20/6/16

Optimal trade policy with monopolistic competition and heterogeneous firms

Jan I. Haaland
Norwegian School of Economics

Anthony J. Venables
University of Oxford

Abstract

This paper derives optimal trade and domestic taxes for a small open economy containing a monopolistically competitive (MC) sector in which firms may have heterogeneous productivity levels. Analysis encompasses cases in which the domestic MC sector is able to expand or contract flexibly, or is constrained to be of fixed size. In the former case domestic protection can bring gains by increasing the number of product varieties on offer; these gains (and the corresponding rates of domestic subsidy or of import tariffs) are reduced by heterogeneity of foreign exporters some of whom may withdraw from the market. In the latter case gains from protection arise from terms-of-trade effects; since various margins of substitution are switched off, only the relative values of domestic taxes, import tariffs and export taxes matter. In general, policies work through both a terms-of-trade and a variety effect, and the paper shows how the relative importance of each depends on the structure of the economy.

Keywords: trade policy; monopolistic competition; heterogeneous firms; terms of trade; variety; productivity.
JEL classification: F12, F13

Acknowledgements: Thanks to Swati Dhingra, Andres Rodriguez-Clare, referees, an editor and participants in seminars in Bari, Bergen, ETSG and Oxford, for comments. This paper is a revised version of CEPR Discussion Paper no. 10219, October 2014 and NHH Discussion Paper SAM 33 2014.

Authors' Addresses:
J.I. Haaland
Department of Economics,
NHH
5045 Bergen, Norway
Jan.haaland@nhh.no

Corresponding author
A.J. Venables
Department of Economics
Manor Road
Oxford OX1 3UQ UK
tony.venables@economics.ox.ac.uk

1. Introduction

What combination of domestic and trade taxes maximises welfare in an open economy? The answer to this turns on two sorts of reason for intervention; one is to manipulate the terms of trade and the other is to mitigate distortions in the economy. The optimal tax and tariff structure depends on the extent to which policy can operate on each of these margins. We revisit this question in an economy in which some activity takes place in a sector with monopolistic competition and in which firms may differ in their productivity. Our objective is to better understand the welfare economics of models with these features by identifying sources of inefficiency in the market equilibrium and the tax policies that correct them.

To achieve this objective we work with a relatively general model structure that encompasses and extends a substantial part of the existing literature. We derive explicit formulae for optimal and constrained optimal taxes on domestic sales, imports, and exports, both when the monopolistically competitive (MC) sector contains symmetric firms (Krugman 1980), and when firms have different productivity levels (Melitz 2003). This distinction matters for the responses of firms in the MC sector to policy. Optimal policy also depends on the general equilibrium response of the economy, in particular the elasticity with which labour (the only sectorally mobile factor) can be shifted between the MC sector and a perfectly competitive (PC) sector. We derive results for intermediate values of this elasticity as well as for the polar cases that are in the literature, i.e. perfectly elastic supply (Krugman 1980, Venables 1982, Flam and Helpman 1987, Baldwin and Forslid 2010, Ossa 2011, Campolmi et al. 2014, Bagwell and Staiger 2015), or perfectly inelastic supply (as in the one-sector economy of Demidova and Rodriguez-Clare 2009). The effects described in this literature operate through different channels, but ultimate welfare change comes from impacts on just two sources of welfare gain – the terms of trade and domestic distortions. By comparing cases – MC with and without heterogeneous firms, alternative specifications of the general equilibrium responses of the economy, and different combinations of admissible policy instruments – we gain considerable insight into what drives results and why they differ across cases.

Throughout the paper our focus is on unilateral policy action by a small open economy. The small open economy assumption states that foreign wages, foreign price indices, and the number of foreign firms are constant. The border prices of imports are therefore constant, although the terms of trade are influenced by domestic policy since the export prices of domestic firms are variable. However, while domestic policy does not influence the number of firms operating in foreign economies, it does (if firms are heterogeneous) influence the number of these firms that choose to export to the domestic economy (as in Demidova and Rodriguez-Clare 2009). This, we think, is quite a good characterisation of the policy margins faced by many countries. Policy changes the number of firms that engage in trade – both the number of domestic firms that export, and the number of foreign firms that select to supply the domestic market. For many countries and many markets this seems to be a realistic approach; foreign activities directly related to the country in question are affected by the country's choices, while foreign macro variables are not.

What are our findings? Welfare gains from policy derive from two opportunities that are not fully exploited by the market. The monopolistic competition (MC) distortion means that the market under-supplies varieties because firms (that are not perfectly discriminating monopolists) are unable to capture the entire consumer surplus associated with a new variety.[1] The possibility of altering the terms of trade (ToT) arises because, while import prices are fixed by the small open economy assumption, export prices are endogenous. Firms set export prices efficiently, maximising the profit that can be extracted from foreign markets.[2] However, policy may, depending on the general equilibrium structure of the economy, bring about a change in the domestic wage, this affecting export prices and the ToT.

Both these mechanisms generate a case for subsidising domestic sales to the domestic market and, in some cases, for a positive tariff on imports. These policies have the effect of switching expenditure to the home MC industry leading, in general, to a price and quantity response. The quantity response increases the number of varieties that are offered, thereby reducing the MC distortion of under-supply of varieties. The price response raises domestic wages relative to foreign and thereby improves the ToT. We provide a decomposition which enables us to attribute the impact of tariffs to a precise combination of MC and ToT effects.

These arguments apply both with symmetric and with heterogeneous firms[3]. The latter creates no fundamentally new arguments for policy since this specification of technology is not itself associated with any market failure. However, the combination of a fixed cost of exporting and heterogeneous productivity levels means that foreign firms' responses to policy become more elastic. In particular, the foreign MC sector may react to policy by changing the number of varieties sold in the domestic market. This reduces the value of using policy to increase the

---

[1] The MC distortion arises as price is greater than marginal cost and, for each variety, total benefit is greater than total cost. The ratios of benefit to cost are the same at the margin and for the totals, this giving the CES property that the market supports an outcome that is efficient, conditional on the level of employment in the sector (Dixit and Stiglitz, 1977). Dhingra and Morrow (2012) establish that this property also holds with heterogeneous firms.

[2] There is no strategic behaviour so 'strategic trade policy' arguments do not apply.

[3] By symmetric firms we mean firms that are identical, except in so far as their products are differentiated, as in Dixit-Stiglitz-Krugman models. Heterogeneous firms differ in their productivity, and hence in their endogenous choices to serve different markets, as in Melitz models.

number of domestic varieties, since more domestic varieties crowd-out some imported foreign varieties.

Related literature contains models which vary in (at least) four respects. Some papers look at the effects of changes in real trade barriers, others at tariffs; some look at unilateral changes, others bilateral; models vary in their general equilibrium structure; and they vary in whether they contain symmetric or heterogeneous firms. Since the objective of the present paper is to understand the welfare economics of monopolistic competition and firm heterogeneity (i.e. the inefficiencies present in the market equilibrium) we focus on tariff and other tax policy instruments that do not have a direct real cost effect, and look at a single open economy (unilateral rather than multilateral policy).

With this focus, we encompass several general equilibrium structures and both symmetric and heterogeneous firms. An older literature from the 1980s studies tax and tariff policy under monopolistic competition, although without firm heterogeneity. Variety and a terms-of-trade reasons for active policies are identified in work by Venables (1982, 1987), Flam and Helpman (1987) and Helpman and Krugman (1989). The newer literature includes papers by Demidova and Rodriguez-Clare (2009) and Felbermayr, Jung and Larch (2013). The former look at unilateral policy in a single sector economy and derive a policy result (taxing trade or subsidising domestic production) which we replicate and generalise. The latter extend the Demidova and Rodriguez-Clare model to large countries, and relate the optimal tariff to MC distortions and ToT effects. A recent paper by Costinot et al. (2016) works with a two-country model, looking at first best policy including the possibility that policies are firm specific (i.e. vary with firm productivity).

Much of the recent literature looks at reductions in real trade barriers. Baldwin and Forslid (2010) study the effect of changes in real trade barriers in a model with an MC and a PC sector, and show that trade liberalisation will have an “anti-variety” effect, an effect that is present in our work. However, their finding that lower trade barriers raise welfare does not generally hold with tariffs rather than real barriers. The substantial new literature on the gains from trade also focuses on real, rather than revenue raising, trade barriers. Arkolakis et al. (2012) and following work point to the importance of trade elasticities in determining the welfare effects of trade. This is developed further in Melitz and Redding (2015a) who study gains from trade and from bilateral liberalisation in a single sector economy, looking at both symmetric and heterogeneous firms. The two cases imply different trade elasticities which shape the gains from liberalisation and – in our work – shape optimal tax and tariff policy.

Our focus is on policy by a single country, maintaining the small open economy assumption. The penultimate section of the paper discusses how our results would extend to a full multi-country model, and hence to the work of Ossa (2011), Campolmi et al. (2014), and Bagwell and Staiger (2015).

The remainder of the paper is organised as follows. Section 2 sets out the model in quite an extensive way, carrying a lot of variables and making few substitutions. Results are derived using comparative static techniques, linearising the model and solving the ensuing equations to derive first order conditions for tax rates. The appendix gives the log-linearised system, and linear substitutions of the full system are readily undertaken by *Mathematica*. They generate explicit optimal (and constrained optimal) tax formulae which are the core of our results. This method enables us to derive optimal (and second best) taxes in a wide range of cases; it is conceptually simple, lending itself to relatively easy interpretation and to application to other issues.

Results are presented and explained in sections 4 and 5. Section 4 looks at the two polar cases, first when the PC sector is such that the supply of labour to the MC sector is perfectly elastic, and then when the PC sector has a fixed labour demand, a limiting case of which is no PC sector at all (Demidova and Rodriguez-Clare 2009). Section 5 places these in a general framework and shows that results are driven by a combination of MC distortion and ToT effects; we present a decomposition that separates these forces and establishes that the former drives results when labour supply is elastic (so quantity effects are large) and the latter when labour supply to the MC sector is inelastic (price effects dominate). Section 6 discusses the implications of relaxing the small open economy assumption and relates our results to literature with full two country models. Section 7 offers concluding comments.

2. The model

We first outline the ingredients of the monopolistically competitive (MC) sector. We do this in a succinct manner since many full expositions are in the literature (e.g. Melitz and Redding 2015b). Each firm in the sector produces a distinct variety of differentiated product. These products generate utility according to a sub-utility function with constant elasticity of substitution σ. E denotes total expenditure on MC products in the domestic economy and P is the price index (the unit expenditure function dual to the sub-utility function). The consumer price of a product is p, demand for the product is $p^{-\sigma}EP^{\sigma-1}$, and the value of its sales is $p^{1-\sigma}EP^{\sigma-1}$.

The marginal cost of a particular firm is W/φ where W is the price of labour, the only input, and φ is the firm's productivity. Firms mark-up price over marginal cost by factor $\sigma/(\sigma-1)$ so the producer price is $(W/\varphi)\sigma/(\sigma-1)$. The consumer price deviates from this according to ad valorem tax factor τ, so $p=\tau(W/\varphi)\sigma/(\sigma-1)$. The value, at consumer prices, of a firm's sales in one market is therefore $[\tau(W/\varphi)\sigma/(\sigma-1)]^{1-\sigma}EP^{\sigma-1}$. The firm captures fraction $1/\tau$ of this, so its revenue is $\tau^{-\sigma}[(W/\varphi)\sigma/(\sigma-1)]^{1-\sigma}EP^{\sigma-1}$. The remainder, fraction $1-1/\tau$, goes to government.[4] The firm's operating profit, π, is fraction $1/\sigma$ of its revenue, so $\pi=\tau^{-\sigma}(W/\varphi)^{1-\sigma}\varsigma EP^{\sigma-1}$ where $\varsigma\equiv\sigma^{-\sigma}(\sigma-1)^{\sigma-1}$.

Entry decisions incur fixed costs that have to be weighed against expected operating profits. In order to produce at all, each home firm pays a fixed cost Wf_E to draw a productivity parameter φ from distribution G_H. If this exceeds cut-off value φ_D the firm will sell in the domestic market after incurring further fixed cost Wf_D, so its expected profits on domestic sales are given by the first term in equation (1) below, where τ_D is the domestic tax rate. Similarly, exporting incurs fixed cost Wf_X, is subject to tax τ_X, and faces foreign demand curve with fixed expenditure and price index $\overline{E},\overline{P}$ (fixed by the small open economy assumption). A firm will export if productivity exceeds cut-off value φ_X, so the firm's expected profits on export sales are the second term in (1). The equation as a whole is the entry condition giving zero expected profits.

$$\int_{\varphi_D}\left\{EP^{\sigma-1}\varsigma\tau_D^{-\sigma}\left(\frac{W}{\varphi}\right)^{1-\sigma}-Wf_D\right\}dG_H+\int_{\varphi_X}\left\{\overline{EP}^{\sigma-1}\varsigma\tau_X^{-\sigma}\left(\frac{W}{\varphi}\right)^{1-\sigma}-Wf_X\right\}dG_H-Wf_E=0. \qquad (1)$$

The survival cut-offs are the lowest levels of productivity at which profits from an activity are non-negative. For domestic sales, φ_D satisfies

$$EP^{\sigma-1}\varsigma\tau_D^{-\sigma}\left(\frac{W}{\varphi_D}\right)^{1-\sigma}=Wf_D, \quad \text{and we define} \qquad \Phi_D\equiv\int_{\varphi_D}\varphi^{\sigma-1}dG_H, \qquad (2)$$

where Φ_D is an aggregate of the productivity of active firms according to their impact on sales. Similarly, the export cut-off is φ_X:

[4] Our results are qualitatively unchanged if, in addition to these revenue raising frictions, there are real trade costs.

$$\overline{E}\overline{P}^{\sigma-1}\varsigma\tau_X^{-\sigma}\left(\frac{W}{\varphi_X}\right)^{1-\sigma} = Wf_X, \qquad \Phi_X \equiv \int_{\varphi_X} \varphi^{\sigma-1} dG_H. \tag{3}$$

In addition to home firms, there are foreign firms some of which supply imports to the domestic market. The foreign wage is fixed at unity, the number of foreign firms is exogenous, and these firms have productivity distribution G_F (possibly different from G_H).[5] They choose whether or not to supply the domestic market. The fixed cost they incur in supplying imports is f_M, and they face an import tariff τ_M, so the importer cut-off is φ_M:

$$EP^{\sigma-1}\varsigma\tau_M^{-\sigma}\left(\frac{1}{\varphi_M}\right)^{1-\sigma} = f_M, \qquad \Phi_M \equiv \int_{\varphi_M} \varphi^{\sigma-1} dG_F. \tag{4}$$

It is convenient to have expressions for the total value of output sold by domestic firms in the domestic market, *D*, in the export market, *X*, and by foreign firms in the domestic market, *M* (all at consumer prices). The mass of domestic firms is denoted *N* and the mass of foreign firms $\overline{N}$ so, integrating over firms' sales at consumer prices and using (2), (3), (4),

$$D \equiv N\int_{\varphi_D} EP^{\sigma-1}\varsigma\sigma\left(\frac{\tau_D W}{\varphi}\right)^{1-\sigma} dG_H = EP^{\sigma-1}\varsigma\sigma(\tau_D W)^{1-\sigma} N\Phi_D \tag{5}$$

$$X \equiv N\int_{\varphi_X} \overline{E}\overline{P}^{\sigma-1}\varsigma\sigma\left(\frac{\tau_X W}{\varphi}\right)^{1-\sigma} dG_H = \overline{E}\overline{P}^{\sigma-1}\varsigma\sigma(\tau_X W)^{1-\sigma} N\Phi_X \tag{6}$$

$$M \equiv \overline{N}\int_{\varphi_M} EP^{\sigma-1}\varsigma\sigma\left(\frac{\tau_M}{\varphi}\right)^{1-\sigma} dG_F = EP^{\sigma-1}\varsigma\sigma\tau_M^{1-\sigma}\overline{N}\Phi_M. \tag{7}$$

Notice that $E = D + M$, and hence the usual definition of the price index follows from adding (5) and (7),

$$P^{1-\sigma} = \varsigma\sigma\left[(\tau_D W)^{1-\sigma} N\Phi_D + \tau_M^{1-\sigma}\overline{N}\Phi_M\right]. \tag{8}$$

We will call $N\Phi_D$ and $\overline{N}\Phi_M$ the 'equivalent number of varieties' supplied to the domestic market by home and foreign firms respectively. The terms Φ_D and Φ_M adjust the number of firms that enter according to the proportion that supply the domestic market and their productivity distribution.

[5] See Demidova (2008) for further development of the implications of different productivity distributions.

Government revenue, R, is earned from each of the tax instruments and, as noted above, is fraction $1 - 1/\tau$ of sales (at consumer prices) so:

$$R = D(1-1/\tau_D) + M(1-1/\tau_M) + X(1-1/\tau_X)\,. \tag{9}$$

Employment in the home MC sector, L, is implicitly defined by the fact that, since firms break even (in expectation), the wage bill in the sector is equal to the value of sales at producer prices,

$$WL = D/\tau_D + X/\tau_X\,. \tag{10}$$

Turning from the MC sector to the general equilibrium of the economy as a whole, we assume that there is a fixed endowment of labour (set at unity) of which L is used in the MC sector and the remainder, $1 - L$, is employed in the PC sector. The PC sector (if it exists) is freely traded with price unity and concave production function $F(1-L)$. The value of national output (at producer prices), Y, is therefore

$$Y = WL + F(1-L)\,. \tag{11}$$

Labour is employed in the PC sector to the point where the wage equals the marginal value product,

$$W = F'(1-L)\,. \tag{12}$$

The elasticity of labour supply from the PC sector with respect to the wage follows from this, and is denoted $\eta \equiv -F'/LF''$.[6]

Consumer income is the value of output plus government revenue, $Y + R$. Utility is Cobb-Douglas with expenditure share on the MC sector μ, giving utility U and MC expenditure E,

$$U = (Y+R)P^{-\mu}\,, \tag{13}$$

$$E = \mu(Y+R)\,. \tag{14}$$

This completes description of the equilibrium; there are 14 equations in $N, E, P, \Phi_D, \Phi_X, \Phi_M, D, X, M, R, Y, L, W$, and U.

The analysis of policy in our general case requires that both the MC and the PC sector are active (although the PC sector becomes inactive in one of the special cases we study). To ensure this, we assume that f_E is small enough for the expected profits of a domestic firm to be positive if $N =$

[6] This elasticity is, in general, not constant. For example, if the PC production function is iso-elastic, $F(1-L) = (1-L)^{\alpha}$, then $\eta = (1-L)/\{(1-\alpha)L\}$. Strict concavity can be interpreted as dependence on a PC sector specific factor of production, with factor share $(1 - \alpha)$.

0, implying that the MC industry is present in the domestic economy; and large enough that expected profits are less than or equal to zero if the entire labour force is employed in the MC sector.[7] We show in Appendix 1 that expected profits are monotonically decreasing with *N*, this giving an interior equilibrium. Profits decline with *N* for two reasons. One is that the wage may increase as the PC sector contracts (equation 12). The other is domestic market crowding; entry of domestic firms increases supply to the domestic market and thereby reduces the price index (equation 8) and $EP^{\sigma-1}$, hence reducing profits. Market crowding is offset by displacement of imports, occurring as lower values of $EP^{\sigma-1}$ increase the importer cut-off, φ_M (equation 4). However, this displacement occurs progressively (because importers are heterogeneous), so increasing *N* is sure to reduce $EP^{\sigma-1}$ and hence the expected profits of domestic firms.

3. Comparative statics

Our primary task is to investigate the effect of changes in tax instruments, and we do this by log differentiation of the equilibrium. Expressions so derived contain proportionate changes in the tax instruments and endogenous variables, together with some further variables capturing relative values of endogenous variables. We define these as follows.

The share of domestic firms' sales in the domestic MC market (at consumer prices), s_D:

$$s_D \equiv D/\left(D+M\right) = D/E. \tag{15}$$

The share of export sales in MC production (at producer prices), s_X:

$$s_X \equiv \left(X/\tau_X\right)/WL. \tag{16}$$

Government revenue as a share of total consumer income,

$$r \equiv R/(Y+R) = \mu R/E. \tag{17}$$

Finally, we define the share of the MC sector in total income, s_Y:

$$s_Y \equiv WL/Y. \tag{18}$$

Given the share of MC spending in income, μ, the shares s_D, s_X, s_Y are not independent. Using (10), (14) - (18) the equation linking them is,

[7] Expected profits are the left hand side of equation (1). The thought experiment here is to evaluate profits given the number of domestic firms, *N*, and with other variables in equilibrium, i.e. satisfying equations (2) – (14)

$$s_D = s_Y(1-s_X)(1-r)\tau_D/\mu . \tag{19}$$

For the remainder of the paper we assume that productivity is Pareto distributed, i.e. $G_i(\varphi) = 1-\varphi^{-k_i}$, so that $\Phi_j = k_i\varphi_j^{\sigma-k_i-1}/(k_i-\sigma+1)$ for i = H, j = D, X and for i = F, j = M. We define the parameter $\lambda_i \equiv \sigma k_i/(\sigma-1)$ and make the standard assumption that $k_i \geq \sigma-1$ so that $\lambda_i \geq \sigma$.

The full log-linearised system is given in Appendix 2, with proportionate changes denoted ^. There are exogenous changes in three tax instruments, and 14 equations giving changes in the endogenous variables (corresponding to equations (1) – (14) above)[8]. Our principal interest is to obtain the coefficients giving the effect of changes in each of the tax rates, $\hat{\tau}_D, \hat{\tau}_X, \hat{\tau}_M$, on utility, $\hat{U}$. These coefficients will in general contain parameters $\sigma, \mu, \lambda_H, \lambda_F$; endogenous variables r, s_D, s_X, s_Y; and tax instruments, τ_D, τ_X, τ_M. Setting these coefficients equal to zero gives the first order conditions for optimal policy and, solving for tax rates, we obtain explicit solutions for optimal policies. These are tabulated and discussed in following sections. While the approach is conceptually straightforward the algebra is cumbersome, even for the linearised system of proportional changes; we use *Mathematica*.

One expression containing total derivatives is helpful for interpretation of results. The welfare effects of changes can be written as (see Appendix 3):

$$\hat{U} = \frac{X}{Y+R}\left[\hat{W} + \hat{\tau}_X\right] + \frac{1}{\sigma-1}\left[\frac{D}{Y+R}(\hat{N} + \hat{\Phi}_D) + \frac{M}{Y+R}\hat{\Phi}_M\right]$$
$$+ \frac{D}{Y+R}\widehat{\left\{\frac{D}{W\tau_D}\right\}}\frac{\tau_D-1}{\tau_D} + \frac{M}{Y+R}\widehat{\left\{\frac{M}{\tau_M}\right\}}\frac{\tau_M-1}{\tau_M} + \frac{X}{Y+R}\widehat{\left\{\frac{X}{W\tau_X}\right\}}\frac{\tau_X-1}{\tau_X}. \tag{20}$$

In this equation ^ denotes a proportional change, and the terms in curly brackets are proportional changes in quantities, i.e. changes in the values *D*, *M*, *X* deflated by their prices. The equation is useful as it shows how changes in the tax instruments and in endogenous variables determine the change in welfare.

The first term on the right-hand side of (20) is the terms-of-trade effect, capturing the welfare effects of an increase in the export price (and remembering that the border price of imports is

[8] The comparative-static equations are numbered (1') to (14') and are found in Appendix 2.

constant by the small open economy assumption). The export price is affected by an export tax and by any change in the domestic wage. The second term we refer to as the MC distortion effect – it captures the welfare effect of changes in the number of varieties available to domestic consumers. Notice that the expression is on changes in the 'equivalent number of varieties', $N\Phi_D$ and $\bar{N}\Phi_M$, so the effects of policy on firm selection enter via changes in Φ_D and Φ_M. The square bracket in this term is multiplied by factor $1/(\sigma-1)$ capturing the welfare gain from an increase in the number of varieties (given total expenditure).[9] Finally, the terms in the second row show the direct welfare effects of policy distortions, generating welfare gain if policy expands an activity where the tax instrument creates a positive wedge between consumer and producer prices.

Optimal policies are found by setting $\hat{U}=0$ in (20), either for one policy instrument at the time or jointly for all three taxes, and with changes in endogenous variables coming from the log-linearised system given in Appendix 2. We derive results both for heterogeneous firms, and for a case which we term symmetric firms. The difference is that in the symmetric case there are no firm selection effects, i.e. the number of firms (and varieties) supplying each market varies only with *N*, the number of active domestic firms. The Krugman-Dixit-Stiglitz model, in which all active firms supply all markets, is an example of this.[10] In our setting, the parameter restriction equivalent to switching off selection effects is $\lambda_H = \lambda_F = \sigma$, this implying that Φ_i, *i* = *D*, *X*, *M*, are constant (see appendix 2 equations 2' – 4'). The point can be illustrated by inspection of the elasticity of the value of imports with respect to a tariff which, using (7') and (4'), is

$$\frac{d\ln M}{d\ln\tau_M} = (1-\sigma) - (\lambda_F - \sigma) = 1 - \lambda_F .$$

Following Chaney (2008), the first element $(1-\sigma)$ shows the intensive margin – the effect of a tariff on imports from the existing (foreign) firms – and the second element is the extensive margin, capturing the effects of changing cut-off levels for the selection into imports.[11] With symmetric firms there is no selection effect so the import elasticity is given by the intensive

[9] Seen most simply for a closed economy with symmetric firms where the price index is $P^{1-\sigma} = Np^{1-\sigma}$. This falls as *N* increases according to $\hat{P} = \hat{N}/(1-\sigma) + \hat{p}$, see also Appendix 2 equation (8').

[10] Melitz and Redding (2014) compare gains from trade with homogenous and heterogeneous firms in a similar way, pointing out that the case with 'symmetric' firms does not require that all firms are identical. The key point is that there is no endogenous selection effect.

[11] Head and Mayer (2015) show how the extensive margin can be decomposed into a selection effect and a composition effect, and Felbermayr et al. (2015) use the same framework to show how the effects of a tariff may differ from the effect of a real trade cost. Our expressions are consistent with Felbermayr et al.

margin (i.e. has $\lambda_F = \sigma$). With heterogeneous firms, both intensive and extensive margin matter, and the combined elasticity is $(1-\lambda_F)$. Similar reasoning applies for export elasticities. Since we allow for the possibility that $\lambda_H \neq \lambda_F$, our setup is general enough for firms in one country to be symmetric and elsewhere heterogeneous. In terms of the welfare effects in (20), the symmetric firm case implies that the variety (MC distortion) effect only appears through changes in the mass of domestic firms, as in standard Krugman-Dixit-Stiglitz models.

4. Optimal policies

We start by presenting results for two special cases, first with a perfectly elastic supply of labour from the PC sector to the MC sector ($\eta = \infty$, section 4.1) and then perfectly inelastic supply ($\eta = 0$, section 4.2). In both sections we look at policy with symmetric firms (case A) and heterogeneous firms (case B). We present optimal tax formulae when all three tax instruments are optimised (denoting values τ^{**}) and where just one instrument is optimised (τ^{*}) with other instruments not used (i.e. set at unity). We call the former first-best policies and the latter second-best. We omit expressions that are excessively complex.[12]

In each section we tabulate results and draw out intuition, concentrating discussion of intuition on the effects of an import tariff. In section 4.1 we are able to develop intuition by some simple arguments which show how policy equates marginal benefits and costs. For other cases intuition is more complex and a crucial issue becomes understanding the different roles of MC distortions and ToT effects. Our discussion of this is contained in section 5.

4.1 Perfectly elastic labour supply to the MC sector

A frequent assumption in the literature is that there is a PC sector that produces a good that has fixed world price with constant returns to labour alone[13]. This fixes the wage in the economy, meaning that the MC sector faces a perfectly elastic labour supply curve. In terms of the model, $\eta = \infty$, $\hat{W} = 0$, and $\hat{L}$ adjusts freely. Table 1 presents results.

The first row of Table 1 gives optimal policy with symmetric firms. The optimum is achieved by a subsidy on domestic sales of home firms, $\tau_D^{**} = (\sigma-1)/\sigma < 1$, with the import tariff and export

---

[12] E.g. reporting τ_X^{*} for the special cases in section 4 but not the general case in section 5. These more complex expressions are available on request from the authors.

[13] As stated in the introduction a number of contributions make this assumption. In section 6, below, we compare their results to ours and show how our decomposition of effects illuminates the underlying welfare economics.

tax set at unity. The effect of the subsidy is to expand home's MC sector, bringing in new varieties and thereby mitigating the MC distortion. The optimal import tariff is zero, $\tau_M^{**} = 1$, since the economy is importing goods at constant price. The optimal export tax is also zero. The economy can vary its export terms of trade, raising their price by an export tax; however, domestic firms have already chosen the price that maximises profit extracted from the foreign market, and any deviation from this is welfare reducing. With reference to equation (20), the optimal policies are determined from the trade-off between reducing the MC distortion through increasing the number of domestic varieties and increasing the policy distortion from subsidising domestic consumption.

Turning to cases in which only one instrument is used, the optimal value of τ_D^* is as above. If the tariff is the only instrument, then it should be positive, $\tau_M^* > 1$. In terms of equation (20) this distorts the domestic price of imports away from the marginal cost at which they are supplied to the economy, but is a second best policy to expand the domestic MC sector, bringing in domestic varieties and mitigating the MC distortion.[14]

The economic intuition underlying this can be developed by establishing the marginal benefits and costs of the quantity changes, and noting that optimal policy equates the ratios of marginal benefit to marginal cost across affected quantities. The first quantity change from an import tariff is a reduction in imports of MC products. Their marginal cost to the economy is their price, and their marginal benefit is price times the factor τ_M, as the tariff raises the marginal value.[15] The second quantity change is that, as the tariff shrinks imports of MC products, so it expands domestic production. As usual in a Dixit-Stiglitz model sales per firm are unchanged, so the quantity change is met entirely by a change in the mass of domestic firms, *N*. Entry of a new variety brings consumer surplus, and the ratio of utility to expenditure is $\sigma/(\sigma-1)$. Since firms break even expenditure on each product equals (expected) costs, so the ratio of benefit to cost for a marginal change in the number of varieties is $\sigma/(\sigma-1)$. Equating these marginal benefit-to-cost ratios for the change in imports and the change in domestic production gives the tariff formula $\tau_M^* = \sigma/(\sigma-1)$.

Continuing in part A of table 1, the final row gives the optimal export tax when other instruments are not used; it should be used to subsidise exports ($\tau_X \leq 1$) with the subsidy rate going to zero as

[14] Notice from (1') that $\hat{E} + (\sigma-1)\hat{P} = 0$, so $\hat{M} = (1-\sigma)\hat{\tau}_M < 0$ (from 7') and $\hat{N} = \hat{D} > 0$ (from (5'), creating the policy distortion and variety effects respectively.

[15] Equivalently, marginal benefit is price times $1 + (\tau_M - 1)$, where the second term is tariff revenue earned.

the share of exports in production increases (the maximum value of τ_X is unity, occurring if $s_X = 1$). This is second best policy shaped by the interaction of two forces. The subsidy attracts entry of domestic firms (an increase in *N*) which increases the number of varieties offered in the domestic market, partially correcting the MC distortion. But, since this policy involves subsidising foreign consumers (ToT loss) it is an inefficient instrument, so the subsidy is relatively small (equal to zero if $s_X = 1$). In terms of equation (20), it is a trade-off between three elements – the ToT loss, the MC distortion gain, and the direct policy distortion from subsidising exports.

Table 1: Optimal policy with perfectly elastic labour supply.

A: Symmetric firms.

All taxes optimally set:		
$\tau_D^{**} = \frac{\sigma-1}{\sigma} < 1$	$\tau_M^{**} = 1$	$\tau_X^{**} = 1$
Fixed taxes	*Optimised taxes*	
$\tau_M = \tau_X = 1$	$\tau_D^{*} = \frac{\sigma-1}{\sigma} < 1$	
$\tau_D = \tau_X = 1$	$\tau_M^{*} = \frac{\sigma}{\sigma-1} > 1$	
$\tau_D = \tau_M = 1$	$\tau_X^{*} = 1 - \frac{(1-s_X)\left[\sigma(1-s_D) + (1-\mu)s_D\right]}{(\sigma-1+\mu)\left[\sigma s_D + \sigma s_X(1-s_D) - s_D(1-s_X)\right] - \sigma\mu s_X} < 1$	

B: Heterogeneous firms: $\lambda_F, \lambda_H > \sigma$

All taxes optimally set:		
$\tau_D^{**} = \frac{\sigma-1}{\sigma} < 1$	$\tau_M^{**} = \left(\frac{\sigma-1}{\sigma}\right)\left(\frac{\lambda_F}{\lambda_F-1}\right) < 1$	$\tau_X^{*} = 1$
Fixed taxes	*Optimised taxes*	
$\tau_M = \tau_X = 1,$	$\tau_D^{*} = \left(\frac{\sigma-1}{\sigma}\right)\left\{1 + \frac{(1-s_D)(\lambda_F-\sigma)(\sigma-1+\mu)}{(\sigma-1)\left[(1-s_D)\lambda_F(\sigma-1+\mu) + \sigma(1-\mu)\right]}\right\} < 1$	
$\tau_D = \tau_X = 1$	$\tau_M^{*} = \left(\frac{\lambda_F}{\lambda_F-1}\right) > 1$	
$\tau_D = \tau_M = 1$	$\tau_X^{*} = 1 - \frac{(1-s_X)\left[\sigma(1-s_D) + (1-\mu)s_D\right]}{(\sigma-1+\mu)\left[\lambda_H s_D + \lambda_F s_x(1-s_D) - s_D(1-s_X)\right] - \sigma\mu s_X} < 1$	

Part B of Table 1, the lower panel, gives policy with heterogeneous firms. The key difference is that expanding the domestic MC sector now reduces the number of foreign firms that select to supply imports (i.e. raises the foreign import cut-off φ_M and reduces Φ_M in equations 4 and 4') so welfare gains from drawing in domestic varieties are offset by loss of imported varieties. First-best optimal policy is therefore the domestic subsidy $\tau_D^{**} < 1$ as before, combined with an import subsidy, $\tau_M^{**} < 1$. The size of this depends on the magnitudes of the expenditure elasticities with respect to tax rates, with the import subsidy less than the subsidy to domestic firms, and collapsing down to the symmetric case when $\lambda_F = \sigma$. Notice that the expression for τ_M^{**} contains λ_F not λ_H so it is heterogeneity of foreign, not domestic firms that makes the case for the import subsidy. With reference to equation (20) again, we now have two forces affecting the number of varieties in the second term of the equation: an increase in the equivalent number of domestic varieties $(\hat{N} + \hat{\Phi}_D > 0)$ and a reduction in imported varieties ($\hat{\Phi}_M < 0$); optimal policy requires two instruments, each set to balance variety gains against the direct policy distortions.

If τ_D is the only instrument used then it should be a subsidy, although at a lower rate (i.e. τ_D^* closer to unity) than in the symmetric case because of the loss of imported varieties; the optimal value depends on foreign selection, λ_F, not on domestic firm selection. The import tariff alone mirrors that in the symmetric case, but with σ replaced by λ_F, implying a lower tariff .[16] The argument for policy is simply the MC distortion, but the number of MC varieties on offer now changes at two margins, *N* and Φ_M .

The final row of the table is the export tax with heterogeneous firms. As in the symmetric case, the export subsidy is determined by tension between ToT loss and variety gain. It collapses to the symmetric case if $\lambda_F = \lambda_H = \sigma$ and is a smaller subsidy if $\lambda_F, \lambda_H > \sigma$. This is because changes in both the import and export cut-offs reduce the impact of the export subsidy on the number of varieties sold in the domestic market; some import varieties are crowded out, and some of the quantity response of the domestic industry is more firms exporting (rather than an increase in *N*). It is noteworthy that this is the only result reported Table 1 in which λ_H appears, i.e. where heterogeneity of domestic firms has any bearing on policy. The reason can be seen by

---

[16] Intuition for this comes from extending the line of reasoning that we used in the symmetric case, but where proportion $(\sigma - \lambda_F)/(1 - \lambda_F)$ of the change in imports comes from the extensive margin, i.e. a change in the number of varieties supplied, and this change carries the premium $1/(\sigma - 1)$.

inspection of (1') - (4'); the productivity cut-offs φ_D and φ_X are unaffected by either τ_M or τ_D (see footnote (13), again), while both these cut-offs are affected by τ_X.

4.2 Inelastic labour supply to MC sector

We now look at situations in which the general equilibrium structure of the economy is such that the supply of labour to the MC sector is perfectly inelastic, this including the one sector economy studied by Demidova and Rodriguez-Clare (2009, henceforth DRC). While tax formulae derived in this case are similar (in several cases identical) to those in the previous section, they are driven by quite different mechanisms. Essentially, when labour supply is elastic quantity changes interact with the MC distortion, but when it is fixed quantity effects are absent and price effects (the ToT) drive results. The distinction is key to understanding the role of policy – and welfare economics more generally – in this class of models.

It turns out that there are two assumptions involved in going from a general setting to the results of DRC. One is that $\eta = 0$ so that the supply of labour to the MC sector is perfectly inelastic. The other is that $\mu = 1$, so domestic consumers only purchase the MC good. Table 2 presents results in three stages; first, the infinitely elastic labour supply case ($\eta = \infty,\ \mu \in (0,1)$, repeating the first rows of each section of Table 1); second, $\eta = 0,\ \mu \in (0,1)$; and finally $\eta = 0,\ \mu = 1$, the DRC case.

The first row of parts A and B of Table 2 gives results for $\eta = \infty,\ \mu \in (0,1)$ in which optimal policy requires setting all three of the instruments at the values indicated.[17] In the second row of each part of the table the size of the domestic MC sector is fixed by inelastic labour supply. This has the consequence that, since domestic MC production cannot change, the only margin is its distribution between domestic and export sales. Hence, policy is achieved by the ratio τ_D / τ_X taking the value indicated (the separate values of τ_D and τ_X being immaterial). With symmetric firms, part A, $\tau_M = \tau_M^{**} = 1$. With firm heterogeneity, part B, it is optimal to subsidise imports as well as domestic consumption to counteract the crowding out of imported varieties due to the foreign selection effect.

In the third row of each block the consumer demand margin is also removed, by assuming that consumption consists solely of the MC good, $\mu = 1$. This means that the only tax instrument that matters is the combination $\tau_D / \tau_X \tau_M$. Varying any one of τ_D, τ_X, τ_M has the same real effect,

[17] Table 2 does not report cases where an instrument is constrained to equal unity.

changing exports and imports together. This is a consequence of Lerner symmetry in what is now a very simple economy. With heterogeneous firms (final row), this gives the result derived by DRC in a single sector economy. We note that it does not require that the economy contains a single sector, or that trade in the MC good be balanced (the PC sector could still exist, simply taking a fixed amount of labour to produce and export a fixed amount of output). But it does require that margins of substitution between the MC sector and the PC sector – on both the supply and demand side – are switched off.

Table 2: First best optimal policy: inelastic labour supply and Lerner symmetry

A: Symmetric firms.

$\eta=\infty,\ \mu\in(0,1)$:	$\tau_D^{**}=\dfrac{\sigma-1}{\sigma}$	$\tau_M^{**}=1$	$\tau_X^{**}=1$
$\eta=0,\ \mu\in(0,1)$:	$\left(\dfrac{\tau_D}{\tau_X}\right)^{**}=\dfrac{\sigma-1}{\sigma}$	$\tau_M^{**}=1$	
$\eta=0,\ \mu=1$:	$\left(\dfrac{\tau_D}{\tau_M\tau_X}\right)^{**}=\dfrac{\sigma-1}{\sigma}$		

B: Heterogeneous firms. $\lambda_F,\lambda_H>\sigma$

$\eta=\infty,\ \mu\in(0,1)$:	$\tau_D^{**}=\dfrac{\sigma-1}{\sigma}$	$\tau_M^{**}=\left(\dfrac{\sigma-1}{\sigma}\right)\left(\dfrac{\lambda_F}{\lambda_F-1}\right)<1$	$\tau_X^{**}=1$
$\eta=0,\ \mu\in(0,1)$:	$\left(\dfrac{\tau_D}{\tau_X}\right)^{**}=\dfrac{\sigma-1}{\sigma}$	$\tau_M^{**}=\left(\dfrac{\sigma-1}{\sigma}\right)\left(\dfrac{\lambda_F}{\lambda_F-1}\right)<1$	
$\eta=0,\ \mu=1$:	$\left(\dfrac{\tau_D}{\tau_M\tau_X}\right)^{**}=\dfrac{\lambda_F-1}{\lambda_F}$		

Despite the similarities in the tax and tariff formulae in Tables 1 and 2, the mechanisms driving policy are completely different in the two cases. Consider first symmetric firms with $\eta=0,\ \mu=1$. Policy cannot be operating through the domestic MC distortion (as it does in Table 1), since the size of the MC sector, number of varieties offered, and output of each firm are completely fixed and invariant to policy. Instead, the domestic wage changes, and it is this that generates a ToT effect. Given this ToT effect, the optimal value of the export tax (or equivalently import tariff) is simply the reciprocal of the foreign elasticity of demand for home's exports.[18] Similarly, with heterogeneous firms and $\eta=0,\ \mu=1$, there is a ToT effect and the optimal tariff is the inverse of the foreign supply elasticity, $\tau_M^{**}-1=1/(\lambda_F-1)$. Allowing some flexibility in consumer demand, i.e. $\eta=0,\ \mu\in(0,1)$, the second row in part B of Table 2, creates the possibility of increasing the number of foreign firms that supply the domestic market, and this is achieved by an import subsidy. In this case there is a combination of ToT and variety effects and the relative importance of each depends on the strength of the selection mechanism. We discuss this more fully in the context of the general case (next section), using our decomposition of the welfare effects of policy, equation (20).

5. The general case: optimal policies and welfare decomposition

Our focus up to this point has been on deriving optimal tariff and tax formulae for the special cases of perfectly elastic and fixed labour supply to the MC sector. We now switch to the general case and investigate two issues. First we derive and discuss optimal policies (section 5.1) and then, focusing on the import tariff, decompose its effect on welfare into a ToT part and an effect through MC distortions (section 5.2).[19]

5.1 Optimal policies in the general case

Table 3 gives first and second best optimal policies for the general case (i.e. $\eta\in[0,\infty),\ \mu\in[0,1]$). At the first best policy optimum (top rows of each part of the table) we can think of each instrument being targeted at a particular margin where welfare gains can be achieved. There are three such margins in this model. (i) The number of domestic varieties offered to domestic consumers; (ii) the number of foreign varieties offered to domestic consumers (endogenous only if foreign firms are heterogeneous and select into exporting); and (iii) the export terms of trade, which depend on the domestic wage and the export tax.

---

[18] Thus, in the symmetric firms case $\tau_X^{**}-1=\tau_M^{**}-1=1/(\sigma-1)$. We have also derived results with distinct notation for home and domestic elasticities of substitution, σ_H,σ_F; expressions in Table 2A contain σ_F not σ_H.

[19] Ossa (2011) discusses the importance of labour market flexibility for policy motives; however, in his analysis he assumes a perfectly elastic labour supply to the MC sector. Crozet and Trionfetti (2008) analyse the importance of concavity in the PC sector for the home-market effect. However, they do not focus on policy issues. See section 6 for further discussion.

At the first best optimum the import tariff is set to target the number of foreign varieties. This requires a subsidy only if this number is endogenous (i.e. foreign firms are heterogeneous), $\tau_M^{**} = \left(\frac{\sigma - 1}{\sigma}\right)\left(\frac{\lambda_F}{\lambda_F - 1}\right) \leq 1$ (Tables 1B, 2B, 3B); it is not used otherwise, since the border price of imports is fixed. The tax on domestic sales of domestic output is set to correct the MC distortion, so $\tau_D^{**} = (\sigma - 1)/\sigma$. The export tax is set to control the ToT, and has value $\tau_X^{**} = 1$.

Several remarks are in order on the last two of these. First, as emphasised in the preceding section, if η is finite then variations in τ_D and in other instruments influence the wage and hence the ToT. However the (net) value of any such ToT change is zero if τ_X has been optimised. This can be seen by noting that the first order condition for τ_D, evaluated at $\tau_M = \tau_M^{**}$, is $\tau_D^* = \frac{\sigma - 1}{\sigma}\left\{\frac{\tau_X + A_D}{1 + A_D}\right\}$ with $A_D \equiv \eta/[(\sigma - 1 - \eta)s_x]$ if firms are symmetric, and $A_D \equiv \eta/[(\lambda_H - 1 - \eta)s_x]$ with heterogeneous firms. As is apparent, τ_D^* depends on η unless τ_X has been optimised, i.e. takes its first best value $\tau_X^{**} = 1$.[20]

Second, why is the ToT margin optimised by setting the export tax at $\tau_X^{**} = 1$? As usual in the monopolistic competition framework, firms have chosen their price to maximise profits earned from the foreign market, so the price of exports is set to equate marginal export revenue with marginal cost. An export tax will change export prices (the ToT) and also the number of varieties produced domestically, but the former has been optimised by firm behaviour, and the latter optimised by setting $\tau_D^{**} = (\sigma - 1)/\sigma$.[21] Notice that if, instead of τ_D, the economy used a production tax, τ_P, the first best policy would involve subsidy $\tau_P = (\sigma - 1)/\sigma$ and tax $\tau_X = \sigma/(\sigma - 1)$ (clawing back the production subsidy on exports), the former targeted at the MC distortion, and the latter at the ToT. We conjecture that global efficiency requires a production subsidy in all countries and no export tax; hence, the unilateral first best optimum given in Table 3 is globally inefficient due to the ToT driven export tax.

---

[20] Or there are no exports, $s_X = 0$. Notice also that this expression collapses to the corresponding expression in Table 2 when $\eta = 0$.

[21] The first order condition for τ_X, evaluated at $\tau_M = 1$, is $\tau_X^* = [\tau_D\sigma/(\sigma - 1) + A_X]/[1 + A_X]$, with $A_X \equiv \eta\sigma/[(\sigma - 1 - \eta)(1 - s_x)\{\sigma - s_D(\mu + \sigma - 1)\}]$. Thus, $\tau_X^* = 1$ only if $\tau_D = (\sigma - 1)/\sigma$.

Table 3: Optimal policy in the general model

A: Symmetric firms.

All taxes optimally set:

$$\tau_D^{**} = \frac{\sigma - 1}{\sigma} < 1 \qquad \tau_M^{**} = 1 \qquad \tau_X^{**} = 1$$

Fixed taxes	*Optimised taxes*
$\tau_M = \tau_X = 1$	$\tau_D^{*} = \frac{\sigma - 1}{\sigma}$.
$\tau_D = \tau_X = 1$	$\tau_M^{*} = \frac{\sigma}{\sigma - 1}\left\{1 - \frac{1-\mu}{\sigma(1-\mu) + (\sigma - 1 + \mu)s_D\left[(1-s_X)\eta + s_X(\sigma - 1)\right]}\right\} \in \left[1, \frac{\sigma}{\sigma - 1}\right]$.

B: Heterogeneous firms: $\lambda_F, \lambda_H > \sigma$

All taxes optimally set:

$$\tau_D^{**} = \frac{\sigma - 1}{\sigma} < 1, \qquad \tau_M^{**} = \left(\frac{\sigma - 1}{\sigma}\right)\left(\frac{\lambda_F}{\lambda_F - 1}\right) < 1, \qquad \tau_X^{**} = 1$$

Fixed taxes	*Optimised taxes*
$\tau_M = \tau_X = 1,$	$\tau_D^{*} = \left(\frac{\sigma - 1}{\sigma}\right)\left\{1 + \frac{(1-s_D)(\lambda_F - \sigma)(\sigma - 1 + \mu)}{(\sigma - 1)\left[(1-s_D)\lambda_F(\sigma - 1 + \mu) + \sigma(1-\mu)\right]}\right\}$
$\tau_D = \tau_X = 1,$	$\tau_M^{*} = \frac{\lambda_F}{\lambda_F - 1}\left\{1 - \frac{1-\mu}{\sigma(1-\mu) + (\sigma - 1 + \mu)s_D\left[(1-s_X)\eta + s_X(\lambda_H - 1)\right]}\right\} \in \left[1, \frac{\lambda_F}{\lambda_F - 1}\right]$

What of the other results in table 3? With symmetric firms the second best domestic subsidy coincides with the first-best subsidy, since $\tau_M = \tau_X = 1$ are the first-best values. With heterogeneous firms, the second best domestic subsidy is lower than the first-best value, the reason being the crowding out of imported varieties, as discussed in section 4.1. It is worth noticing that τ_D^{*} does not depend on η. Again, this follows from the fact that there is no additional ToT motive as long as τ_X is at its optimal value ($\tau_X = 1$).

With symmetric firms the second best optimal import tariff (with $\tau_D = \tau_X = 1$) is given in the third row of the table. This is an import tax that expands the number of domestic varieties on offer and, if η is finite, also changes the ToT. We look at this in detail in the next subsection, and decompose the impact of the MC distortion and ToT effect.

Finally, firm heterogeneity affects the magnitude, but not the sign of second-best policies. The role of firm heterogeneity comes through by comparing parts A and B of Table 3. Selection effects for importers $(\lambda_F > \sigma)$ leads to a lower optimal tariff or a lower domestic subsidy, since domestic varieties crowd out imported ones. Selection at home $(\lambda_H > \sigma)$ on the other hand, increases the optimal (second-best) tariff. The reason is that with heterogeneous firms, the change in exports following a tariff will in part come as a reduction in the share of firms exporting (the extensive margin). This dampens the wage effect and increases the variety effect of an import tariff.

5.2 Decomposition of the tariff effect

We have shown that policy is determined by a combination of MC and ToT effects, the former dominant if the elasticity of labour supply to the MC sector is high, the latter if it is low. We now give the precise decomposition of these effects, building on equation (20) and looking not at the first best optimum but just at an import tariff, with other instruments not used, so $\tau_D = \tau_X = 1$ and $\hat{\tau}_D = \hat{\tau}_X = 0$. From (20), the condition for the optimal tariff then becomes (see also Appendix 3)

$$\hat{U} = \frac{X}{Y+R}\hat{W} + \frac{1}{\sigma-1}\left[\frac{D}{Y+R}\left(\hat{N}+\hat{\Phi}_D\right) + \frac{M}{Y+R}\hat{\Phi}_M\right] + \frac{M}{Y+R}\left[1-\frac{1}{\tau_M}\right]\left(\hat{M}-\hat{\tau}_M\right) = 0 \qquad (21)$$

The changes $\hat{W}, \hat{N}, \hat{\Phi}_D, \hat{\Phi}_M, \hat{M}$ come from differentiation of the equilibrium with respect to τ_M, and follow from equations (1') – (14') of Appendix 2, Making the substitutions yields a complex expression which it is convenient to summarise as $\hat{U} = \{\text{ToT} + \text{MC} + \text{TW}\}\hat{\tau}_M = 0$, where ToT is the terms-of-trade effect, MC is the variety effect, and TW is the effect of the tariff wedge. If the tariff is optimally chosen the term in curly brackets is equal to zero, $\text{TW} = -\{\text{ToT} + \text{MC}\}$, so

$$\hat{U} = (\text{ToT} + \text{MC})\left\{\frac{\text{ToT}}{\text{ToT}+\text{MC}} + \frac{\text{MC}}{\text{ToT}+\text{MC}} - 1\right\}\hat{\tau}_M = 0. \qquad (22)$$

The first two terms in curly brackets give the relative contributions of the TOT and MC effects. The expression for the share of the ToT and MC effects in the general case (evaluated at the optimal second-best tariff rate, i.e. τ_M^* from the final row of Table 3B) are (see Appendix 3),

$$\frac{\text{ToT}}{\text{ToT}+\text{MC}} = \frac{(\sigma-1+\mu)s_D s_x(\lambda_F - 1)}{(\sigma-1+\mu)s_D\left[\eta(1-s_X) + s_X(\lambda_H - 1)\right] - (1-\mu)(\lambda_F - \sigma)}, \qquad (23)$$

$$\frac{\text{MC}}{\text{ToT}+\text{MC}} = 1 - \frac{\text{ToT}}{\text{ToT}+\text{MC}}. \tag{24}$$

It is clear from (23) that a more elastic supply of labour to the MC sector (higher η) reduces the contribution of the ToT effect. As $\eta \to \infty$ so (23) goes to zero and (24) to unity, confirming that the results of Table 1 are driven entirely the MC distortion, not the ToT effect. The polar opposite case of Table 2 has $\eta = 0, \mu = 1$, so (23) and (24) become

$$\frac{\text{ToT}}{\text{ToT}+\text{MC}} = \frac{\lambda_F - 1}{\lambda_H - 1}, \qquad \frac{\text{MC}}{\text{ToT}+\text{MC}} = \frac{\lambda_H - \lambda_F}{\lambda_H - 1}. \tag{25}$$

If the distributions of productivity are the same in both countries, $\lambda_H = \lambda_F$, then, with $\eta = 0, \mu = 1$, welfare effects are entirely due to ToT arguments. Although individual firms take into account the slope of foreign demand curves for their varieties they do not internalise the fact that changes in output change the wage, and it is this that creates the ToT effect. However, if $\lambda_H > \lambda_F$ then there is also a positive MC distortion effect. Essentially, the (positive) response of domestic varieties to a tariff is greater in absolute value than the (negative) response of foreign varieties, meaning that (even in the one sector economy of section 4), there is a variety gain from an import tariff. Hence, for the single-sector economy, the main message from this decomposition is that, while some MC distortion effects are present if $\lambda_H \neq \lambda_F$, the case for the tariff is driven principally by the fact that it improves the ToT. As the tariff cuts imports, so the wage has to rise in order to cut exports in line. More generally, the share of the ToT effect is strictly decreasing in both η and μ, as can be seen from equation (23) [22].

6. Small versus large economy analysis

In the analysis so far we have focused on a small open economy. What additional effects come into play if this assumption is relaxed? Domestic policy can then, in principle, change the foreign wage, mass of active firms, and price index. This will change equilibrium responses (the derivatives of Appendix 2), and adds two new terms to our expression for domestic welfare change, equation (20). A change in the foreign wage changes the price of domestic imports, having a ToT effect, and a change in the mass of foreign firms changes the number of varieties sold in the domestic economy, so has an MC distortion effect. Formally, equation (20) becomes

[22] This thought experiment holds the share variables, s_x, s_D constant.

$$\hat{U}_H = \left[\frac{X}{Y+R}\left(\hat{W}_H + \hat{\tau}_X\right) - \frac{M/\tau_M}{Y+R}\hat{W}_F\right] + \frac{1}{\sigma - 1}\left[\frac{D}{Y+R}(\hat{N}_H + \hat{\Phi}_D) + \frac{M}{Y+R}\left(\hat{N}_F + \hat{\Phi}_M\right)\right]$$

$$+ \frac{D}{Y+R}\widehat{\left\{\frac{D}{W_H \tau_D}\right\}}\frac{\tau_D - 1}{\tau_D} + \frac{M}{Y+R}\widehat{\left\{\frac{M}{W_F \tau_M}\right\}}\frac{\tau_M - 1}{\tau_M} + \frac{X}{Y+R}\widehat{\left\{\frac{X}{W_H \tau_X}\right\}}\frac{\tau_X - 1}{\tau_X}. \qquad (26)$$

Subscripts H and F are added to identify home and foreign variables as appropriate. These terms combine into ToT and MC effects, together with the policy distortions, as before. However, the ToT effect (the first square bracket) is now the weighted sum of changes in export and import prices, both measured at international (border) prices, where the weights are the value of exports and imports, respectively, measured at border prices. Thus, an increase in foreign wages, W_F, reduces domestic welfare. The MC distortion effect (the second square bracket) depends on the change in the mass of domestic firms (N_H) and the selection of domestic and foreign firms into supplying the domestic market (Φ_D, Φ_M) as before, and now also on the change in the mass of foreign firms, N_F.

It is useful to compare this expression to papers analysing the large economy case. For example, Campolmi, Faldinger and Forlati (2014, henceforth CFF) study a large economy Krugman model with perfectly elastic labour supply, and discuss four possible effects of policy. The first is the MC distortion, operating via $\hat{N}_H$. The second is a production relocation effect (sometimes known as a home market effect and referred to by CFF as a delocation motive). This occurs as, in a full two country model, a shock will in general change both $\hat{N}_H$ and $\hat{N}_F$ (CFF has symmetric firms, so selection effects $\hat{\Phi}_D, \hat{\Phi}_M$ are absent). This production relocation effect changes welfare through its impact on the supply of varieties, i.e. interacting with the MC-distortion and impacting welfare through the second square brackets in (26). It has been extensively studied in the two-country Krugman model (for example Venables 1987, CFF, Ossa 2011, Bagwell and Staiger 2015). Thus, a home import tariff raises N_H and, in the full two-country model, reduces N_F. This may raise domestic welfare, as can be seen by noting that, in equation (26), a symmetric change ($\hat{N}_H$ = - $\hat{N}_F > 0$) will have a positive effect on welfare if domestic firms' sales are a larger share of expenditure than are imports, $D/(Y+R) > M/(Y+R)$.

The third mechanism analysed by CFF is a ToT effect. In their framework wages are constant so a ToT effect arises only through a change in domestic export taxes (as in the first term of equation (26)) or wage taxes (not present in our model). If labour supply to the MC sector is less than perfectly elastic, as in the present paper, a shock will also impact the ToT via changes in wages, $\hat{W}_H$ and $\hat{W}_F$. Thus, production relocation will impact the ToT as well as the MC-distortion. As we have seen the relative importance of ToT versus MC-distortion effects depends on the elasticity of labour supply to the MC sector. This is clearly stated (although not formally modelled) by Ossa (2011). While our approach is to endogenise the wage response via diminishing returns to scale (or equivalently the presence of a fixed factor) in the PC sector, Crozet and Trionfetti (2008), in a paper about the home market effect, achieve wage endogeneity by making an Armington assumption on the output of the PC sector.

CFF's fourth mechanism is called a fiscal-burden-shifting effect, and corresponds to the policy induced distortions in the final row of (26). It appears when more than one policy instrument is in use. As an example, if there is an export subsidy in place and a wage tax alters the quantity exported, then the changes in export subsidies should be taken into account when determining the effect of the wage tax.

This comparison illustrates the consistency of our analysis with the results of these papers, and also makes an important general point. Ultimate gains from policy in this class of models arise from MC-distortion and ToT effects, but these can be driven by different policy instruments and occur via different channels. Thus, ToT effects occur through export subsidies or, if labour is supplied inelastically, through a broader set of measures that change demand for labour in the MC sector and hence change wages. MC-distortion effects arise as the number of domestic firms change and also as the number of foreign firms supplying the domestic market may be affected by domestic policy. This arises in large economy models (production relocation), and also in the presence of firm level heterogeneity when, as we have shown, domestic policy will influence the number of foreign firms choosing to supply the domestic market[23]. We think it important that the numerous different effects and channels described in the literature are understood in terms of the ultimate mechanisms through which they change welfare.

Finally, our focus has been on welfare in the economy implementing the policy. An implication of our results is that the effect of policy in the domestic MC sector on other countries depends on the overall structure of the economy – and hence whether it just changes quantities, or also changes prices and the ToT. Building on this, a large open economy model could be used to analyse the response of foreign policy instruments to domestic policy. This is beyond the scope of the present paper, although we note the contributions of Ossa (2011) and Bagwell and Staiger (2015) to this issue, illustrating how relocation and terms-of-trade effects may play key roles in

[23] Costinot et al. (2016) confirms this result by pointing out that firm heterogeneity lowers trade protection if and only if there is active selection of foreign firms into exporting.

international trade policy competition. As should be clear from our discussion of CFF above, the policy motives are the same in our small open economy model as for large economies; however, with active policies in several countries, non-cooperative policy solutions will typically differ from the unilateral policies of our model.

7. Concluding comments

The model of trade and Dixit-Stiglitz monopolistic competition, with or without heterogeneous firms, is the workhorse model of trade theory. The structure is simple enough to yield explicit formulae for optimal policies, yet also complicated enough for the algebra involved in deriving these for the general case to explode. This has given rise to a literature of special cases and incomplete general understanding. The present paper has derived the optimal tax and tariff formulae in a model that encompasses these special cases and thereby draws out the underlying arguments for policy.

The main messages are that there are potential gains from using policy to support the home MC sector, either through a subsidy to firms' domestic sales or through trade taxes. The gains are driven by a combination of two effects, one through quantities, the other through prices. The quantity effect arises from the interaction of trade policy with the MC distortion; supporting the domestic industry increases the range of products on offer with beneficial variety effects. The price effect arises through the general equilibrium of the model; if labour supply to the MC sector is inelastic then supporting the MC sector raises wages and this brings a ToT improvement. The presence of heterogeneous productivity tempers results since the number of foreign varieties sold in the domestic market is endogenous and foreign reactions become more price elastic. However, this heterogeneity creates no qualitatively new arguments for or against policy interventions.

Our results encompass the findings of Demidova and Rodriguez-Clare (2009) and Felbermayr et al. (2013) for the one-sector economy. However, by decomposing the welfare effects, we highlight the mechanisms and draw out similarities with the older literature on homogeneous firms, e.g. Flam and Helpman (1987) and Venables (1982 and 1987). With more than one sector, there is a trade-off between terms-of-trade effects and variety effects, and we demonstrate how this depends on the elasticity of labour supply to the MC sector. Furthermore, our results show that although firm heterogeneity does not give qualitatively new arguments for policy interventions, the size of the optimal taxes or subsidies are affected. And by distinguishing between domestic and foreign selection effects, the importance of foreign firms' selection effects for product variety in the home market becomes clear in our results.

When comparing our results with those from studies of optimal policies in large economies, we conclude that the underlying forces are the same: gains from policies ultimately arise from the MC-distortion and the ToT effects. Finally, we note that our modelling of a small, open economy with firm heterogeneity captures important and realistic features in many markets, i.e. that domestic policies may affect the number of foreign suppliers choosing to service the market, even if the country is too small to have an impact on macro-conditions in the rest of the world.

Appendix 1. Expected profits and the mass of firms.

We assume that f_E is small enough for the expected profits of a domestic firm to be positive if $N = 0$, and large enough that expected profits are less than or equal to zero if the entire labour force is employed in the MC sector. There is a unique value of N at which a domestic entrant expects to break even if expected profits are monotonically declining with N. We establish this relationship using the comparative statics of section 3, with one extension. The left hand side of equation (1) is expected profits, which we now denote π. We take N as exogenous, all tax rates constant and equal to unity, and use (2') – (14') to investigate how π changes with N. From *Mathematica*,

$$\hat{\pi} = -\sigma\hat{N}\left[\frac{s_Y[\lambda_H s_D + \lambda_F(1-s_D)] + \mu s_D[s_D(1-\lambda_H) - s_Y] + \eta\mu s_D^2}{\lambda_H\{s_Y[\lambda_H s_D + \lambda_F(1-s_D)] + \mu s_D[s_D(1-\lambda_H) - s_Y]\} + \eta s_Y\{\lambda_F(1-s_D) + \lambda_H s_D\}}\right]$$

For the analyses of section 4.1 and 4.2 respectively,

$$\hat{\pi} = -\sigma\hat{N}\left[\frac{\mu s_D^2}{s_Y\{\lambda_F(1-s_D) + \lambda_H s_D\}}\right] < 0 \text{ and } \hat{\pi} = -\frac{\sigma\hat{N}}{\lambda_H} < 0.$$

Appendix 2. Comparative statics

Differentiation of the equilibrium conditions (1) – (14) gives the following equations with equation numbers matching level equations (1) – (14)

$$0 = (1-s_X)\{\hat{E} + (\sigma-1)\hat{P} - \sigma(\hat{W} + \hat{\tau}_D)\} - s_X\sigma(\hat{W} + \hat{\tau}_X). \quad (1')$$

$$(1-\sigma)\hat{\varphi}_D = \hat{E} + (\sigma-1)\hat{P} - \sigma(\hat{W} + \hat{\tau}_D), \quad \rightarrow \quad \hat{\Phi}_D = \frac{\lambda_H - \sigma}{\sigma}\{\hat{E} + (\sigma-1)\hat{P} - \sigma(\hat{W} + \hat{\tau}_D)\}, \quad (2')$$

$$(1-\sigma)\hat{\varphi}_X = -\sigma(\hat{W} + \hat{\tau}_X), \quad \rightarrow \quad \hat{\Phi}_X = (\sigma - \lambda_H)(\hat{W} + \hat{\tau}_X), \quad (3')$$

$$(1-\sigma)\hat{\varphi}_M = \hat{E} + (\sigma-1)\hat{P} - \sigma\hat{\tau}_M, \quad \rightarrow \quad \hat{\Phi}_M = \frac{\lambda_F - \sigma}{\sigma}\{\hat{E} + (\sigma-1)\hat{P} - \sigma\hat{\tau}_M\}. \quad (4')$$

Changes in the values of sales (domestic, export and imports) satisfy;

$$\hat{D} = \hat{E} + (\sigma-1)\hat{P} + (1-\sigma)(\hat{W} + \hat{\tau}_D) + \hat{N} + \hat{\Phi}_D, \quad (5')$$

$$\hat{X} = (1-\sigma)(\hat{W} + \hat{\tau}_X) + \hat{N} + \hat{\Phi}_X, \quad (6')$$

$$\hat{M} = \hat{E} + (\sigma-1)\hat{P} + (1-\sigma)\hat{\tau}_M + \hat{\Phi}_M. \quad (7')$$

The price index (8) changes according to:

$$\hat{P} = s_D\left[\hat{W} + \hat{\tau}_D + \frac{\hat{N} + \hat{\Phi}_D}{1-\sigma}\right] + (1-s_D)\left[\hat{\tau}_M + \frac{\hat{\Phi}_M}{1-\sigma}\right]. \tag{8'}$$

The change in government revenue is *dR* and, since *R* can equal zero, we use *dR/E* not *dR/R*.

$$\frac{dR}{E} = \left(\frac{\hat{D}(\tau_D - 1) + \hat{\tau}_D}{\tau_D}\right)s_D + \left(\frac{\hat{M}(\tau_M - 1) + \hat{\tau}_M}{\tau_M}\right)(1-s_D) + \left(\frac{\hat{X}(\tau_X - 1) + \hat{\tau}_X}{\tau_D}\right)\left(\frac{s_D s_X}{1-s_X}\right). \tag{9'}$$

From (10), employment in the MC sector changes according to[24]

$$\hat{L} + \hat{W} = (1-s_X)\left(\hat{D} - \hat{\tau}_D\right) + s_X\left(\hat{X} - \hat{\tau}_X\right). \tag{10'}$$

Turning to the general equilibrium, the change in the value of output and wages are

$$\hat{Y} = s_Y \hat{W}\ . \tag{11'}$$

$$\hat{W} = \hat{L}/\eta\ . \tag{12'}$$

The change in utility and expenditure are :

$$\hat{U} = (1-r)\hat{Y} + r\hat{R} - \mu\hat{P} = (1-r)s_Y\hat{W} + \mu dR/E - \mu\hat{P}\ , \tag{13'}$$

$$\hat{E} = (1-r)\hat{Y} + r\hat{R} = (1-r)s_Y\hat{W} + \mu dR/E\ . \tag{14'}$$

Appendix 3. Decomposition of the welfare effects of an import tariff.

Substituting (8'), (9') and (11') in (13') gives

$$\hat{U} = (1-r)s_Y\hat{W} + \mu\left(\frac{\hat{D}(\tau_D - 1) + \hat{\tau}_D}{\tau_D}\right)s_D + \mu\left(\frac{\hat{M}(\tau_M - 1) + \hat{\tau}_M}{\tau_M}\right)(1-s_D)$$
$$+ \mu\left(\frac{\hat{X}(\tau_X - 1) + \hat{\tau}_X}{\tau_D}\right)\left(\frac{s_D s_X}{1-s_X}\right) - \mu\left\{s_D\left[\hat{W} + \hat{\tau}_D + \frac{\hat{N} + \hat{\Phi}_D}{1-\sigma}\right] + (1-s_D)\left[\hat{\tau}_M + \frac{\hat{\Phi}_M}{1-\sigma}\right]\right\}.$$

Rearranging this and using $\mu s_D = D/(Y+R)$, $\mu(1-s_D) = M/(Y+R)$, and $\left[(1-r)s_Y - \mu s_D\right] = X/(Y+R)$, gives equation (20) of the text.

---

[24] With the Pareto distribution and equal price cost mark-ups in both markets, employment is proportional to the number of firms, so $\hat{L} = \hat{N}$, as can be seen from (5') and (6'), using (1'), (2') and (3').

Holding $\tau_D = \tau_X = 1$ gives equation (24), with changes in endogenous variables coming from differentiation of the system with respect to τ_M. The welfare change can be expressed as $\hat{U} = \{ToT + MC + TW\}\hat{\tau}_M$, where ToT is the terms-of-trade effects, MC is the variety effect, and TW corresponds to the tariff wedge. With the optimal tariff in place TW is set such that $ToT + MC + TW = 0$, giving (21) of the text. The expressions for ToT and MC are found by using *Mathematica* to find terms in (20) (evaluated at the second-best optimal tariff in the bottom row of Table 3B).

$$ToT = \frac{Ks_x(\lambda_F - 1)}{(\sigma - 1 + \mu)\{s_D[(1 - s_X)(\eta + 1) + s_X\lambda_H] + (1 - s_D)\lambda_F\} - \sigma\mu},$$

$$MC = \frac{K[\eta(1 - s_X) + s_X(\lambda_H - \lambda_F)] - \mu(1 - \mu)(1 - s_D)(\lambda_F - \sigma)}{(\sigma - 1 + \mu)\{s_D[(1 - s_X)(\eta + 1) + s_X\lambda_H)] + (1 - s_D)\lambda_F\} - \sigma\mu},$$

where $K \equiv \mu(\sigma - 1 + \mu)s_D(1 - s_D)$ and hence

$$ToT + MC = \frac{K[\eta(1 - s_X) + s_X(\lambda_H - 1)] - \mu(1 - \mu)(1 - s_D)(\lambda_F - \sigma)}{(\sigma - 1 + \mu)\{s_D[(1 - s_X)(\eta + 1) + s_X\lambda_H)] + (1 - s_D)\lambda_F\} - \sigma\mu}.$$

By inspection, ToT is positive (both the numerator and denominator are positive) while MC may be positive or negative. Equations (23 and (24) follow from these expressions.

References

Arkolakis, C., Costinot, A., Rodríguez-Clare, A., 2012. New trade models, same old gains? American Economic Review 102 (1), 94–130.

Bagwell, K., Staiger, R. 2015, Delocation and trade agreements in imperfectly competitive markets, Research in Economics, 69, 132-156

Baldwin, R.F., Forslid, R., 2010. Trade liberalization with heterogeneous firms. Review of Development Economics 14 (2), 161–176.

Campolmi, A., Fadinger, H., Forlati, C. 2014. Trade policy; home market effect versus terms of trade externality', Journal of International Economics, 93, 92-107.

Chaney, T., 2008. Distorted Gravity: The intensive and extensive margins of international trade. American Economic Review 98(4), 1707-1721.

Costinot, A., Rodriguez-Clare A., Werning, I. 2016. Micro to Macro: Optimal Trade Policy with Firm Heterogeneity, NBER working paper 21989

Crozet, M., Trionfetti, F., 2008, Trade costs and the Home Market Effect, Journal of International Economics, 76, 309-321.

Demidova, S. 2008. Productivity improvements and falling trade costs; boon or bane? International Economic Review 49 (4), 1437-62.

Demidova, S., Rodríguez-Clare, A., 2009. Trade policy under firm-level heterogeneity in a small economy. Journal of International Economics 78 (1), 100–112

Dhingra, S., Morrow J., 2012. Monopolistic Competition and Optimum Product Diversity under Firm Heterogeneity, London School of Economics, mimeograph.

Dixit, A.K., Stiglitz J.E. 1977. Monopolistic competition and optimum product diversity, American Economic Review 67, 297-308,

Felbermayr, G., Jung, B., Larch, M., 2013. Optimal tariffs, retaliation and the welfare loss from tariff wars in the Melitz model. Journal of International Economics, 89, 13-25.

Felbermayr, G., Jung, B. Larch, M., 2015. The welfare consequences of import tariffs; a quantitative perspective'. Journal of International Economics, 97, 295-309.

Flam, H., Helpman, E., 1987. Industrial policy under monopolistic competition. Journal of International Economics 22, 79–102.

Head, K., Mayer, T., 2015. Gravity Equations: Workhorse, Toolkit, and Cookbook. Handbook of International Economics, vol. 4.

Helpman, E., Krugman, P.R., 1989. Trade Policy and Market Structure. The MIT Press, Cambridge MA.

Krugman, P.R., 1980. Scale economies, product differentiation, and the pattern of trade. American Economic Review 70, 950–959.

Melitz, M.J., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica 71, 1695–1725.

Melitz, M.J., Redding S. 2015a. New Trade Models, New Welfare Implications. American Economic Review, 105, 1105-1146.

Melitz, M.J., Redding S., 2015b. Heterogeneous Firms and Trade. Handbook of International Economics, vol. 4.

Ossa, R., 2011. A 'new trade' theory of GATT/WTO negotiations. Journal of Political Economy 119, 122–152.

Venables, A.J., 1982. Optimal tariffs for trade in monopolistically competitive commodities. Journal of International Economics, 12, 225-241.

Venables, A.J., 1987. Trade and trade policy with differentiated products: a Chamberlinian–Ricardian model. The Economic Journal 97, 700–717.