetts.

HELPMAN, E., AND P. R. KRUGMAN (1989): *Trade Policy and Market Structure*. MIT Press, Cambridge, Massachussetts.

HELPMAN, E., M. MELITZ, AND Y. RUBINSTEIN (2008): "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 2(5), 441–487.

HERTEL, T. (1997): *Global Trade Analysis: Modeling and Applications*. Cambridge University Press, Cambridge.

HOLMES, T. J., W.-T. HSU, AND S. LEE (2010): "Plants and Productivity in Regional Agglomeration," *mimeo, University of Minnesota*.

HSIEH, C.-T., AND R. OSSA (2011): "A Global View of Productivity Growth in China and India," *NBER Working Paper*.

HUMMELS, D., AND R. HILLBERRY (2012): *Trade Elasticity Parameters for a CGE Model*, Handbook of Computable General Equilibrium Modeling. Elsevier.

IMBS, J., AND I. MEJEAN (2011): "Elasticity Optimism," *mimeo*.

IRWIN, D. A. (1998): "Changes in U.S. Tariffs: The Role of Import Prices and Commercial Policies," *American Economic Review*, 88(4), 1015–1026.

———— (2005): "The Welfare Costs of Autarky: Evidence from the Jeffersonian Embargo, 1807-1809," *Review of International Economics*, 13, 631–645.

JOHNSON, H. (1960): "The Cost of Protection and the Scientific Tariff," *Journal of Political Economy*, 68, 327–345.

KATZ, L., AND K. MURPHY (1992): "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 107(1), 35–78.

KEE, H. L., A. NICITA, AND M. OLARREAGA (2008): "Import Demand Elasticities and Trade Distortions," *Review of Economics and Statistics*, 90(4), 666–682.

KEHOE, T. J. (2005): "An Evaluation of the Performance of Applied General Equilibrium Models of the Impact of NAFTA," in *Frontiers in Applied General Equilibrium Modeling*, ed. by T. J. Kehoe, T. Srinivasan, and J. Whalley, pp. 341–377. Cambridge University Press, New York.

KLENOW, P. J., AND A. RODRÍGUEZ-CLARE (1997): "Quantifying Variety Gains from Trade Liberalization," Manuscript, University of Chicago.

KRUGMAN, P. (1979): "Increasing Returns Monopolistic Competition and International Trade," *Journal of International Economics*, 9(4), 469–479.

KRUGMAN, P. (1980): "Scale Economies, Product Differentiation, and the Pattern of Trade," *The American Economic Review*, 70(5), 950–959.

———— (1995): *Increasing Returns, Imperfect Competition, and the Positive Theory of International Trade*, vol. 3 of *Handbook of International Economics*. Elsevier.

KRUGMAN, P., AND A. J. VENABLES (1995): "Globalization and the Inequality of Nations," *Quarterly Journal of Economics*, 110(4), 857–880.

LAI, H., AND D. TREFLER (2002): "The Gains from Trade with Monopolistic Competition: Specification, Estimation, and Mis-Specification," *NBER Working Paper*.

LEAMER, E. E., AND J. E. LEVINSOHN (1995): *International Trade Theory: The Evidence*, vol. 3 of *Handbook of International Economics*. Elsevier.

LEVCHENKO, A. A., AND J. ZHANG (2011): "The Evolution of Comparative Advantage: Measurement and Welfare Implications," *mimeo University of Michigan*.

MCDANIEL, C., AND E. J. BALISTRERI (2003): "A review of Armington Trade Substitution Elasticities," *Integration and Trade*, 18(7), 161–173.

MELITZ, M., AND S. REDDING (2013): *Heterogeneous Firms and Trade*, vol. 4 of *Handbook of International Economics*. Elsevier.

MELITZ, M., AND D. TREFLER (2012): "Gains from Trade when Firms Matter," *Journal of Economic Perspectives*, 26(2).

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

MELITZ, M. J., AND G. I. P. OTTAVIANO (2008): "Market Size, Trade, and Productivity," *The Review of Economic Studies*, 75(1), 295–316.

OLLEY, S., AND A. PAKES (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64(6), 1263–1298.

OSSA, R. (2011a): "A "New Trade" Theory of GATT/WTO Negotiations," *Journal of Political Economy*, 119(1), 112–152.

——— (2011b): "Trade Wars and Trade Talks with Data," *NBER Working Paper*.

——— (2012): "Why Trade Matters After All," *mimeo*.

PARRO, F. (2012): "Capital-Skill Complementarity and the Skill Premium in a Quantitative Model of Trade," *mimeo*.

PAVCNIK, N. (2002): "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants," *The Review of Economic Studies*, 69(1), 245–276.

PRESCOTT, E. C., AND E. MCGRATTAN (2010): "Technology Capital and the US Current Account," *American Economic Review*, 100(4), 1493–1522.

RAMONDO, N. (2012): "A Quantitative Approach to Multinational Production," *mimeo*.

RAMONDO, N., AND A. RODRÍGUEZ-CLARE (2013): "Trade, Multinational Production, and the Gains from Openness," *Journal of Political Economy*, 121(2), 273–322.

RUTHERFORD, T. F. (2002): "Lecture Notes on Constant Elasticity Functions," *mimeo University of Colorado*.

SAMUELSON, P. A. (1939): "The Gains from International Trade," *Canadian Journal of Economics*, 5(2), 195–205.

SHAPIRO, J. (2012): "Trade, CO2, and the Environment," *mimeo MIT*.

SHIKHER, S. (forthcominga): "Predicting the Effects of NAFTA: Now we can do it better!," *Economic Inquiry*.

———— (forthcomingb): "Putting Industries into the Eaton-Kortum Model," *Journal of International Trade and Economic Development*.

SIMONOVSKA, I., AND M. WAUGH (2011): "The Elasticity of Trade: Estimates and Evidence," *mimeo*.

———— (2012): "Different Trade Models, Different Trade Elasticities?," *mimeo*.

SOBARZO, H. (1992): "A General Equilibrium Analysis of the Gains from NAFTA for the Mexican Economy," *The World Economy*, 15, 83–100.

TIMMER, M. (2012): "The World Input-Output Database (WIOD): Contents, Sources and Methods," *WIOD Working Paper*.

TINBERGEN, J. (1962): *Shaping the World Economy*. New York: The Twentieth Century Fund.

TREFLER, D. (1993): "International Factor Price Differences: Leontief Was Right!," *Journal of Political Economy*, 101(6), 961–987.

———— (2004): "The Long and Short of the Canada-US Free Trade Agreement," *American Economic Review*, 94(4), 870–895.

TYBOUT, J. R. (2001): "Plant- and Firm-Level Evidence on "New" Trade Theories," *NBER Working Paper*, 8418.

WAUGH, M. (2010): "International Trade and Income Differences," *American Economic Review*, 100(5), 2093–2124.

	One Sector (12)	Multiple Sectors, No Intermediates (23)		Multiple Sectors, with Intermediates (29)			
		Perfect Competition	Monopolistic Competition	Perfect Competition (data alphas)	Perfect Competition	Monop. Comp. (Krugman)	Monop. Comp. (Melitz)
Country	1	2	3	4	5	6	7
AUS	2.3%	8.6%	3.7%	15.8%	15.7%	6.9%	6.8%
AUT	5.7%	30.3%	30.5%	49.5%	49.0%	57.6%	64.3%
BEL	7.5%	32.7%	32.4%	54.6%	54.2%	63.0%	70.9%
BRA	1.5%	3.7%	4.3%	6.3%	6.4%	9.7%	12.7%
CAN	3.8%	17.4%	15.3%	30.2%	29.5%	33.0%	39.8%
CHN	2.6%	4.0%	4.0%	11.5%	11.2%	28.0%	77.9%
CZE	6.0%	16.8%	21.2%	34.0%	37.2%	65.1%	86.7%
DEU	4.5%	12.7%	17.6%	21.3%	22.5%	41.4%	52.9%
DNK	5.8%	30.2%	24.8%	41.4%	45.0%	42.0%	44.8%
ESP	3.1%	9.0%	9.5%	18.3%	17.5%	24.4%	30.5%
FIN	4.4%	11.1%	10.5%	20.2%	20.3%	24.2%	28.0%
FRA	3.0%	9.4%	11.1%	17.2%	16.8%	25.8%	32.1%
GBR	3.2%	12.9%	11.7%	21.6%	22.4%	22.2%	23.5%
GRC	4.2%	16.3%	4.7%	23.7%	24.7%	6.8%	6.1%
HUN	8.1%	29.8%	31.3%	53.5%	55.3%	75.7%	91.0%
IDN	2.9%	5.5%	4.0%	13.1%	11.6%	11.2%	14.6%
IND	2.4%	4.6%	4.3%	9.2%	8.6%	9.5%	11.7%
IRL	8.0%	23.5%	14.2%	37.1%	38.9%	28.1%	29.1%
ITA	2.9%	8.7%	9.2%	16.4%	16.2%	21.7%	26.5%
JPN	1.7%	1.4%	3.7%	4.6%	3.5%	20.7%	32.7%
KOR	4.3%	3.9%	8.6%	12.5%	11.4%	44.1%	70.2%
MEX	3.3%	11.1%	12.1%	18.4%	18.6%	24.3%	28.4%
NLD	6.2%	24.3%	23.1%	40.1%	39.8%	43.4%	47.6%
POL	4.4%	18.4%	19.7%	33.8%	34.5%	46.9%	57.0%
PRT	4.4%	23.8%	20.6%	35.9%	37.4%	36.7%	40.3%
ROM	4.5%	17.7%	12.7%	26.4%	29.2%	20.8%	20.7%
RUS	2.4%	18.0%	0.9%	35.9%	30.7%	-2.1%	-7.1%
SVK	7.6%	22.2%	23.6%	48.3%	50.5%	78.6%	96.4%
SVN	6.8%	39.6%	39.3%	57.8%	61.6%	71.3%	79.7%
SWE	5.1%	12.5%	14.5%	24.4%	23.6%	36.6%	45.8%
TUR	2.9%	11.9%	13.3%	20.0%	20.9%	26.4%	29.5%
TWN	6.1%	9.6%	9.9%	19.9%	19.4%	28.6%	37.8%
USA	1.8%	4.4%	3.8%	8.3%	8.0%	8.6%	10.3%
RoW	5.2%	15.2%	7.3%	33.3%	28.4%	18.1%	21.8%
Average	4.4%	15.3%	14.0%	26.9%	27.1%	32.3%	40.0%

Gj expressed in percentages computed using:

Table 1: The numbers in parenthesis indicate the equation used for the computation. All data is from WIOD and trade elasticities are from Caliendo and Parro (2010). Perfect competition and monopolistic competition are obtained from the formulas using $\delta_s = 0$ for all s and $\delta_s = 1$ for all s, respectively. Results for the Krugman and Melitz models are obtained setting $\eta_s = 0$ for all s and setting $\eta_s = 0.65$ for all s, respectively.

	Welfare Effect of Tariffs under Perfect Competition						
	Unilateral US 40% Tariff				Uniform Worldwide 40% Tariff		
	One Sector	Multiple Sectors			One Sector	Multiple Sectors	
Country		Without Intermediates	Without Intermediates, With Dispersion	With Intermediates		Without Intermediates	With Intermediates
	1	2	3	4	5	6	7
AUS	-0.10%	-0.13%	-0.11%	-0.23%	-1.26%	-1.38%	-2.82%
AUT	-0.09%	-0.06%	-0.05%	0.01%	-2.98%	-2.04%	-4.42%
BEL	-0.16%	-0.12%	-0.10%	-0.18%	-3.96%	-2.63%	-6.85%
BRA	-0.10%	-0.08%	-0.07%	-0.14%	-0.81%	-0.43%	-0.85%
CAN	-1.20%	-1.16%	-0.97%	-2.21%	-2.06%	-2.14%	-4.20%
CHN	-0.22%	-0.14%	-0.12%	-0.16%	-1.56%	-0.43%	-1.98%
CZE	-0.05%	-0.03%	-0.02%	0.10%	-3.16%	-1.34%	-4.75%
DEU	-0.16%	-0.10%	-0.08%	0.15%	-2.48%	-0.74%	-1.63%
DNK	-0.20%	-0.09%	-0.08%	-0.15%	-3.04%	-1.32%	-3.78%
ESP	-0.06%	-0.04%	-0.03%	-0.33%	-1.47%	-0.71%	-2.24%
FIN	-0.09%	-0.04%	-0.03%	0.01%	-2.36%	-0.94%	-2.80%
FRA	-0.09%	-0.07%	-0.05%	-0.19%	-1.51%	-0.60%	-1.58%
GBR	-0.16%	-0.15%	-0.13%	-0.33%	-1.66%	-1.50%	-3.28%
GRC	-0.08%	-0.02%	-0.01%	-0.60%	-1.84%	-1.65%	-3.76%
HUN	-0.13%	-0.06%	-0.05%	-0.26%	-4.19%	-2.54%	-7.75%
IDN	-0.09%	-0.06%	-0.05%	-0.11%	-1.56%	-0.82%	-2.34%
IND	-0.16%	-0.13%	-0.11%	-0.34%	-1.17%	-0.71%	-1.78%
IRL	-0.91%	-0.56%	-0.48%	-1.22%	-4.41%	-2.17%	-6.61%
ITA	-0.07%	-0.03%	-0.03%	-0.12%	-1.47%	-0.46%	-1.44%
JPN	-0.11%	-0.06%	-0.05%	-0.07%	-0.92%	0.24%	0.06%
KOR	-0.21%	-0.14%	-0.12%	-0.27%	-2.31%	0.22%	-1.14%
MEX	-1.08%	-0.87%	-0.73%	-1.72%	-1.74%	-1.11%	-2.48%
NLD	-0.22%	-0.16%	-0.13%	-0.06%	-3.33%	-1.71%	-3.99%
POL	-0.04%	-0.03%	-0.03%	-0.17%	-2.21%	-1.28%	-3.47%
PRT	-0.06%	-0.05%	-0.04%	-0.46%	-2.13%	-1.85%	-4.25%
ROM	-0.03%	0.00%	0.01%	-0.48%	-2.08%	-2.15%	-5.25%
RUS	-0.03%	-0.06%	-0.04%	0.03%	-1.30%	-2.84%	-4.83%
SVK	-0.05%	-0.01%	-0.01%	0.00%	-3.97%	-2.52%	-6.88%
SVN	-0.06%	-0.04%	-0.03%	-0.22%	-3.50%	-2.44%	-6.31%
SWE	-0.15%	-0.08%	-0.07%	0.01%	-2.71%	-1.23%	-3.14%
TUR	-0.03%	-0.01%	-0.01%	-0.24%	-1.34%	-0.45%	-1.54%
TWN	-0.46%	-0.34%	-0.29%	-0.52%	-3.40%	-1.85%	-5.03%
USA	0.21%	0.41%	0.27%	0.43%	-0.80%	-0.44%	-1.17%
ROW	-0.49%	-0.43%	-0.37%	-1.14%	-2.69%	-2.45%	-6.09%
Average	-0.20%	-0.14%	-0.12%	-0.33%	-2.27%	-1.37%	-3.54%

Table 2: Counterfactual results are computed using the exact hat algebra. All data is from WIOD and trade elasticities are from Caliendo and Parro (2010). See online Appendix for details.

	Welfare Effect of a 40% Worldwide Tariff					
	Without intermediates			With intermediates		
	Perfect Competition	Monopolistic Competition		Perfect Competition	Monopolistic Competition	
Region		Krugman	Melitz		Krugman	Melitz
	1	2	3	4	5	6
Pacific Ocean	-0.1%	-1.7%	-1.7%	-0.6%	-6.4%	-6.6%
Western Europe	-0.8%	-3.6%	-3.5%	-1.7%	-10.0%	-13.3%
Eastern Europe	-2.2%	-2.1%	-1.9%	-4.5%	-7.2%	-10.3%
Latin America	-0.7%	-2.3%	-2.2%	-1.5%	-5.2%	-6.5%
North America	-0.6%	-0.7%	-0.7%	-1.2%	-1.7%	-2.0%
China	-0.7%	-1.7%	-1.6%	-2.6%	-14.5%	-40.0%
Southern Europe	-0.8%	-2.5%	-2.5%	-1.8%	-6.8%	-8.4%
Northern Europe	-1.6%	-3.5%	-3.3%	-3.3%	-8.6%	-9.5%
Indian Ocean	-0.8%	-1.3%	-1.2%	-1.9%	-4.0%	-5.7%
RoW	-3.2%	-2.1%	-1.8%	-6.6%	-6.7%	-8.9%
Average	-1.2%	-2.1%	-2.0%	-2.6%	-7.1%	-11.1%

Table 3: Counterfactual results are computed using the exact hat algebra. All data is from WIOD and trade elasticities are from Caliendo and Parro (2010). See online Appendix for details.