Economies of Scale and Industrial Policy: A View from Trade *

Dominick Bartelme, University of Michigan

Arnaud Costinot, MIT

Dave Donaldson, MIT

Andres Rodriguez-Clare, UC Berkeley

December 2018

Abstract

When sector size goes up, does productivity go up as well? If it does, are productivity gains larger in some sectors than others? And if they are, what are the gains from industrial policies that subsidize these sectors at the expense of others? In this paper we develop a new empirical strategy to estimate economies of scale using trade data and provide answers to these questions. Across 2-digit manufacturing sectors, our baseline estimates of scale elasticities range from 0.07 to 0.25 and average 0.13. Viewed through the lens of a Ricardian model with external economies of scale, these estimates imply gains from optimal industrial policy that are around 0.61% on average across countries, a bit smaller than the gains from optimal trade policy.

*Contact: dbartelm@umich.edu, costinot@mit.edu, ddonald@mit.edu, andres@econ.berkeley.edu.

We are grateful to Susantu Basu, Jonathan Eaton and David Weinstein for comments, and to Daniel Haanwinckel and Walker Ray for outstanding research assistance.

1 Introduction

When sector size goes up, does productivity go up as well? If it does, are productivity gains larger in some sectors than others? And if they are, what are the gains from industrial policies that subsidize these sectors at the expense of others? The goal of our paper is to address these questions.

Section 2 presents our theoretical framework. We study a Ricardian economy with multiple sectors, each subject to external economies of scale. Our focus on this environment is motivated by its long intellectual history in the field, dating back to Graham's (1923) famous argument for trade protection, as well as the recent emergence of the Ricardian model as a workhorse model for quantitative work, as discussed by Eaton and Kortum (2012). Within each sector, external economies of scale may affect both the physical productivity of firms as well as the quality of the goods that they produce. This creates a rationale for Pigouvian taxation at the sector-level, which we refer to as industrial policy. In a competitive equilibrium, firms do not internalize that by increasing sector size, they raise its (quality adjusted) productivity. The optimal production subsidy, equal to the elasticity of productivity with respect to size, exactly compensates the firm for the marginal effect of its output decision on sector productivity.

Section 3 turns to identification. We show that external economies of scale are non-parametrically identified in this environment from commonly available trade data and standard orthogonality conditions. The starting point of our approach is the observation that in each destination and within each sector, trade flows from different origins reflect the optimal demand for labor services from these countries. Provided that this demand system is invertible, changes in trade flows therefore reveal changes in the effective prices of these services. Once the prices of labor services have been revealed, we can estimate external economies of scale by measuring the extent to which an exogenous increase in sector size lowers such prices.

Section 4 imposes parametric restrictions to implement the previous general strategy. Within each sector, we assume that (i) productivity is a log-linear function of size, so that we have constant *scale elasticities*, and that (ii) the demand for labor services from different countries has Constant Elasticity of Substitution (CES), so that we have constant *trade elasticities*. Both parametric restrictions are satisfied by the multi-sector gravity models with external economies of scale analyzed in Kucheryavyy et al. (2017). Under these assumptions, the (log of the) price of labor services from a country is proportional to (the log of) its sector size, with a slope given by the scale elasticity; and the revealed (log of the) price of labor services is proportional to (the log of) its bilateral exports, with a slope given by the inverse of the trade elasticity. Given existing estimates of sector-level trade elasticities in the literature, we can therefore estimate sector-level scale elasticities using a log-linear regression of bilateral exports, adjusted by the trade elasticity, on sector size.

Since exogenous productivity differences across countries and sectors affect both sector size and bilateral exports, identification requires an instrument that is positively correlated with sector size yet uncorrelated with idiosyncratic productivity shocks. To construct such an instrument, we first estimate the upper-level elasticity of substitution between goods from different sectors. Given an estimate of this elasticity, we then compute the demand residuals that rationalize observed expenditure shares across sectors and countries. Under the assumption that idiosyncratic productivity shocks are uncorrelated with sector-level demand shocks, the product of these demand residuals and country's population provides a valid instrument.

Not surprisingly, this leads to a very strong first stage, since both domestic expenditure shares and country size matter for sector size. Reassuringly, in every sector the IV estimate is lower than the OLS estimate, as would be expected if sector size responds positively to productivity. Our findings point to positive and significant scale elasticities in manufacturing sectors. As mentioned above, these range from 0.07 to 0.25, with an average of 0.13.

Section 5 uses our empirical estimates to compute the welfare gains from industrial policy. Our baseline analysis focuses on a small open economy that can only affect the price of its own good relative to goods from other countries: relative prices in the rest of the world, employment, and expenditure across sectors are taken as exogenously given by its government. Although external economies of scale are large, gains from industrial policy are only 0.61% for the average country. This is a bit smaller than the average gains from optimal trade policy implied by the same model.

Our analysis is related to a large empirical literature on the estimation of production functions in industrial organization and macroeconomics, see Ackerberg et al. (2007) and Basu (2008). Compared to the former, we make no attempt at estimating internal economies of scale at the firm-level. Rather, we focus on external economies at the sector-level, which sector-level trade flows reveal. Our focus on economies of scale at the sector-level is closer in spirit to Caballero and Lyons (1992) and Basu and Fernald (1997). A key difference between our approach and theirs is that we do not rely on measures of real output, or price indices, collected by statistical agencies. Instead, we use estimates of the demand for foreign inputs, as in Adao et al. (2017), to infer the effective prices for inputs. This provides a theoretically-grounded way to adjust for quality differences across origins within the same sector. We come back to these issues in Section 3.2.

The general idea of using trade data to infer economies of scale bears a direct relationship to empirical tests of the home-market effect; see e.g. Davis and Weinstein (2003), Head and Ries (2001), and Costinot et al. (2016). Indeed, a home-market effect, that is, a positive effect of demand on exports, implies the existence of economies of scale at the sector level. Our empirical strategy is also closely related to previous work on revealed comparative advantage; see e.g. Costinot et al. (2012) and Levchenko and Zhang (2016). The starting point of these papers, like ours, is that trade flows contain information about costs, a point also emphasized by Antweiler and Trefler (2002).

A large literature in international trade uses gravity models for counterfactual analysis. As discussed by Costinot and Rodríguez-Clare (2013) and Kucheryavyy et al. (2017), the quantitative predictions of these models hinge on two key elasticities: trade elasticities and scale elasticities. While the former have received significant attention in the empirical literature, as discussed in Head and Mayer (2013), the latter have not. Scale economies, when introduced in gravity models, are instead indirectly calibrated using information about the elasticity of substitution across goods in monopolistically competitive environments; see e.g. Balistreri et al. (2011). One of the goals of our paper is to offer more direct and credible evidence about the extent of sector-level economies of scale.

Finally, while a number of theoretical and empirical papers have discussed the rationale and potential consequences of industrial policy, as reviewed in Harrison and Rodríguez-Clare (2010) , there have been few attempts at connecting theory and data. A notable exception is the work of Lashkaripour and Lugovskyy (2018). Their quantitative exploration of the gains from industrial policy is very similar to ours in spirit. The main substantial difference between their paper and ours is the empirical strategy used to estimate scale elasticities. They study a monopolistically competitive environment à la Krugman (1980) where the elasticity of substitution between domestic varieties may differ from the elasticity of substitution between domestic and foreign varieties. In this model, the scale elasticity coincides with the elasticity of substitution between domestic varieties, whereas the trade elasticity coincides with the elasticity of substitution between domestic and foreign varieties. Using this particular feature of the model, they can jointly infer scale and trade elasticities by estimating the elasticity of firm-level exports with respect to firm-level prices—which uncovers the trade elasticity—as well as their elasticity with respect firm-level export shares from a given origin country—which uncovers the difference in substitutability between domestic and foreign varieties, and so, according to their model, the scale elasticity. In contrast, our empirical strategy directly identifies scale elasticities from the responses of sector-level exports to changes in employment across countries caused by variation in domestic demand.

2 Theory

2.1 Environment

Consider an economy comprising many origin countries, indexed by $i = 1, ..., I$, many destination countries, indexed by $j = 1, ..., J$, and many sectors, indexed by $k = 1, ..., K$. Each sector itself comprises many goods, indexed by ω.

Technology. Technology is Ricardian. In any origin country i, the same composite input, equipped labor, is used to produce all goods in all sectors.[1] We let L_i denote the fixed supply of labor in country i. For any sector k, output of good ω in country i that is available for consumption in country j is given by

$$q_{ij,k}(\omega) = A_{ij,k}(\omega) l_{ij,k}(\omega),$$

where $l_{ij,k}(\omega)$ denotes the amount of labor used by firms from an origin country i to produce and deliver good ω to a destination country j.[2] Transportation costs, if any, are reflected in $A_{ij,k}(\omega)$. At the sector-level, production may be subject to economies of scale,

$$A_{ij,k}(\omega) = \alpha_{ij,k}(\omega) A_{ij,k} E_k^A(L_{i,k}),$$

where $L_{i,k} = \sum_j \int l_{ij,k}(\omega) d\omega$ is the total amount of labor used in country i and sector k. For expositional purposes, we shall simply refer to $L_{i,k}$ as sector size.

Preferences. There is a representative agent with weakly separable preferences in each country. The utility of the representative agent in a destination country j is given by

$$U_j = U_j(U_{j,1}, .., U_{j,K}), \tag{1}$$

with $U_{j,k}$ the subutility associated with goods from sector k,

$$U_{j,k} = U_{j,k}(\{B_{ij,k}(\omega) q_{ij,k}(\omega)\}).$$

[1] This rules out cross-sectoral differences in either factor intensity or input-output linkages in our baseline analysis.

[2] The above specification assumes constant returns to scale at the good level, but does not require constant returns to scale at the firm level. As is well understood, constant returns to scale at the good level ω may reflect the free entry of heterogeneous firms, each subject to decreasing returns to scale, as in Hopenhayn (1992). Appendix A.1 makes that point explicitly.

$q_{ij,k}(\omega)$ denotes the total amount of good ω from sector k produced in country i and sold to consumers in country j and $B_{ij,k}(\omega)$ is an origin-destination-sector-specific taste shock that captures quality differences. We assume that subutility $U_{j,k}$ is homothetic, that standard Inada conditions hold, and that demand for goods within a sector satisfies the connected substitutes property, as defined in Arrow and Hahn (1971). We also allow quality to be affected by sector size,

$$B_{ij,k}(\omega) = \beta_{ij,k}(\omega) B_{i,k} E_k^B(L_{i,k}).$$

Taxes. There are three types of taxes in all countries. Production in a given sector k may be subject to an ad-valorem production subsidy, $s_{j,k}$, which creates a wedge between the prices faced by firms and consumers in country j. Imports and exports in a given sector k may also be subject to an import tariff, $t^m_{ij,k}$, and an export tax, $t^x_{ji,k}$. The first trade tax creates a wedge between the price paid by consumers in country j and the price received by firms in country $i \neq j$, whereas the second creates a wedge between the price received by firms in country j and the price paid by consumers in country i. Net revenues from taxes and subsidies are rebated through a lump-sum transfer, T_j, to the representative agent in country j.

2.2 Competitive Equilibrium

We focus on a competitive equilibrium with external economies of scale. In equilibrium, consumers maximize utility taking as given good prices, wages, taxes, and the size of each sector; firms maximize their profits, also taking as given good prices, wages, taxes, and the size of each sector; and all markets clear. The formal definition of a competitive equilibrium can be found in Appendix A.2.

To prepare our analysis of optimal policy, it is convenient to focus on the exchange of labor services between countries, as in Adao, Costinot and Donaldson (2017). Let $L_{ij,k}$ denote the demand, in efficiency units, for labor from country i in country j within a given sector k, and let $V_j(\{L_{ij,k}\}_{i,k})$ denote the utility of the representative agent in country j associated with a given vector of input demand,

$$V_j(\{L_{ij,k}\}_{i,k}) \equiv \max_{\{q_{ij,k}(\omega), l_{ij,k}(\omega)\}} U_j(\{U_{j,k}(\{\beta_{ij,k}(\omega) q_{ij,k}(\omega)\}_{i,\omega})\}_k)$$
$$q_{ij,k}(\omega) \leq \alpha_{ij,k}(\omega) l_{ij,k}(\omega) \text{ for all } \omega,\ i, \text{ and } k,$$
$$\int l_{ij,k}(\omega) d\omega \leq L_{ij,k} \text{ for all } i \text{ and } k.$$

In a competitive equilibrium, the labor services demanded by country j from different origins ana d sectors, $\{L_{ij,k}\}_{i,k}$, the labor services exported by country j towards different destinations, $\{L_{ji,k}\}_{i\neq j,k}$, and the sector sizes in country j, $\{L_{j,k}\}_k$, must solve

$$\max_{\{\tilde{L}_{ij,k}\}_{i,k},\{\tilde{L}_{ji,k}\}_{i\neq j,k},\{\tilde{L}_{j,k}\}_k} V_j(\{\tilde{L}_{ij,k}\}_{i,k}) \tag{2a}$$

$$\sum_{i\neq j} c_{ij,k}(1+t^m_{ij,k})\tilde{L}_{ij,k} \leq \sum_{i\neq j} c_{ji,k}(1-t^x_{ji,k})\tilde{L}_{ji,k} + T_j, \tag{2b}$$

$$\sum_i \eta_{ji,k}\tilde{L}_{ji,k} \leq (1+s_{j,k})E_k(L_{j,k})\tilde{L}_{j,k}, \text{ for all } k, \tag{2c}$$

$$\sum_k \tilde{L}_{j,k} \leq L_j, \tag{2d}$$

where $E_k(L_{j,k}) \equiv E^A_k(L_{j,k})E^B_k(L_{j,k})$ captures the joint effect of external economies of scale on the supply and demand sides; $\eta_{ji,k} \equiv 1/(A_{ij,k}B_{ij,k})$ captures systematic productivity and quality differences; and $c_{ij,k} \equiv \eta_{ij,k}w_i/[(1+s_{i,k})(1-t^x_{ij,k})E_k(L_{i,k})]$ corresponds to the effective price of labor from country i in country j and sector k, that is the wage w_i adjusted by the export tax $t^x_{ij,k}$, the production subsidy $s_{i,k}$, the systematic productivity and quality differences $\eta_{ij,k}$, and the external economies of scale $E_k(L_{i,k})$.

Equation (2b) is the trade balance condition. It states that the value of inputs imported by country j is no greater than the value of its exports. Equations (2c) captures technological constraints; it states total demand for inputs across destinations i, adjusted by the bilateral exogenous efficiency term $\eta_{ji,k}$ can be no greater than the total supply, in efficient units, in country j and sector k. The term $E_k(L_{j,k})$ reflects the fact that because of economies of scale, an increase in sector size leads either to larger quantities or higher quality goods being produced with a given amount of inputs, and hence an increase the number of inputs supplied in efficiency units. Since firms do not internalize this effect, $L_{j,k}$ is taken as given in the above problem. Equation (2d) is the labor market clearing condition; it states that the sum of labor allocated across sectors k can be no greater than the total labor supply in country j.

For future reference, we let $x_{ij,k} = [(1+t^m_{ij,k})c_{ij,k}L_{ij,k}]/(\sum_{i'}[(1+t^m_{i'j,k})c_{i'j,k}L_{i'j,k}])$ denote the share of expenditure in destination j on labor services from country i in sector k. In a Ricardian environment, this also corresponds to the share of expenditure on goods from sector k produced in country i. In what follows, we shall simply refer to $\{x_{ij,k}\}$ as trade shares. As shown in Appendix A.3, trade shares in a perfectly competitive equilibrium can be expressed as

$$x_{ij,k} = \chi_{ij,k}((1+t^m_{1j,k})c_{1j,k}, ..., (1+t^m_{Ij,k})c_{Ij,k}), \tag{3}$$

where $\chi_{j,k} \equiv (\chi_{1j,k}, ..., \chi_{Ij,k})$ is homogeneous of degree zero, invertible, and a function of, and only of, $U_{j,k}$, $\{\alpha_{i,k}(\omega)\}$ and $\{\beta_{ij,k}(\omega)\}$. This is the sector-level counterpart of factor demand in Adao, Costinot and Donaldson (2017); it will play a key role in our identification of external economies of scale.

2.3 Optimal Policy

We now turn to the analysis of optimal policy. By optimal, we mean the vector of trade and production taxes or subsidies that maximize the utility of the representative agent in a given country j, taking as given policies in other countries. We further assume that country j is a small open economy that can only affect the price of its own good relative to goods from other countries: relative prices in the rest of the world, sector-level employment, and sector-level expenditure are taken as exogenously given by its government. As argued below, the restriction to a small open economy is irrelevant for the structure of optimal industrial policy, which is our main focus in this paper.

We proceed in two steps. First, we consider the problem of a government that can directly choose consumption and production in order to maximize utility in country j. Second, we show how the solution to that planning problem can be decentralized through sector-level production and trade taxes.

Government Problem. The problem of country j's government is

$$\max_{\{\tilde{L}_{ij,k}\}_{i,k}, \{\tilde{L}_{ji,k}\}_{i\neq j,k}, \{\tilde{L}_{j,k}\}_k} V_j(\{\tilde{L}_{ij,k}\}_{i,k}) \tag{4a}$$

$$\sum_{i\neq j,k} c_{ij,k}\tilde{L}_{ij,k} \leq \sum_{i\neq j,k} c_{ji,k}(\tilde{L}_{ji,k})\tilde{L}_{ji,k}, \tag{4b}$$

$$\sum_{i} \eta_{ji,k}\tilde{L}_{ji,k} \leq E_k(\tilde{L}_{j,k})\tilde{L}_{j,k}, \text{ for all } k, \tag{4c}$$

$$\sum_{k} \tilde{L}_{j,k} \leq L_j. \tag{4d}$$

There are two key differences between problems (2) and (4).

First, country j's government internalizes sector-level economies of scale, $E_k(\tilde{L}_{j,k})$, whereas firms and consumers do not. This explains why $E_k(\tilde{L}_{j,k})$ depends on the choice variable, $\tilde{L}_{j,k}$, rather than the equilibrium sector size, $L_{j,k}$, as in equation (2c). This creates a rationale for Pigouvian taxation, that is production subsidies, $\{s_{j,k}\}$, that may be non-zero at the optimum.

Second, the government recognizes its market power on foreign markets, whereas firms and consumers do not. In the small open economy case that we focus on, country j's government takes import prices, $c_{ij,k} \equiv \eta_{ij,k} w_i / [(1 - t^x_{ij,k}) E_k(L_{i,k})]$, as given for any origin country $i \neq j$, but it internalizes the fact that export prices, $c_{ji,k}(\tilde{L}_{ji,k})$, are a function of its own exports, $\tilde{L}_{ji,k}$, implicitly given by the solution to

$$\chi_{ji,k}((1+t^m_{1i,k})c_{1i,k}, ..., (1+t^m_{Ii,k})c_{Ii,k}) = \frac{(1+t^m_{ji,k})c_{ji,k}\tilde{L}_{ji,k}}{\sum_{i' \neq j}(1+t^m_{i'i,k})c_{i'i,k}L_{i'i,k} + (1+t^m_{ji,k})c_{ji,k}\tilde{L}_{ji,k}}, \quad (5)$$

with the equilibrium costs of other exporters, $\{c_{i'i,k}\}_{i' \neq j}$, as well as their exports of labor services, $\{L_{i'i,k}\}_{i',k}$, taken as given. The fact that firms and consumers ignore such effects creates a rationale for export taxes, $\{t^x_{ji,k}\}_{i,k}$, that manipulate country j's terms-of-trade.

Implementation. To characterize the structure of optimal policy, we compare the solutions to (2) and (4) and derive necessary conditions on production subsidies and trade taxes such that the two solutions coincide.

Consider first the solution to (4). The first-order conditions with respect to $\{\tilde{L}_{j,k}\}_k$, $\{\tilde{L}_{ji,k}\}_{i \neq j,k}$, and $\{\tilde{L}_{ij,k}\}_{i,k}$ imply

$$\begin{aligned}
[E'_k(L_{j,k})L_{j,k} + E_k(L_{j,k})]\rho_{j,k} &= \rho_j, \\
\lambda_j[c'_{ji,k}(L_{j,k})L_{ji,k} + c_{ji,k}(L_{j,k})] &= \eta_{ji,k}\rho_{j,k}, \\
dV_j(\{L_{ij,k}\}_{i,k})/dL_{ij,k} &= \lambda_j c_{ij,k}, \text{ if } i \neq j, \\
dV_j(\{L_{ij,k}\}_{i,k})/dL_{ij,k} &= \eta_{ij,k}\rho_{j,k}, \text{ if } i = j.
\end{aligned}$$

where λ_j, $\{\rho_{j,k}\}$ and ρ_j denote the values of the Lagrange multipliers associated with constraints (4b)-(4d) at the optimal allocation.

Now suppose that the same allocation arises at the solution to (2). The first-order conditions associated with this problem imply

$$\begin{aligned}
&(1+s_{j,k})E_k(L_{j,k})\rho^e_{j,k} = \rho^e_j, \\
&\lambda^e_j(1-t^x_{ji,k})c_{ji,k}(L_{j,k}) = \eta_{ji,k}\rho^e_{j,k}, \\
&dV_j(\{L_{ij,k}\}_{i,k})/dL_{ij,k} = \lambda^e_j(1+t^m_{ij,k})c_{ij,k}(L_{i,k}), \text{ if } i \neq j, \\
&dV_j(\{L_{ij,k}\}_{i,k})/dL_{ij,k} = \eta_{ij,k}\rho^e_{j,k}, \text{ if } i = j,
\end{aligned}$$

where λ^e_j, $\{\rho^e_{j,k}\}$ and ρ^e_j denote the values of the Lagrange multipliers associated with constraints (2b)-(2d). A comparison of these two sets of first-order conditions leads to the following proposition.

Proposition 1. *For a small open economy j, the unilaterally optimal policy consists of a combination of production and trade taxes such that, for some $s_j, t_j > -1$,*

$$1 + s_{j,k} = (1 + s_j)(1 + \frac{d \ln E_k}{d \ln L_{j,k}}), \text{for all } k,$$
$$1 - t^x_{ji,k} = (1 + t_j)(1 + \frac{d \ln c_{ji,k}}{d \ln L_{ji,k}}), \text{for all } i \text{ and } k,$$
$$1 + t^m_{ij,k} = 1 + t_j, \text{for all } i \text{ and } k.$$

The two shifters, s_j and t_j, reflects two distinct sources of tax indeterminacy in our model. First, since labor supply is perfectly inelastic, a uniform production tax or subsidy s_j only affects the level of input prices in country j, but leaves the equilibrium allocation unchanged. Second, a uniform increase in all trade taxes again affects the level of prices in country j, but leaves the trade balance condition and the equilibrium allocation unchanged, an expression of Lerner Symmetry. In the rest of our analysis, we normalize both s_j and t_j to zero.

It is worth noting that while we have focused on the case of a small open economy, this restriction is only relevant for the structure of optimal trade policy, which would depend, in general, on the entire vector of imports and exports by country j. The optimal Pigouvian tax, in contrast, is always given by $\frac{d \ln E_k}{d \ln L_{j,k}}$. Formally, this can be seen easily from the fact that the technological constraints (2c) and (4c) would be unchanged in the case of a large open economy, as described in Appendix A.4.

3 Identification

Section 2 highlights the importance of two structural objects for optimal policy design: (i) $\chi_{j,k}$, which determines trade shares in the rest of the world and, in turn, export prices for country j; and (ii) E_k, which determines external economies of scale across sectors. Under the assumption that demand in each sector satisfies the connected substitutes property, $\chi_{j,k}$ is invertible and non-parametrically identified under standard orthogonality conditions, as discussed in Adao, Costinot and Donaldson (2017). Our goal in this section is to provide conditions under which, given knowledge of $\chi_{j,k}$, E_k is non-parametrically identified as well.

The basic idea is to start by inverting demand in order to go from the trade shares, that are observed, to the effective input prices, that are not. Once the prices having been inferred, we can then estimate external economies of scale by measuring the extent to which an exogenous increase in sector size lowers such prices.

3.1 Non-Parametric Identification of External Economies of Scale

Formally, let $\chi_{ij,k}^{-1}(x_{1j,k}, ..., x_{Ij,k})$ denote the effective price of input from country i in country j and sector k, up to some normalization. For any pair of origin countries, i_1 and i_2, and any sector k_1, equation (3) implies

$$\ln \frac{\chi_{i_1 j,k_1}^{-1}(x_{1j,k_1}, ..., x_{Ij,k_1})}{\chi_{i_2 j,k_1}^{-1}(x_{1j,k_1}, ..., x_{Ij,k_1})} = \ln \frac{E_{k_1}(L_{i_2,k_1})}{E_{k_1}(L_{i_1,k_1})} + \ln \frac{w_{i_1}}{w_{i_2}} + \ln \frac{\tilde{\eta}_{i_1 j,k_1}}{\tilde{\eta}_{i_2 j,k_1}},$$

with $\tilde{\eta}_{ij,k} \equiv [\eta_{ij,k}(1 + t_{ij,k}^m)]/[(1 - t_{ij,k}^x)(1 + s_{i,k})]$. Taking a second difference relative to another sector k_2, we therefore have

$$\ln \frac{\chi_{i_1 j,k_1}^{-1}(x_{1j,k_1}, ..., x_{Ij,k_1})}{\chi_{i_2 j,k_1}^{-1}(x_{1j,k_1}, ..., x_{Ij,k_1})} - \ln \frac{\chi_{i_1 j,k_2}^{-1}(x_{1j,k_2}, ..., x_{Ij,k_2})}{\chi_{i_2 j,k_2}^{-1}(x_{1j,k_2}, ..., x_{Ij,k_2})}$$
$$= \ln \frac{E_{k_1}(L_{i_2,k_1})}{E_{k_1}(L_{i_1,k_1})} - \ln \frac{E_{k_2}(L_{i_2,k_2})}{E_{k_2}(L_{i_1,k_2})} + \ln \frac{\tilde{\eta}_{i_1 j,k_1}}{\tilde{\eta}_{i_2 j,k_1}} - \ln \frac{\tilde{\eta}_{i_1 j,k_2}}{\tilde{\eta}_{i_2 j,k_2}}. \tag{6}$$

Given two origin countries, i_1 and i_2, two sectors, k_1 and k_2, and a destination country j, equation (6) is a nonparametric regression model with endogenous regressors and a linear error term,

$$y = h(l) + \epsilon,$$

where the endogenous variables, y and l, the function to be estimated, $h(\cdot)$, and the error term, ϵ, are given by

$$y \equiv \ln \frac{\chi_{i_1 j,k_1}^{-1}(x_{1j,k_1}, ..., x_{Ij,k_1})}{\chi_{i_2 j,k_1}^{-1}(x_{1j,k_1}, ..., x_{Ij,k_1})} - \ln \frac{\chi_{i_1 j,k_2}^{-1}(x_{1j,k_2}, ..., x_{Ij,k_2})}{\chi_{i_2 j,k_2}^{-1}(x_{1j,k_2}, ..., x_{Ij,k_2})},$$
$$l \equiv (L_{i_1,k_1}, L_{i_2,k_1}, L_{i_1,k_2}, L_{i_2,k_2}),$$
$$h(l) \equiv \ln \frac{E_{k_1}(L_{i_2,k_1})}{E_{k_1}(L_{i_1,k_1})} - \ln \frac{E_{k_2}(L_{i_2,k_2})}{E_{k_2}(L_{i_1,k_2})},$$
$$\epsilon \equiv \ln \frac{\tilde{\eta}_{i_1 j,k_1}}{\tilde{\eta}_{i_2 j,k_1}} - \ln \frac{\tilde{\eta}_{i_1 j,k_2}}{\tilde{\eta}_{i_2 j,k_2}}.$$

Economically speaking, the endogeneity of the regressors, $E[\epsilon|l] \neq 0$, simply reflects the fact that sectors with higher productivity, higher quality, or lower trade costs in a given origin country will also tend to have larger sizes. The nonparametric identification of $h(\cdot)$

therefore requires a vector of instruments.

Newey and Powell (2003) provide general conditions for nonparametric identification in such environments. Specifically, if there exists a vector of instruments z that satisfies the exclusion restriction, $E[\epsilon|z] = 0$, as well as the completeness condition, $E[g(l)|z] = 0$ implies $g = 0$ for any g with finite expectation, then $h(\cdot)$ is nonparametrically identified. As shown in Appendix A.5, once $h(\cdot)$ is identified, both E_{k_1} and E_{k_2} are also identified, up to a normalization. In the next section, we will propose such a vector of instruments and use it to estimate sector-level external economies of scale.

3.2 Discussion

So far we have established that one can use data on trade shares, $\{x_{ij,k}\}$, and sector sizes, $\{L_{i,k}\}$, to identify external economies of scale in a perfectly competitive environment. An obvious benefit of this empirical strategy is that trade data are easily available for a large number of countries, sectors, and years. Output data, however, may be available as well. If so, one could use micro-level data, that records firm's physical output and input use, in order to estimate firm-level production functions directly,

$$q = E_k^A(L_{i,k})F(l, \phi),$$

with ϕ is an index of productivity that may vary across firms producing the same good ω in country i and sector k, as discussed further in Appendix A.1.

One could also use macro-level data, that records sector-level quantity indices for real output and real input uses, in order to estimate sector-level production functions. Before turning to our empirical analysis, we briefly discuss the relative costs and benefits of these alternative empirical strategies. We focus our discussion on differences in terms of robustness—that is, the strength of the assumptions required for inferences about the magnitude of external economies of scale to be valid—as well as data requirements.

Perfect versus Imperfect Competition The estimation of production functions, either using micro or macro data, does not require any assumption on good market structure. With output data and exogenous variation in input use, one can directly estimate the elasticity of output with respect to input, and hence economies of scale, regardless of whether good markets are perfectly competitive or not. In contrast, the nonparametric identification of external economies of scale in Section 3 is conducted under the assumption of perfect competition. Under this assumption, prices are equal to unit costs. This allows us to infer how variation in sector sizes affects costs, and hence economies of scale, by estimating how the variation in sector sizes affects prices, as revealed by trade shares.

The previous discussion might suggest that perfect competition is critical for our empirical strategy. In an economy where the pass-through from costs into prices is incomplete, one might expect our approach to systematically misinterpret changes in markups as changes in costs. This is not the case.

This is best seen through an extreme example. Consider an economy where production is as described in Section 2.1, but there is now an imperfectly competitive retail sector that buys goods at marginal costs and sell them at a profit. We assume that retailers take sector-level expenditure as given and that there are no taxes. In this economy, retailers will impose different markups on different goods,

$$p_{ij,k}(\omega) = \frac{\mu_{ij,k}(\omega) w_i}{\alpha_{ij,k}(\omega) A_{ij,k} E_k^A(L_{i,k})}.$$

However, as we formally demonstrate in Appendix A.6, markups in sector k and country j will be a function of $(c_{1j,k}, ..., c_{Ij,k})$, and hence we can still express trade shares as a function of input prices, $\chi_{j,k}(c_{1j,k}, ..., c_{Ij,k})$. Thus, given knowledge of $\chi_{j,k}$, external economies are nonparametrically identified under the same condition as under perfect competition. The reason why the lack of market power by firms is not critical for our empirical strategy can be understood as follows. If we have access to an observable exogenous shifter of $c_{ij,k}$, like freight costs, which is what the knowledge of $\chi_{j,k}$ requires, then one can compare the elasticity of trade shares with respect to this observable cost shifter to the elasticity of trade shares with respect to sector size. The ratio of the latter to the former then identifies by how much sector sizes has affected costs, i.e. the extent of economies of scale. Whether or not good prices are equal to their marginal costs, the exact same inference remains valid.[3]

Physical Productivity versus Quality The economic environment of Section 2.1 features two types of external economies of scale. As a sector expands, both physical productivity and quality may change, as captured by $E_k^A(L_{i,k})$ and $E_k^B(L_{i,k})$, respectively.

By using micro data, one could estimate these two functions sector by sector. Specifically, one could first use data on firm's physical output and input use to estimate firm-

[3]This establishes that perfect competition is not critical for our empirical strategy, not that there does not exist imperfectly competitive models under which variation in markups would affect our inferences about the magnitude of external economies of scale. Costinot et al. (2016) discuss such an example. In their model, an increase in the number of firms producing in a given origin country and sector lowers the markup charged by those firms everywhere, leading to a decrease in the prices faced by importing countries, absent any external economies of scale.

level production functions, $F(l, \phi)$. Given such estimates, one could then infer $E_k^A(L_{i,k})$ by investigating how much of the firms' productivity residuals can be explained by sector size. Similarly, one could use data on firms' physical output and prices to estimate the demand for all goods within a sector and then infer $E_k^B(L_{i,k})$ by estimating how much of the demand residuals can be explained by sector size.

Compared to this strategy, our approach proposes to: (i) fold the estimation of firm-level production functions and demand functions into a single object, the demand for inputs from country i in sector s; (ii) recover the quality adjusted price of these inputs by inverting that demand system; and (iii) estimate the relationship between quality-adjusted prices and sector sizes. The main benefit of our approach is in terms of data requirements. All we need are data on sector-level trade flows, sector sizes, and an instrument for those. While our approach does not allow us to separately identify $E_k^A(L_{i,k})$ and $E_k^B(L_{i,k})$, it allows us to estimate the combination of these joint effects, $E_k(L_{i,k}) = E_k^A(L_{i,k})E_k^B(L_{i,k})$, which is all that will matter for optimal industrial policy.

In this regard, our approach is similar to the one that would use macro data, on quantity and price indices, in order to estimate sector-level economies of scale. Such an approach consists in estimating directly the impact of exogenous changes in sector sizes, $L_{i,k}$, on a sector-level quantity index, $Q_{i,k}$. Provided that price indices used to go from revenue to real output properly adjusts for quality, this alternative empirical strategy would also identify the joint effect of sector sizes on physical productivity and quality. The key difference between this macro approach and ours therefore boils down to the the nature of the quality adjustment. In our case, it derives from the estimation of demand for inputs from different countries and the associated residuals. In the case of the macro approach, it is left to the statistical agency in charge of computing price deflators.[4]

Internal versus External Economies of Scale As we have already noted, our model is consistent with the existence of internal economies of scale at the firm-level, provided that there is free entry in the production of each good, as in Hopenhayn (1992). If so, as the total number of workers employed to produce a good ω increases, the measure of entering firms increases in a proportional manner, while the number of workers per firm remains unchanged, making firm-level economies of scale irrelevant for our results.

[4]The distinction here is potentially more severe than the distinction between an exact price index, given some specific assumptions about demand, and a first-order approximation, that would be valid regardless of whether these specific assumptions hold or not. In the economic environment that we consider in Section 2.1, there may not exist a single-output technology at the country-sector level. The reason is that within a sector, different goods may be sold by the same country to different destinations. In such cases, there is no theoretically grounded expenditure function that the measured price index would be a first-order approximation to.

Absent free entry, good-level production functions may no longer be constant returns and economies of scale estimated at the sector-level may therefore reflect a mixture of both internal and external economies of scale. To control for internal economies of scale, without assuming that they necessarily vanish at the good level, one would need micro data. This is the same issue that one faces when estimating sector-level production functions.

To sum up, the main cost of our approach is that it requires restrictions on market structure and constant returns to scale at the good level. The lack of market power can be relaxed substantially, but the lack of internal economies (or diseconomies) of scale at the good (though not firm) level is critical to identify economies of scale as external. The main benefit of our approach is that it only requires commonly available data on trade flows and sector sizes as well as a simple, theoretically-consistent way to control for quality differentiation across goods.

4 Parameter Estimation

The results in Section 3 demonstrate how knowledge of external economies of scale could be obtained—that is, how the function E_k is nonparametrically identified from conventional data and exogenous variation. In this section we describe the empirical procedure that we use to obtain estimates of this function, before using them in Section 5 to assess the efficacy of optimal policy.

4.1 Parametric Restrictions

The results presented in Section 3 are asymptotic in nature. They answer the question of whether, in theory, one could point-identify external economies of scale, E_k, in all sectors with a dataset that includes an infinite sequence of economies. In this context, we have established that given an exogenous shifter of sector sizes, one can identify external economies of scale by tracing out the impact of changes in sector sizes on prices, as revealed by changes in equilibrium trade shares.

In practice, realistic datasets only include a small number of observations. For example, as we discuss below, the dataset we use here includes only 4 time periods and 61 countries. So, estimation inevitably needs to proceed parametrically. In the rest of our analysis, we impose the following functional-form assumptions at all times t:

$$\chi^t_{ij,k}((1+t^{m,t}_{1j,k})c^t_{1j,k}, ..., (1+t^{m,t}_{Ij,k})c^t_{Ij,k}) = \frac{((1+t^{m,t}_{ij,k})c^t_{ij,k})^{-\theta_k}}{\sum_{i'}((1+t^{m,t}_{i'j,k})c^t_{i'j,k})^{-\theta_k}}, \tag{7}$$

$$E^t_k(L^t_{i,k}) = (L^t_{i,k})^{\gamma_k}. \tag{8}$$

These choices have the advantage of focusing on the two main within-sector elasticities that matter for optimal policy. Equation (7) states that bilateral trade shares between an origin country i and a destination j in any sector k satisfy a gravity equation with trade elasticity, θ_k. Costinot et al. (2012) provide a multi-sector extension of Eaton and Kortum (2002) that provide micro-theoretical foundations for such functional form. The same micro-theoretical foundations can be invoked in the presence of external economies of scale, as in Kucheryavyy et al. (2017). Equation (8) allows external economies of scale to vary across sectors, but restricts the elasticity of external economies γ_k to be constant within each sector.

In addition, for the construction of our instruments, we will need to estimate demand residuals across countries and sectors. We will do so under the assumption that the elasticity of substitution across manufacturing sectors is constant as well. Hence, we can express country j's share of expenditure on a manufacturing sector $k \in M$, across all origins, as

$$x^t_{j,k} = \frac{\exp(\varepsilon^t_{j,k})(P^t_{j,k})^{1-\rho}}{\sum_{l \in M}(P^t_{j,l})^{1-\rho}}, \tag{9}$$

where ρ is the elasticity of substitution between sectors, $\varepsilon^t_{j,k}$ is an exogenous preference parameter, and $P^t_{j,k}$ is sector k's price index in country j,

$$P^t_{j,k} \equiv \left[\sum_i ((1+t^{m,t}_{ij,k})c^t_{ij,k})^{-\theta_k}\right]^{-1/\theta_k}. \tag{10}$$

One feature to note about the about the functional forms in equations (7)-(10) is that we allow all level-shifters to change over time, but the elasticities ($\{\theta_k\}$, $\{\gamma_k\}$ and ρ) do not. This means that, while we use data from multiple time periods, this is not necessary for identification; that is, we could proceed with data from just one time period, but choose to take advantage of the increased statistical precision that comes from pooling the data from all available cross-sections.

4.2 Empirical Strategy

We now discuss our empirical strategy for obtaining estimates of the external economies of scale elasticity γ_k and how, in the process, we will obtain estimates of the trade elasticity θ_k and and the cross sectoral elasticity of substitution ρ.

4.2.1 Baseline Specification

Let $x^t_{ij,k}$ denote the trade share of exporter i for importer j in sector k in period t. Given equations (7) and (8), equation (6) simplifies into

$$\frac{1}{\theta_{k_2}} \ln(\frac{x^t_{i_1j,k_2}}{x^t_{i_2j,k_2}}) - \frac{1}{\theta_{k_1}} \ln(\frac{x^t_{i_1j,k_1}}{x^t_{i_2j,k_1}}) = \gamma_{k_1} \ln(\frac{L^t_{i_2,k_1}}{L^t_{i_1,k_1}}) - \gamma_{k_2} \ln(\frac{L^t_{i_2,k_2}}{L^t_{i_1,k_2}}) + \ln \frac{\tilde{\eta}^t_{i_1j,k_1}}{\tilde{\eta}^t_{i_2j,k_1}} - \ln \frac{\tilde{\eta}^t_{i_1j,k_2}}{\tilde{\eta}^t_{i_2j,k_2}}.$$

The equivalent fixed-effect specification is

$$\frac{1}{\theta_k} \ln(x^t_{ij,k}) = \delta^t_{ij} + \nu^t_{j,k} + \gamma_k \ln L^t_{i,k} + \epsilon^t_{ij,k}, \tag{11}$$

where δ^t_{ij} and $\nu^t_{j,k}$ represent exporter-importer-year and importer-sector-year fixed effects, respectively, and $\epsilon^t_{ij,k} \equiv -\ln \tilde{\eta}^t_{ij,k}$.[5]

4.2.2 Construction of the Dependent Variable

To construct the dependent variable in equation (11), we need estimates of the parameter θ_k, separately for each sector k. Equation (7) implies the following gravity relationship between trade shares and trade costs

$$\ln x^t_{ij,k} = \delta^t_{j,k} - \theta_k \ln(1 + t^{m,t}_{ij,k}) - \theta_k c^t_{ij,k}, \tag{12}$$

where $\delta^t_{j,k}$ is an importer-sector-year fixed effect. A large literature has sought to estimate the trade elasticity θ_k in this equation through the use of exogenous variation in either

[5] Fixing some exporter i_2 and sector k_2, these fixed effects satisfy the following structural relationships:

$$\delta^t_{ij} \equiv \frac{1}{\theta_{k_2}} \ln(\frac{x^t_{ij,k_2}}{x^t_{i_2j,k_2}})) - \gamma_{k_2} \ln(L^t_{i,k_2}) + \ln \frac{\tilde{\eta}^t_{ij,k_2}}{\tilde{\eta}^t_{i_2j,k_2}},$$

$$\nu^t_{j,k} \equiv \frac{1}{\theta_k} \ln(x^t_{i_2j,k}) - \gamma_k \ln(L^t_{i_2,k}) + \gamma_{k_2} \ln(L^t_{i_2,k_2}) + \ln \tilde{\eta}^t_{i_2j,k}.$$

$\left(1 + t_{ij,k}^{m,t}\right)$ or $c_{ij,k}^{t}$. We draw on such estimates here by calculating the median estimate of θ_k, within each sector k, among a set of studies that we describe below.

4.2.3 Construction of the Instrument

As discussed in Section 3, ordinary least squares (OLS) estimates of Equation (11) would be biased because of the fact that an exporter's size in any sector $L_{i,k}^{t}$ would respond endogenously to the idiosyncratic productivity shocks that are part of $\epsilon_{ij,k}^{t}$. Put differently, estimation of the supply-side parameter γ_k requires demand-side instrumental variables for sector size $L_{i,k}^{t}$. We now describe a procedure for constructing such variables. We proceed in two steps.

Step 1: Estimation of the upper-tier demand elasticity. Let $X_{j,k}^{t} \equiv \sum_k X_{ij,k}^{t}$ denote the expenditure by importer j on all goods (from all origins i) in manufacturing sector k at time t and let $x_{j,k}^{t} \equiv X_{j,k}^{t} / \sum_{s \in M} X_{j,s}^{t}$ be the share of expenditures in sector k as a share of total manufacturing expenditures. The CES preferences of equation (9) imply that such expenditures will depend on prices as follows

$$\ln x_{j,k}^{t} = (1 - \rho) \ln P_{j,k}^{t} + \delta_j^t + \delta_k^t + \varepsilon_{j,k}^{t}, \tag{13}$$

where δ_j^t is a country-year fixed effect (that controls for the upper-tier manufacturing price index Estimates of the price indices $P_{j,k}^{t}$ can be obtained from the estimated importer-sector-year fixed effect $\delta_{j,k}^{t}$ in a relaxed version of our main estimating equation 11,

$$\frac{1}{\theta_k}(\ln x_{ij,k}^{t} - \ln x_{j,k}^{t}) = \delta_{i,k}^{t} + \delta_{ij}^{t} + \delta_{j,k}^{t} + \xi_{ij,k.}^{t} \tag{14}$$

With such estimates in hand, which we denote $\widehat{P}_{j,k}^{t} = \hat{\delta}_{j,k}^{t}$, we estimate ρ in the demand equation (13), for which an instrumental variables (IV) procedure is necessary to circumvent simultaneity bias.[6]

Because any valid IVs for this demand estimation problem would come from supply-side variation, we draw on the supply-based logic of our economies of scale model. In the presence of $\gamma_k > 0$, we know that a country j's productivity in any sector k will be increasing in $L_{j,k}^{t}$. While sector size $L_{j,k}^{t}$ is endogenously determined, a natural predictor of such sector scale (especially for the empirically relevant case of low import penetration in most sectors) is the country's overall population $\overline{L}_j^t$. This overall country size will have

[6] Although our baseline model does not account for intermediates, for this equation we use only expenditure shares in final goods, as the goal here is to estimate a preference parameter.

a differential impact on productivity, and hence price reduction, across sectors depending on their relative strength of economies of scale γ_k. We would therefore suspect that relevant IVs could be constructed from an interaction between $\overline{L}_j^t$ and γ_k. However, γ_k is unknown at this stage—indeed, our procedure for estimating γ_k below relies on knowledge of the parameter ρ that is the goal here. So we simply construct instruments from the interaction of $\overline{L}_j^t$ and a set of sector indicators, and then later confirm that there is a strong (inverse) correlation between the first stage coefficients here, on each sector interaction variable, and the sector's corresponding estimate of γ_k.

Summarizing this discussion, the first stage of our IV upper-tier demand elasticity estimation procedure is

$$\ln \widehat{P}_{j,k}^t = \sum_s \beta_s \mathbf{1}_{s=k} \cdot \ln \overline{L}_j^t + \widetilde{\delta}_j^t + \widetilde{\delta}_k^t + \widetilde{\varepsilon}_{j,k}^t \tag{15}$$

where $\mathbf{1}_{s=k}$ denotes an indicator variable for the event that $s = k$, and $\widetilde{\delta}_j^t$ and $\widetilde{\delta}_k^t$ represent country-year and sector-year fixed effects, respectively. The exclusion restriction corresponding to this IV requires that countries with large populations do not have systematically greater demand, relative to smaller countries, in some sectors than others. We find this plausible, especially in the light of the finding (described below) that $corr(\gamma_k, \beta_k) < 0$, as we would expect from predominantly supply-driven variation. Finally, we note that because of serial correlation over time we report standard errors of ρ that are clustered by country-sector.

Step 2: Combining Demand Residuals and Population. Equation (13) posits that the demand-shifter $\varepsilon_{j,k}^t$ captures variation in demand across sectors and countries that is not a function of prices $P_{j,k}^t$. We use estimates of these demand-shifters to construct an IV for $\ln L_{i,k}^t$ in equation (11), via a procedure that works as follows. First, we estimate $\varepsilon_{j,k}^t$ from the residuals of equation (13), and denote this estimate by $\widehat{\varepsilon}_{j,k}^t$. Because these residuals govern expenditure shares, rather than levels, we then formulate a prediction for expenditure by multiplying $\widehat{\varepsilon}_{j,k}^t$ by the total population $\overline{L}_j^t$, a variable that our model views as an exogenous country characteristic. The total demand facing a producing country i in sector k could then be constructed from a trade cost-adjusted sum of the predicted demands $\widehat{\varepsilon}_{j,k}^t \overline{L}_j^t$ in each destination market j. However, in practice, import penetration ratios in most cases are so low that the bulk of demand comes from home sources, so we simply use $\widehat{\varepsilon}_{i,k}^t \overline{L}_i^t$ to form the basis of a prediction for the demand faced by country i in sector k. We note that this predicted demand is not the full structural model's best prediction (since, among other considerations, that prediction itself would involve the unknown parameters γ_k), but this limitation affects only the strength and not the the validity of the IV procedure (since that validity rests on the exogeneity of $\widehat{\varepsilon}^t_{j,k}$ and $\overline{L}^t_j$, not the particular function of these variables used to construct an IV that predicts demand).

4.2.4 Discussion

Summarizing, our IV formulates a demand-side predictor for (log) sector size $\ln L^t_{i,k}$ from the (log of the) interaction between demand share residuals and population, $\ln\left(\widehat{\varepsilon}^t_{i,k}\overline{L}^t_i\right)$. However, since we aim to estimate a separate, sector-specific coefficient γ_k on the variable $\ln L^t_{i,k}$ in equation (11), the appropriate econometric procedure is to use a (just-identified) 2SLS system in which the instruments are $\ln\left(\widehat{\varepsilon}^t_{i,k}\overline{L}^t_i\right)$ interacted with a full set of sector indicators, and the endogenous variables are $\ln L^t_{i,k}$ interacted with that same full set of sector indicators.[7]

Before proceeding, it is instructive to think about the reduced-form regression associated with our IV approach. Abusing notation so that we can use the same labels for coefficients in the reduced-form and in the second-stage equation (11) above, the reduced-form equation can be written as

$$\frac{1}{\theta_k}\ln(x^t_{ij,k}) = \delta^t_{ij} + \nu^t_{j,k} + \eta_k \ln\left(\widehat{\varepsilon}^t_{i,k}\overline{L}^t_i\right) + \epsilon^t_{ij,k}, \tag{16}$$

with the reduced-form coefficients from each sector given by η_k. This is not a structural equation in our model,but we still expect the coefficients η_k to depend on upper-tier preferences, scale elasticities, and the extent to which countries are trading internationally. In particular, for the case of $\rho > 1$, and relatively closed economies, we expect the reduced-form coefficients $\hat{\eta}_k$ to line up with the structural second-stage coefficients $\hat{\gamma}_k$—that is, whenever a country i has relatively large home demand $\ln\left(\widehat{\varepsilon}^t_{i,k}\overline{L}^t_i\right)$ in sector k it should be expected to have relatively large elasticity-adjusted exports $\frac{1}{\theta_k}\ln(x^t_{ij,k})$ in that sector to any destination j. We explore this implication of our model when we present the empirical results below.

[7]One unusual feature of our 2SLS estimation system of equations is that the first-stage equation involves a more aggregate level of variation than the (bilateral) second-stage equation. However, this poses no difficulties of interpretation or inference given that we cluster the standard errors in all of the following regressions (first-stage and second-stage) at the exporter-sector level. In addition to correcting for unrestricted forms of serially correlated errors over time, this clustering procedure has the advantage of correcting for the purely mechanical within-group (that is, within-exporter-sector-year) correlation in the first-stage.

4.3 Data

Our main estimation procedure seeks to estimate the external economies of scale elasticity γ_k within each sector k. This requires data on bilateral trade flows $X^t_{ij,k}$ and sector size $L^t_{i,k}$, as well as data on population $\overline{L}^t_i$. We discuss each of these in turn.

We obtain data on bilateral trade flows $X^t_{ij,k}$ from the OECD's Inter-Country Input-Output (ICIO) tables. This source documents bilateral trade among 61 major exporters i and importers j, within each of 34 sectors k (27 of which are traded, with 15 in manufacturing) defined at a similar level to the 2-digit SIC, and for each year $t = 1995, 2000, 2005$, and 2010. The 15 manufacturing sectors k are those for which we aim to estimate γ_k.[8] .

We lack comparable international data on the number of (efficiency-adjusted) workers $L^t_{i,k}$ in each country and sector. However sector-level value added $Y^t_{i,k}$ is observable and, according to our model, satisfies $Y^t_{i,k} = w^t_i L^t_{i,k}$. This implies that $w^t_i = \frac{\sum_k Y^t_{i,k}}{\sum_k L^t_{i,k}}$, and so we can measure $L^t_{i,k}$ as $Y^t_{i,k}\frac{\sum_k L^t_{i,k}}{\sum_s Y^t_{i,s}}$. In our baseline without intermediate goods, $Y^t_{i,k}$ is measured as $\sum_j X^t_{ij,k}$.

Finally, we take our preferred measure of population $\overline{L}^t_i$ from the "POP" variable in the Penn World Tables version 9.0; in practice this variable is highly correlated with alternative measures such as the total labor force.

4.4 Parameter Estimates

4.4.1 Estimates of trade elasticities

As described above, we obtain estimates of θ_k from a body of prior work that has estimated this parameter for each manufacturing sector k. Specifically, we take the median estimate, within each sector, from the following studies: Bagwell et al. (2018), Caliendo and Parro (2015), Giri et al. (2018), and Shapiro (2016). The resulting estimates are detailed in Table 1.

4.4.2 Estimate of the upper-tier demand elasticity

We estimate the demand parameter ρ following the IV procedure outlined above. Table B1 reports the first-stage coefficients β_k from estimating equation (15), which have an overall F-statistic (clustered by country-sector) of 8.606. As expected, these first-stage coefficient estimates are negative (in all sectors but one), a phenomenon that we would expect if

[8] We omit sector 18, Recycling and Manufacturing NEC from the estimation.

Table 1: Trade Elasticity Estimates from Prior Studies

Sector	Shapiro (1)	BSY (2)	CP (3)	GYY (4)	Median (5)
Food, Beverages and Tobacco	5.3	10.7	2.6	3.6	4.4
Textiles	18.6	7.3	8.1	4.4	7.7
Wood Products	5.9	12.0	11.5	4.2	8.7
Paper Products	5.8	9.9	16.5	3.0	7.8
Coke/Petroleum Products	9.0	13.9	64.9	3.8	11.4
Chemicals	1.6	7.7	3.1	3.8	3.4
Rubber and Plastics	1.6	9.5	1.7	4.1	2.9
Mineral Products	12.9	8.6	2.4	5.1	6.8
Basic Metals	12.9	6.9	3.3	8.9	7.9
Fabricated Metals	12.9	5.8	7.0	5.1	6.4
Machinery and Equipment	10.8	9.0	1.5	3.3	6.2
Computers and Electronics	10.8	8.0	13.0	3.3	9.4
Electrical Machinery, NEC	10.8	9.4	12.9	3.3	10.1
Motor Vehicles	6.9	7.5	1.8	4.5	5.7
Other Transport Equipment	6.9	6.4	0.4	4.5	5.4

Notes: This table reports estimates of the trade elasticity θ_k from prior studies, matched as closely as possible to our sector classification. Column (1) refers to Table 4, column 2 in Shapiro (2016); column (2) to Table 2 in Bagwell et al. (2018); column (3) to Table 1, column 4 in Caliendo and Parro (2015); column (4) to Table 4 in Giri et al. (2018); and column 5 reports the median of columns (1)-(4).

country population causes lower prices through the logic of scale economies. Further, the correlation between $\widehat{\beta}_k$ and the estimates of $\widehat{\gamma}_k$ that we report below is -0.94, which is strongly consistent with the negative correlation one would again expect to obtain if sectors with stronger scale economies see larger price reductions due to scale.

Estimates of the parameter ρ itself are reported in Table 2. The OLS estimate, in column (1), implies that $\widehat{\rho} = 3.35$, whereas our preferred IV estimate in column (2) reveals $\widehat{\rho} = 1.47$. That is, demand for the sectors k within manufacturing appear in our data to be substitutes, albeit relatively weak ones—in particular, the standard error (clustered on country) of 0.47 in column (2) implies that we reject completely inelastic demand (i.e. $\rho = 0$), but not the range of complements (i.e. $\rho \leq 1$), at standard levels of statistical significance.

Finally, we note that the fact that our OLS estimate of ρ is lower than our IV estimate is precisely what one should expect in an increasing returns setting. When supply curves slope downwards, demand shocks lead to reductions in prices. This means that the OLS estimate, which confounds a truly downward-sloping demand curve with the negative correlation between demand shocks (in the error term) and prices, will be an underesti-

Table 2: Estimate of Upper-Tier Elasticity of Substitution $(1-\rho)$

	log (sectoral expenditure share)	log (sectoral expenditure share)
	OLS (1)	IV (2)
log (sectoral price index)	-2.35	-0.47
	(0.21)	(0.47)
Within R^2	0.13	0.05
Observations	3,660	3,660
First-stage F-statistic		8.606

Notes: This table reports the OLS and IV estimates of the upper-tier elasticity of substitution $(1-\sigma)$. The instruments are the natural log of country population interacted with sector dummies. All regressions include sector-time and country-time fixed effects. Appendix Table B1 reports the corresonding first-stage coefficients from the specification in column (2). Standard errors in parentheses are clustered at the country-sector level.

mate of the true demand elasticity.

4.4.3 Estimates of scale elasticities

Finally, we turn to the estimates of γ_k for each sector k. As described above, these estimates involve the logic of the nonparametric identification argument in Section 3, in its parametric form of equation (11), as well as the instrumental variable approach in Section 4.2.1. But we begin by reporting OLS estimates, which are reported in column (1) of Table 3. All of these estimates imply precisely-estimated economies of scale (i.e. $\gamma_k > 0$) but, as discussed, we expect these to be overestimates of true economies of scale.

For this reason we turn to the IV estimation procedure documented above. This amounts to estimating a 2SLS system in which there are 15 endogenous variables (the variable $\ln L_{i,k}^t$ interacted with an indicator variable for each sector) and 15 instruments (the variable $\ln\left(\widehat{\varepsilon}_{i,k}^t \overline{L}_i^t\right)$, again interacted with an indicator variable for each sector). While this means that there are 15 first-stage equation estimates to report (each with 15 coefficients), the F-statistics from each of those first-stage equations are large, as reported in columns (4) and (5), so potential concerns about finite-sample bias from weak instruments seem not to apply here.[9]

[9] Specifically, column (4) reports the conventional F-statistic from the 15 instruments in each first-stage equation, whereas column (5) reports the corresponding Sanderson-Windmeijer F-statistics, which assess the extent to which each first-stage is affected by independent variation in the instruments from that in the other 14 first-stages.

Table 3: Estimates of Scale Elasticities (γ_k)

Sector	OLS (1)	IV (2)	Reduced-form (3)	First-stage F-stat (4)	SW F-stat (5)
Food, Beverages and Tobacco	0.19	0.16	0.10	87.20	394.3
	(0.01)	(0.02)	(0.02)		
Textiles	0.14	0.12	0.06	56.70	349.9
	(0.01)	(0.01)	(0.02)		
Wood Products	0.13	0.11	0.05	15.50	210.7
	(0.01)	(0.02)	(0.01)		
Paper Products	0.14	0.11	0.05	55.60	661.9
	(0.01)	(0.02)	(0.01)		
Coke/Petroleum Products	0.09	0.07	0.03	14.20	299.1
	(0.01)	(0.01)	(0.01)		
Chemicals	0.23	0.20	0.17	31.10	335.8
	(0.01)	(0.02)	(0.02)		
Rubber and Plastics	0.29	0.25	0.22	39.13	436.0
	(0.02)	(0.03)	(0.03)		
Mineral Products	0.16	0.13	0.08	40.50	405.0
	(0.01)	(0.02)	(0.01)		
Basic Metals	0.13	0.11	0.07	14.40	254.0
	(0.01)	(0.01)	(0.01)		
Fabricated Metals	0.16	0.13	0.07	57.10	421.1
	(0.01)	(0.02)	(0.01)		
Machinery and Equipment	0.15	0.13	0.07	66.40	401.6
	(0.01)	(0.01)	(0.01)		
Computers and Electronics	0.10	0.09	0.04	18.60	290.5
	(0.01)	(0.01)	(0.01)		
Electrical Machinery, NEC	0.11	0.09	0.03	45.90	419.5
	(0.01)	(0.01)	(0.01)		
Motor Vehicles	0.17	0.15	0.15	39.80	390.2
	(0.01)	(0.01)	(0.02)		
Other Transport Equipment	0.17	0.16	0.11	24.00	381.6
	(0.01)	(0.02)	(0.02)		

Notes: Column (1) reports the OLS estimate, and column (2) the IV estimate, of equation (15). Column (3) reports the reduced form coefficients. The instruments are the log of (country population × sectoral expenditure share residual), interacted with sector dummies. Column (5) reports the Sanderson-Windmeijer F-statistic from the first-stage regression corresponding to each row. All regressions control for importer-sector-year fixed-effects and (asymmetric) trading pair-year effects. Standard errors in parentheses are clustered at the exporter-sector level. The overall Kleibergen-Paap F-statistic for the IV regression is 17.5 and the number of obserations is 207,542. The correlation between the coefficients in columns (3) and (4) is 0.96.

The resulting IV estimates of γ_k themselves are reported in column (2) of Table 3. These are our preferred estimates of the strength of economies of scale within each of the 15 manufacturing sectors in our sample. Evidently, economies of scale that are statistically significantly different from zero are present in all of these sectors. But there is substantial heterogeneity, with estimates ranging from $\gamma_k = 0.07$ in the Coke/Petroleum Products sector to $\gamma_k = 0.25$ in the Rubber and Plastics sector. We can also easily reject the hypothesis of coefficient equality at the 1% level. This heterogeneity is important for the scope for industrial policy, as we discuss in Section 5 below. Notably, as expected, in each sector we see that the OLS estimate of γ_k is larger than its corresponding IV estimate, as we would expect from OLS estimation that confounds movement along supply curves and movement of those curves themselves.

Finally, column (3) reports the reduced-form parameter estimates η_k from equation (16). Recall that an interesting feature of our empirical strategy is that our theoretical framework gives a prediction for the magnitude of the reduced form coefficients η_k relative to the structural coefficients γ_k. In particular, we expect a positive correlation—and indeed, $corr(\eta_k, \gamma_k) = 0.96$ in our estimates, which suggests an internally consistent logic in the model's interpretation of the available data.

5 Quantitative Results

5.1 Preliminaries

According to Proposition 1, optimal industrial and trade policy only require knowledge of two types of elasticities, $\frac{d \ln E_k}{d \ln L_{j,k}}$ and $\frac{d \ln c_{ji,k}}{d \ln L_{ji,k}}$. Under the parametric restrictions imposed in Section 4—equations (7) and (8)—we have $\frac{d \ln E_k}{d \ln L_{j,k}} = \gamma_k$ and $\frac{d \ln c_{ji,k}}{d \ln L_{ji,k}} = -\frac{1}{1+\theta_k}$, where the second expression uses the fact that $\frac{d \ln c_{ji,k}(L_{ji,k})}{d \ln L_{ji,k}} = -\frac{1}{1-\frac{d \ln \chi_{ji,k}}{d \ln c_{ij,k}}}$, by equation (5). Under our normalization, $s_j = t_j = 0$, this leads to

$$s_{j,k} = \gamma_k, \text{ for all } k,$$

$$t_{ji,k} = \frac{1}{1+\theta_k}, \text{ for all } k \text{ and } i \neq j,$$

$$t_{ij,k} = 0, \text{ for all } k \text{ and } i.$$

In Section 4, we have estimated γ_k for all manufacturing sectors, as well as the elasticity of substitution in consumption across those sectors, ρ. To quantify the welfare gains from trade and industrial policy, we again will use the median value of the trade elasticities in manufacturing sectors estimated in recent studies. We also need to take a stand on the value of scale and trade elasticities in non-manufacturing sectors as well on the shape of the upper-level utility function between manufacturing and non-manfacturing. In our baseline analysis, we set $\theta_k = 4$ for non-manufacturing – this is the simple average of trade elasticities across manufacturing sectors. We also set $\gamma_k = 0$ in non-manufacturing. This implies that there are welfare gains from reallocating resources from non-manufacturing to manufacturing sectors which have $\gamma_k > 0$, and so the overall gains from industrial policy will be higher than if we had set γ_k in non-manufacturing to the average γ_k across manufacturing sectors. We consider this alternative case in the sensitivity analysis in Section 5.3 and show that the gains from industrial policy are significantly lower in that case. Finally, we assume that upper-tier preferences across manufacturing and non-manufacturing are CES with the same elasticity of substitution that we estimated across manufacturing sectors, $\rho = 1.5$. Recent studies estimate values of this elasticity that are below one, and so – as we explain below – our baseline analysis will again lead to higher gains from industrial policy than if we had set the elasticity of substitution between manufacturing and non-manufacturing following this evidence. Overall, our choices are on the more aggressive side, to give a chance for industrial policy to yield high gains – as we see below, even with these choices, the gains are on the low side.

As in Dekle et al. (2007) we allow for transfers across countries so that country i has a trade deficit D_i, with $\sum_i D_i = 0$ and compute counterfactuals using exact hat algebra under the assumption that the data comes from an equilibrium without taxes or subsidies, as formally described in Appendix A.7.

5.2 Baseline Results

In column 1 of Table 4 we report the gains from optimal policy for some selected countries assuming that each of them is a small open economy. In column 2 we report the gains from imposing the export taxes that are part of the optimal policy, i.e., $1/(1+\theta_k)$, but now assuming that there are no production subsidies. One can think of these gains as those that would arise if the government rightly thought that it had the ability to manipulate the price of its exports, but wrongly thought that there were no external economies of scale. For ease of exposition, we simply refer to these as the gains from optimal trade policy.[10] In column 3 we show the simple difference between columns 1 and 2 – this gives

[10] One should keep in mind, however, that the previous vector of export taxes is only optimal when the optimal vector of production subsidies is in place. When those are constrained to be zero, the optimal trade policy would also take into account its impact on the resource allocation across sectors with different

Table 4: Gains from Optimal Policies, Selected Countries, Zero Non-Manufacturing SE

Country	Optimal Policy (1)	Classic Trade Pol. (2)	Add Industrial Policy (3)	Constrained Industrial Pol. (4)	Global Efficient Pol. (5)
United States	0.66%	0.28%	0.37%	0.38%	0.42 %
China	0.75%	0.26%	0.49%	0.51%	0.22 %
Germany	1.12%	0.48%	0.64%	0.49%	-0.36 %
Ireland	2.32%	1.20%	1.12%	1.37%	-1.81 %
Vietnam	1.78%	1.08%	0.69%	1.04%	1.41 %
Avg, Unweighted	**1.42%**	**0.81%**	**0.61%**	**0.80%**	**0.29%**
Avg, GDP Weighted	**0.88%**	**0.41%**	**0.47%**	**0.50%**	**0.22%**

Notes: Each column reports the gains, expressed as a share of initial real national income, that could be achieved by each type of policy. See the text for detailed descriptions of the exercises.

the gains that such a government would realize if it suddenly learnt the existence and magnitude of external economies of scale, and imposed the optimal Pigouvian subsidies γ_k that internalize them. We refer to these as the gains from optimal industrial policy.

The results in Table 4 reveal that the gains from optimal policy (column 1) are on average 1.42%. The gains from optimal industrial policy (column 3) are a bit smaller than those from optimal trade policy (column 2): averaging across countries, these gains are 0.61% and 0.81%, respectively. Smaller countries gain more from optimal trade and industrial policy than larger ones – this is revealed by the fact that, for each of columns 1 to 3, the simple average is higher than the corresponding GDP-weighted average. As an example, Ireland has gains from optimal policy that are almost four times higher than those of the United States (2.32% vs 0.66%). The gap is particularly large for the gains from trade policy (1.2% vs 0.28%), but the pattern also holds for the gains from industrial policy (1.12% vs 0.37%). The reason why smaller countries gain more from optimal trade policy is simple: such policy improves a country's terms of trade, and since small countries tend to trade more, they benefit more from that improvement.[11] Being more open also explains why smaller countries gain more from optimal industrial policy, but the fuller explanation is more subtle, as we discuss further below.

To provide perspective on our estimates of the gains from industrial policy, it is convenient to start from the familiar Harberger triangles. Formally, let us assume that the

external economies of scale.

[11] In our baseline analysis, all countries are assumed to face the same iso-elastic demand curves for their exports. More generally, larger countries may face more inelastic demand curves for their exports, leading to more room for manipulating their terms of trade.

economy is closed and that preferences are quasi-linear, with all income effects absorbed by the non-manufacturing sector. In the absence of general equilibrium effects, Harberger triangles capture the welfare costs of introducing a small tax or subsidy starting from the social optimum. Expressed as a fraction of total income, they can be expressed as

$$\Delta W = -\frac{1}{2} \sum_{k \in M} \left(\frac{L_k}{L} \right) \frac{s_k^2}{\epsilon_k^d + \epsilon_k^s},$$

where ϵ_k^d and ϵ_k^s are the inverse of the demand and supply elasticities in sector k, respectively. In our model, $\epsilon_k^d = 1/\rho$, whereas $\epsilon_k^s = -\gamma_k$ for all $k \in M$. If we think of the gains from industry policy as the negative of the losses from removing the optimal subsidies $s_k = \gamma_k$, we therefore obtain

$$\Delta W = \frac{1}{2} \sum_{k \in M} \left(\frac{L_k}{L} \right) \frac{\gamma_k^2}{1/\rho - \gamma_k}.$$

According to this simple formula, the gains from industrial policy depends on three statistics: sector sizes, L_k, scale elasticities, γ_k, and the demand elasticity, ρ.[12] Intuitively, for gains from industrial policy to be large, there must be either a large subsidy γ_k—which increase the height of the triangle—or a large quantity response of employment to the subsidy, due to a high initial level of employment or large demand and supply elasticities—which increases the base of the triangle. The quadratic term, γ_k^2, captures the first consideration, whereas $L_k/(1/\rho - \gamma_k)$ capture the second. On average, using our estimates of γ_k and ρ, the simple Harberger formula predicts gains from industrial policy around 0.55%, not far from the average that we estimate using the full structure of the model (0.61%).

The previous average results, however, masks substantial heterogeneity across countries. As mentioned above, Table 5 suggests that the gains from industrial policy tend to be larger in open economies such as Ireland and Vietnam than ones that are relatively closed such as the United States and China. This is confirmed by Figure 1, which shows a scatter plot of the gains from industrial policy (vertical axis) against openness measured as exports plus imports over gross output (horizontal axis).

Intuitively, inelastic domestic demand exerts a weaker restraint on labor reallocation in more open economies, and so there is more scope for industrial policy to generate gains. One way to see this formally is to consider the utility associated with domestic employ-

[12] Provided that $\gamma_k < 1/\rho$, so that there exists a unique interior equilibrium, this formula implies positive gains, $dW > 0$.

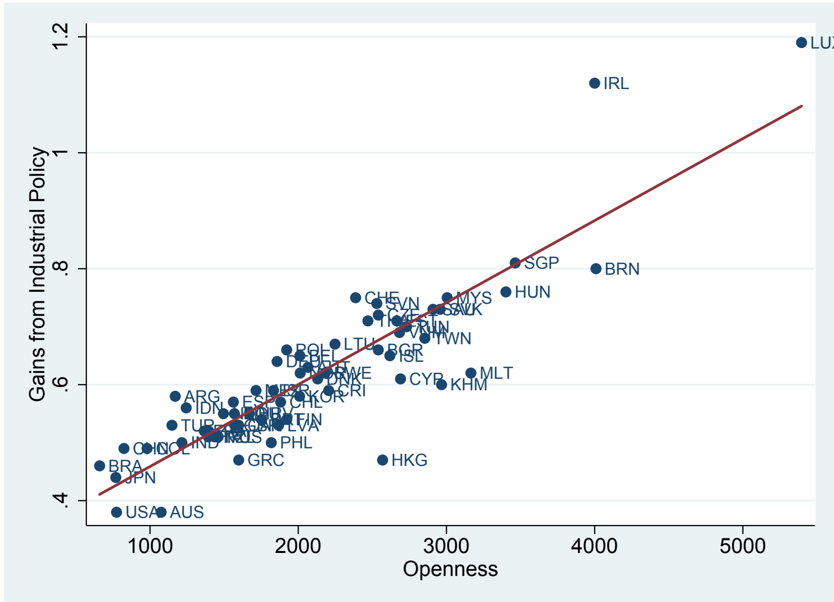

Figure 1: Gains from optimal industrial policy and specialization

Notes: The y-axes displays the gains from optimal industrial policy for an SOE, and the x-axis the openness of country, defined as $\frac{imports+exports}{GDP+Expenditure}$.

ment in an open economy. Compared to a closed economy, it no longer corresponds to the utility derived from domestic factor services, since some of those are exported, while foreign factor services are imported. Specifically, we have

$$
\begin{aligned}
V_j^{open}(\{\{L_{j,k}\}_k\}) \quad & \max_{\{\tilde{L}_{ij,k}\}_{i,k},\{\tilde{L}_{ji,k}\}_{i\neq j,k}} V_j(\{\tilde{L}_{ij,k}\}_{i,k}) \\
& \sum_{i\neq j,k} c_{ij,k}\tilde{L}_{ij,k} \leq \sum_{i\neq j,k} c_{ji,k}(\tilde{L}_{ji,k})\tilde{L}_{ji,k}, \\
& \sum_{i} \eta_{ji,k}\tilde{L}_{ji,k} \leq E_k(L_{j,k})\tilde{L}_{j,k}, \text{ for all } k.
\end{aligned}
$$

Since V_j^{open} is given by the upper-envelope, the associated indifference curves must be less convex than indifference curve associated with V_j. Hence, demand for domestic factor services must be more elastic in the open economy, a version of Le Châtelier's Principle. If so, the areas of Harberger triangles should be bigger as well, leading to greater gains from industrial policy.

In column 4 of Table 4 we present the results of a different exercise: we compute the gains from production subsidies chosen to maximize a country's welfare assuming that it cannot use trade taxes. This is motivated by the fact that such taxes are beggar thy neighbor policies, and so in principle one would expect international trade agreements to prevent countries from using them. We solve for this numerically by finding the production subsidies that maximize utility conditional on zero trade taxes. Intuitively, these constrained-optimum production subsidies involve a compromise between internalizing the production externalities via Pigouvian subsidies and improving the country's terms of trade by taxing the sectors with the lowest trade elasticities.[13] The results in column 4 tend to be lower than those in column 3.

Finally, in column 5 of Table 4 we show the gains that each country derives if all countries follow the policy that maximizes world welfare $\sum_i \pi_i U_i$ for any set of Pareto weights π_i – this entails $s_{i,k} = \gamma_k$ for all i, k. We see that there are large distributional implications associated with the imposition of globally efficient production subsidies. For example, Vietnam experiences a welfare gain of 1.41% while Ireland suffers a welfare loss of 1.81%. This is because of terms of trade changes: countries that specialize in sectors with high scale elasticities experience a deterioration of their terms of trade since those sectors expand everywhere thanks to positive production subsidies.[14]

5.3 Sensitivity Analysis

In this section we explore the sensitivity of the results regarding the gains from industrial policy to (1) allowing for scale economies in non-manufacturing, (2) considering different values of γ_k in manufacturing sectors, and (3) considering different values of ρ.

Scale Economies in Non-Manufacturing. We start by studying the implications of our assumption that there are no scale economies in non-manufacturing. In particular, we now assume that $\gamma_k = \bar{\gamma}_{j,M}$ for all $k \notin M$, with $\bar{\gamma}_{j,M} \equiv \sum_{s \in M} \frac{\pi_{j,s}}{\sum_{s' \in M} \pi_{j,s'}} \gamma_s$, where $\pi_{j,s}$ is the share of gross output in sector s in total gross output in country j. Roughly speaking, this implies that optimal policy only internalizes differences in scale elasticities within manufacturing. The results are reported in Table 5. As one would expect, the gains from industrial policy decrease significantly, going from a simple average of 0.61% to 0.15%.

[13] An illustrative case to consider is the one in which all production is exported – in this case production subsidies replicate the effect of both the production subsidies and export taxes in the unconstrained policy case, with the subsidies equal to $(1 + \gamma_k) \frac{\theta_k}{1+\theta_k} - 1$. As long as there is some sector in which part of domestic production is sold at home, however, the constrained-optimal production subsidies would deviate from these production subsidies and the corresponding gains in column 4 would be lower than those from the unconstrained policy in column 1.

[14] To confirm this intuition, we computed the correlation between the gains from industrial policy in column 5 and the country-level correlation between sectoral net exports and scale elasticities. The correlation is -0.4.

Table 5: Gains from Optimal Policies, Selected Countries, Average Non-Manufacturing SE

Country	Optimal Policy (1)	Classic Trade Pol. (2)	Add Industrial Policy (3)	Constrained Industrial Pol. (4)	Global Efficient Pol. (5)
United States	0.33%	0.27%	0.07%	0.02%	0.31 %
China	0.37%	0.26%	0.11%	0.08%	0.00 %
Germany	0.71%	0.56%	0.16%	0.03%	-0.46 %
Ireland	1.67%	1.24%	0.43%	0.63%	-1.66 %
Vietnam	1.06%	0.84%	0.22%	0.29%	0.70 %
Avg, Unweighted	**0.84%**	**0.69%**	**0.15%**	**0.18%**	**-0.23%**
Avg, GDP Weighted	**0.49%**	**0.39%**	**0.11%**	**0.07%**	**0.01%**

Notes: Each column reports the gains, expressed as a share of initial real national income, that could be achieved by each type of policy. See the text for detailed descriptions of the exercises.

Alternative Elasticity of Substitution Across Sectors. Next, we compute the gains from industrial policy for alternative values of rho, $\rho = 0.5$ and $\rho = 3$. As explained above, we expect that a higher ρ leads to higher gains from industrial policy, and this is confirmed in Table 6. We see that the average gain from industrial policy increases from 0.61% to 1.05% as ρ increases from the baseline 1.5 to 3, while they fall from 0.61% to 0.42% as ρ decreases from the baseline 1.5 to 0.5.

Alternative Scale Elasticities. Finally, we compute the gains from industrial policy if we simply set $\gamma_k = 0.95/\theta_k$, which is close to what we would have in the multi-sector model with monopolistic competition (e.g., Krugman, 1980) where we would have $\gamma_k = 1/\theta_k$. The average gains from industrial policy would increase from 0.61% in the baseline to 0.8%, as can be seen from the last row in Table 6.

Taking these results together, we conclude that the gains from industrial policy could vary from the numbers presented in Section 5.2, but the number is unlikely to be much higher than 1%, and could be as low as 0.15%.

6 Concluding Remarks

Perennial arguments for industrial policy rest on three beliefs. First, that production processes display external economies of scale—such that a nation's productivity in a given sector is increasing in its scale in that sector. Second, that such scale economies differ

Table 6: Gains from Optimal Policies, Further Sensitivity Analysis

Assumption	Optimal Policy (1)	Classic Trade Pol. (2)	Add Industrial Policy (3)	Constrained Industrial Pol. (4)	Global Efficient Pol. (5)
Upper EOS $\rho = 3$	1.78 %	0.73%	1.05%	1.10%	0.83 %
Upper EOS $\rho = 0.5$	1.27%	0.85%	0.42%	0.68%	0.09 %
High SE (γ)	1.14%	0.34%	0.80%	0.38%	0.29 %

Notes: Each column reports the unweighted average gains, expressed as a share of initial real national income, that could be achieved by each type of policy. The rows represent different configurations of parameter values. The first row sets the upper tier elasticity of substituion $\rho = 3$, while the second row sets $\rho = 0.5$. The third row takes the trade elasticities from Table 1 as given and sets $\gamma_k = 0.95/\theta_k$. The fourth row takes the estimated γ_k from Table 3 as given, and sets $\theta_k = 0.95/\gamma_k$. All scenarios assume that the SE in non-manufacturing is zero and the TE in non-manufacturing is 4.

across sectors—such that any productivity-enhancing expansion of scale in one sector does not just lead to an equal and opposite contraction of productivity in some other sector. And third, that countries produce highly substitutable and tradable goods—such that a country can simultaneously expand scale in one sector without driving down the price of its own output, and find useful foreign alternative versions of the goods in the sector that it chooses to shrink.

In this paper we have set out to estimate and quantify these three forces and in that way arrive at a better understanding of when and where industrial policy might succeed. Methodologically, our main contribution has been to show how international trade data can be used to circumvent two well-known obstacles to credible estimation of aggregate economies of scale: the difficulties of measuring aggregate productivity when products proliferate in their unobserved quality levels and their dauntingly complex patterns of substitutability; and the simultaneity bias caused when observed scale is codetermined by both supply and demand forces.

We find that external economies of scale do indeed exist, and do indeed differ substantially across sectors (ranging from an elasticity of 0.07 to 0.25), but the gains from unilateral industrial policy for all countries in our sample are never particularly large (and equal to just 0.61% of GDP on average across all countries) because countries cannot much expand in attractive sectors without both depressing the price of their goods and forcing consumers to import, often at high trade costs, imperfect substitutes for these goods. These gains are a bit smaller than the gains from optimal trade policy (0.81% of GDP on average).

References

Ackerberg, D., L.C. Benkard, S. Berry, and A. Pakes, *Econometric Tools for Analyzing Market Outcomes*, Vol. 6 of *Handbook of Econometrics*, Elsevier, 2007.

Adao, Rodrigo, Arnaud Costinot, and Dave Donaldson, "Nonparametric counterfactual predictions in neoclassical models of international trade," *The American Economic Review*, 2017, *107* (3), 633–689.

Antweiler, Werner and Daniel Trefler, "Increasing Returns and All That: A View from Trade," *American Economic Review*, 2002, *92* (1), 93–119.

Arrow, K.J. and F.H. Hahn, *General competitive analysis*, San Francisco: Holden-Day, 1971.

Bagwell, Kyle, Robert W. Staiger, and Ali Yurukoglu, "Quantitative Analysis of Multi-Party Tariff Negotiations," 2018.

Balistreri, E.J., R.H Hillberry, and T.F. Rutherford, "Structural Estimation and Solution of International Trade Model with Heterogeneous Firms," *Journal of Intermational Economics*, 2011, *83* (1), 95–108.

Basu, S. and J.G. Fernald, "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, 1997, *105*, 249–283.

Basu, Susanto, "Returns to Scale Measurement," *New Palgrave Dictionary of Economics*, 2008.

Caballero, Ricardo J. and Richar K. Lyons, "External Effects in U.S. Procyclical Productivity," *Journal of Monetary Economics*, 1992, *29* (209-225).

Caliendo, Lorenzo and Fernando Parro, "Estimates of the Trade and Welfare Effects of NAFTA," *Review of Economic Studies*, 2015, *82* (1), 1–44.

Costinot, Arnaud and Andres Rodríguez-Clare, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," in Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, eds., *Handbook of International Economics*, Vol. 4, New York: Elsevier, 2013.

_, Dave Donaldson, and Ivana Komunjer, "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," *Review of Economic Studies*, 2012, *79* (2), 581–608.

— , — , **Margaret Kyle, and Heidi Williams**, "The More We Die, The More We Sell? A Simple Test of the Home-Market Effect," *mimeo MIT*, 2016.

Davis, Donald R. and David Weinstein, "Market access, economic geography and comparative advantage: an empirical test," *Journal of International Economics*, 2003, *59*, 1–23.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum, "Unbalanced Trade," *NBER working paper 13035*, 2007.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

— **and** — , "Putting Ricardo to Work," *Journal of Economic Perspectives*, 2012, *26* (2), 65–90.

Giri, Rahul, Kei-Mu Yi, and Hakan Yilmazkuday, "Gains from Trade: Does Sectoral Heterogeneity Matter?," 2018.

Graham, Frank D., "Some Aspects of Protection Further Considered," *The Quarterly Journal of Economics*, 1923, *37* (2), 199–227.

Harrison, Ann E. and Andres Rodríguez-Clare, *Trade, Foreign Investment, and Industrial Policy for Developing Countries*, Vol. 5 of *Handbook of Development Economics*, Elsevier, 2010.

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

— **and Thierry Mayer**, "Gravity Equations: Toolkit, Cookbook, Workhorse," in Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, eds., *Handbook of International Economics*, Vol. 4, New York: Elsevier, 2013.

Hopenhayn, Hugo A., "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 1992, *60* (5), 1127–1150.

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

Kucheryavyy, Konstantin, Gary Lin, and Andres Rodríguez-Clare, "Grounded by Gravity: A Well-Behaved Trade Model with Industry-Level Economies of Scale," *mimeo UC Berkeley*, 2017.

Lashkaripour, Ahmad and Vova Lugovskyy, "Scale Economies and the Structure of Trade and Industrial Policy," *mimeo Indiana University*, 2018.

Levchenko, Andrei A. and Jing Zhang, "The Evolution of Comparative Advantage: Measurement and Welfare Implications," *Journal of Monetary Economics*, 2016, *78*, 96–111.

Newey, W.K. and J.L. Powell, "Instrumental Variable Estimation in Nonparametric Models," *Econometrica*, 2003, *71* (5), 1565–1578.

Shapiro, Joseph, "Trade Costs, CO2, and the Environment," *American Economic Journal: Economic Policy*, 2016, *8* (4), 220–54.

A Proofs

A.1 Firm-Level Economies of Scale

In Section 2.1, we have argued that our model, which assumes constant returns to scale at the good level, is consistent with firm-level economies of scale. We now make this point formally.

In any origin country i, suppose that there is a large pool of perfectly competitive firms. Like in Section 2.1, firms can use the same composite input to produce any good in any sector. Unlike in Section 2.1, firms must pay a fixed entry cost, $f_{ij,k}(\omega)$, to start producing in sector k for country j. Once this fixed cost has been paid, firms get access to a production function,

$$q = A_{ij,k} E_k^A(L_{i,k}) F(l, \phi),$$

where l is the amount of the composite input used by the firm; ϕ is a firm-specific productivity shock, randomly drawn from a distribution, $G_{ij,k}(\cdot|\omega)$; and $F(l, \phi)$ determines the extent of internal economies of scale. We assume that they are such that profits, $\pi_{ij,k}(l, \phi, \omega) = p_{ij,k}(\omega) A_{ij,k} E_k^A(L_{i,k}) F(l, \phi) - w_i l$, is single-peaked.

In a competitive equilibrium with free entry: (i) firms choose l in order to maximize profits taking input prices, $\{w_i\}$, and prices, $\{p_{ij,k}(\omega)\}$, as given,

$$\pi_{ij,k}(w_i, p_{ij,k}(\omega), \phi) = \max_l p_{ij,k}(\omega) A_{ij,k} E_k^A(L_{i,k}) F(l, \phi) - w_i l;$$

and (ii) expected profits are zero for all goods with positive output,

$$\int \pi_{ij,k}(w_i, p_{ij,k}(\omega), \phi) dG_{ij,k}(\phi|\omega) = w_i f_{ij,k}(\omega), \text{ if } l > 0 \text{ for some firm.}$$

The two previous observations imply that producer prices must satisfy

$$p_{ij,k}(\omega) = \frac{w_i}{\alpha_{ij,k}(\omega) A_{ij,k} E_k^A(L_{i,k})}, \text{ if } q_{ij,k}(\omega) > 0,$$

where $\alpha_{ij,k}(\omega)$ is a function of, and only of, $f_{ij,k}(\omega)$, $G_{ik,k}(\cdot|\omega)$, and $F(l, \phi)$. This is the dual of the production function with constant returns to scale at the good-level assumed in Section 2.1.

A.2 Competitive Equilibrium

Let w_i denote the wage in country i and let $p_{ij,k}$ denote the price of good ω sold by a firm from country i in sector k and destination j.

Profit Maximization. For any origin country i, any destination country j, any sector k, and any good ω, profit maximization determines supply,

$$q_{ij,k}(\omega) \in \operatorname{argmax}_{\tilde{q}_{ij,k}(\omega)} [p_{ij,k}(\omega)(1+s_{i,k})(1-t^x_{ij,k}) - \frac{w_i}{\alpha_{ij,k}(\omega) A_{ij,k} E^A_k(L_{i,k})}] \tilde{q}_{ij,k}(\omega), \tag{17}$$

with the convention $t^x_{ii,k} = 0$ and sector size in country i and sector k given by

$$L_{i,k} = \sum_j \int \frac{q_{ij,k}(\omega)}{\alpha_{ij,k}(\omega) A_{ij,k} E^A_k(L_{i,k})} d\omega. \tag{18}$$

Utility Maximization. For any destination country j and any sector k, utility maximization determines demand,

$$\{q_{ij,k}(\omega)\}_{i,\omega} \in \operatorname{argmax}_{\{\tilde{q}_{ij,k}(\omega)\}_{i,\omega}} \{U_{j,k}(\{\beta_{ij,k}(\omega) B_{ij,k} E^B_k(L_{i,k}) \tilde{q}_{ij,k}(\omega)\}_{i,\omega}) \tag{19}$$

$$| \sum_{i,\omega} p_{ij,k}(\omega)(1+t^m_{ij,k}) \tilde{q}_{ij,k}(\omega) = X_{j,k}\},$$

$$\{U_{j,k}\}_k \in \operatorname{argmax}_{\{\tilde{U}_{j,k}\}_k} \{U_j(\tilde{U}_{j,1},..,\tilde{U}_{j,K}) | \sum_k P_{j,k} \tilde{U}_{j,k} = w_j L_j + T_j\}, \tag{20}$$

with the convention $t^m_{ii,k} = 1$, total expenditure in country j and sector j given by

$$X_{j,k} = P_{j,k} U_{j,k}, \tag{21}$$

and the price index in country j and sector k given by

$$P_{j,k} = \min_{\{\tilde{q}_{ij,k}(\omega)\}_{i,\omega}} \{\sum_{i,\omega} p_{ij,k}(\omega) \tilde{q}_{ij,k}(\omega) | U_{j,k}(\{\beta_{ij,k}(\omega) B_{ij,k} E^B_k(L_{i,k}) \tilde{q}_{ij,k}(\omega)\}_{i,\omega}) = 1\}. \tag{22}$$

Market Clearing. For any country i, labor demand equals labor supply,

$$\sum_{j,k} \int p_{ij,k}(\omega) q_{ij,k}(\omega) d\omega = w_i L_i. \tag{23}$$

Government Budget Balance. For any country i, the government's budget is balanced,

$$T_i = \sum_{j,k,\omega} t^m_{ji,k} p_{ji,k}(\omega) q_{ji,k}(\omega) + \sum_{j,k,\omega} t^x_{ij,k} p_{ij,k}(\omega)(1+s_{i,k}) q_{ij,k}(\omega) - \sum_{j,k,\omega} s_{i,k} p_{ij,k}(\omega) q_{ij,k}(\omega). \tag{24}$$

Definition. A competitive equilibrium with production subsidies, $\{s_{j,k}\}$, import tariffs, $\{t^m_{ij,k}\}$, export taxes, $\{t^x_{ij,k}\}$, and lump-sum transfers, $\{T_j\}$, corresponds to quantities, $\{q_{ij,k}(\omega)\}$, with sector sizes, $\{L_{i,k}\}$, sector expenditures, $\{X_{i,k}\}$, good prices, $\{p_{ij,k}(\omega)\}$, sector price indices, $\{P_{i,k}\}$, and wages, $\{w_i\}$, such that equations (17)-(24) hold.

A.3 Factor Demand

In Section 2.2, we have argued that trade shares in a perfectly competitive equilibrium satisfy equation (3), with: $\chi_{ij,k}$ homogeneous of degree zero, invertible, and a function of, and only of, $U_{j,k}$, $\{\alpha_{ij,k}(\omega)\}$, and $\{\beta_{ij,k}(\omega)\}$; and $E_k(L_{i,k}) = E_k^A(L_{i,k})E_k^B(L_{i,k})$. We now establish this result formally.

By condition (17), equilibrium quantities and prices must satisfy

$$p_{ij,k}(\omega) = \frac{w_i}{(1+s_{i,k})(1-t_{ij,k}^x)\alpha_{ij,k}(\omega)A_{ij,k}E_k^A(L_{i,k})} \text{ if } q_{ij,k}(\omega) > 0. \tag{25}$$

By condition (19), since $U_{j,k}$ is homothetic and taste shocks, $\beta_{ij,k}(\omega)B_{ij,k}E_k^B(L_{i,k})$, enter utility multiplicatively, optimal quantities consumed must satisfy

$$q_{ij,k}(\omega)\beta_{ij,k}(\omega)B_{ij,k}E_k^B(L_{i,k}) = \delta_{ij,k}(\{p_{i'j,k}(\omega')(1+t_{i'j,k}^m)/(\beta_{i'j,k}(\omega')B_{i'j,k}E_k^B(L_{i',k}))\}_{i',\omega'}|\omega)X_{j,k} \tag{26}$$

where $\delta_{ij,k}(\cdot|\omega)$ only depends on $U_{j,k}$ and $\{p_{i'j,k}(\omega')(1+t_{i'j,k}^m)/(\beta_{i'j,k}B_{i'j,k}E_k^B(L_{i',k})\}_{i',\omega'}$ represents the vector of quality-adjusted prices faced by the representative consumer in destination j and sector k.

Now consider the share of expenditure, $x_{ij,k} = \sum_\omega (1+t_{ij,k}^m)p_{ij,k}(\omega)q_{ij,k}(\omega)/X_{j,k}$, in destination j on goods from sector k produced in country i. Equations (25) and (26) imply

$$x_{ij,k} = \chi_{ij,k}((1+t_{1j,k}^m)c_{1j,k}, ..., (1+t_{Ij,k}^m)c_{Ij,k}),$$

with

$$\begin{aligned}
\chi_{ij,k}(c_{1j,k}, ..., c_{Ij,k}) &= \sum_\omega \frac{c_{ij,k}}{\alpha_{ij,k}(\omega)\beta_{ij,k}(\omega)}\delta_{ij,k}(\{\frac{c_{i'j,s}}{\alpha_{i'j,k}(\omega')\beta_{i'j,k}(\omega')}\}_{i',\omega'}|\omega),\\
c_{ij,k} &= \frac{\eta_{ij,k}w_i}{(1+s_{i,k})(1-t_{ij,k}^x)E_k(L_{i,k})},\\
\eta_{ij,k} &= \frac{1}{A_{ij,k}B_{ij,k}},\\
E_k(L_{i,k}) &= E_k^A(L_{i,k})E_k^B(L_{i,k}).
\end{aligned}$$

The fact that $\chi_{j,k}$ is homogeneous of degree zero derives from the fact that the Marshallian demand for goods is homogeneous of degree zero in prices and income. The fact that $\chi_{j,k}$ is invertible derives from the fact that demand for goods within a sector satisfies the connected substitute property and standard Inada conditions hold, as in Adao et al. (2017).

A.4 Optimal Policy in a Large Economy

In the case of a large economy, the problem of country j's government generalizes to

$$\max_{\{\tilde{L}_{ij,k}\}_{i,k},\{\tilde{L}_{ji,k}\}_{i\neq j,k},\{L_{j,k}\}_k} V_j(\{\tilde{L}_{ij,k}\}_{i,k}) \tag{27a}$$

$$\sum_{i\neq j,k} c_{ij,k}(\{\tilde{L}_{ij,k}\}_{i\neq j,k},\{\tilde{L}_{ji,k}\}_{i\neq j,k})\tilde{L}_{ij,k} \leq \sum_{i\neq j,k} c_{ji,k}(\{\tilde{L}_{ij,k}\}_{i\neq j,k},\{\tilde{L}_{ji,k}\}_{i\neq j,k})\tilde{L}_{ji,k}, \tag{27b}$$

$$\sum_i \eta_{ji,k}\tilde{L}_{ji,k} \leq E_k(\tilde{L}_{j,k})\tilde{L}_{j,k}, \text{ for all } k, \tag{27c}$$

$$\sum_k \tilde{L}_{j,k} \leq L_j. \tag{27d}$$

where both import and export prices, $c_{ij,k}(\{\tilde{L}_{ij,k}\}_{i\neq j,k},\{\tilde{L}_{ji,k}\}_{i\neq j,k})$ and $c_{ji,k}(\{\tilde{L}_{ij,k}\}_{i\neq j,k},\{\tilde{L}_{ji,k}\}_{i\neq j,k})$, are now a function of the entire vector of imports and exports.

A.5 Nonparametric Identification

In Section 3, we have argued that if there exists a vector of instruments z that satisfies the exclusion restriction, $E[\epsilon|z]=0$, as well as the completeness condition, $E[g(l)|z]=0$ implies $g=0$ for any g with finite expectation, then for any k, E_k is identified, up to a normalization. We now establish this result formally.

Fix i_1, i_2, k_1, k_2, and j. Starting from equation (6), the exclusion restriction implies

$$E[\ln\frac{\chi^{-1}_{i_1j,k_1}(x_{1j,k_1},...,x_{Ij,k_1})}{\chi^{-1}_{i_2j,k_1}(x_{1j,k_1},...,x_{Ij,k_1})} - \ln\frac{\chi^{-1}_{i_1j,k_2}(x_{1j,k_2},...,x_{Ij,k_2})}{\chi^{-1}_{i_2j,s_2}(x_{1j,s_2},...,x_{Ij,k_2})}|z] = -E[\ln\frac{E_{k_1}(L_{i_2,k_1})}{E_{k_1}(L_{i_1,k_1})} - \ln\frac{E_{k_2}(L_{i_2,k_2})}{E_{k_2}(L_{i_1,k_2})}|z].$$

Now suppose that there are two solutions (E_{k_1},E_{k_2}) and $(\tilde{E}_{k_1},\tilde{E}_{k_2})$ that solve the previous equation. Then we must have

$$E[\ln\frac{E_{k_1}(L_{i_2,k_1})}{E_{k_1}(L_{i_1,k_1})} - \ln\frac{E_{k_2}(L_{i_2,k_2})}{E_{k_2}(L_{i_1,k_2})} - \ln\frac{\tilde{E}_{k_1}(L_{i_2,k_1})}{\tilde{E}_{k_1}(L_{i_1,k_1})} + \ln\frac{\tilde{E}_{k_2}(L_{i_2,k_2})}{\tilde{E}_{k_2}(L_{i_1,k_2})}|z] = 0$$

By the completeness condition, we therefore have

$$\ln\frac{E_{k_1}(L_{i_2,k_1})}{E_{k_1}(L_{i_1,k_1})} - \ln\frac{E_{k_2}(L_{i_2,k_2})}{E_{k_2}(L_{i_1,k_2})} = \ln\frac{\tilde{E}_{k_1}(L_{i_2,k_1})}{\tilde{E}_{k_1}(L_{i_1,k_1})} - \ln\frac{\tilde{E}_{k_2}(L_{i_2,k_2})}{\tilde{E}_{k_2}(L_{i_1,k_2})},$$

which can be rearranged as

$$\ln\frac{\tilde{E}_{k_1}(L_{i_1,k_1})}{E_{k_1}(L_{i_1,k_1})} = \ln\frac{\tilde{E}_{k_1}(L_{i_2,k_1})}{E_{k_1}(L_{i_2,k_1})} + \ln\frac{E_{k_2}(L_{i_2,k_2})}{E_{k_2}(L^n_{i_1,k_2})} - \ln\frac{\tilde{E}_{k_2}(L_{i_2,k_2})}{\tilde{E}_{k_2}(L_{i_1,k_2})}.$$

Since the right-hand side does not depend on L_{i_1,k_1}, the left-hand side cannot depend on L_{i_1,k_1} either. This implies that $\ln(\tilde{E}_{k_1}(L_{i_1,k_1})/E_{k_1}(L_{i_1,k_1}))$ is a constant, i.e., that E_{k_1} is identified up to a normalization. The same argument implies that $E_{k_2}$is identified up to a normalization as well.

A.6 Imperfect Competition

In the main text, we have discussed the case of an economy with an imperfectly competitive retail sector that buys goods at marginal costs and sell them at a profit. In this alternative environment, we have argued that we can still express trade shares as a function of input prices, $\chi_{j,k}(c_{1j,k}, ..., c_{Ij,k})$. We now establish this result formally.

From our analysis in Appendix A.3, we know that the price at which the retailer from sector k in destination j can buy goods is given by

$$p_{ij,k}(\omega) = \frac{w_i}{\alpha_{ij,k}(\omega) A_{ij,k} E_k^A(L_{i,k})}.$$

Let $\bar{p}_{ij,k}(\omega)$ denote the price at which the same retailer sells to consumers. From our analysis in Appendix A.3, we also know that the demand of the consumer in destination j for goods from sector k can be expressed as

$$q_{ij,k}(\omega)\beta_{ij,k}(\omega)B_{ij,k}E_k^B(L_{i,k}) = \delta_{ij,k}(\{p_{i'j,k}(\omega')/(\beta_{i'j,k}(\omega')B_{i'j,k}E_k^B(L_{i',k}))\}_{i',\omega'}|\omega)X_{j,k}.$$

Accordingly, we can express the profit maximization problem of the retailer as

$$\max_{\{\bar{p}_{ij,k}(\omega)\}} \sum_{\omega,i} \left[\bar{p}_{ij,k}(\omega) - \frac{\tau_{ij,k}}{\alpha_{ij,k}(\omega)A_{ij,k}E_k^A(L_{i,k})}\right] \left[\frac{\delta_{ij,k}(\{\bar{p}_{i'j,k}(\omega')/(\beta_{i'j,k}(\omega')B_{i',k}E_k^B(L_{i',k}))\}_{i',\omega'}|\omega)X_{j,k}}{\beta_{ij,k}(\omega)B_{ij,k}E_k^B(L_{i,k})}\right]$$

or, in terms of quality adjusted prices, $\tilde{p}_{ij,k}(\omega) \equiv \bar{p}_{ij,k}(\omega)/(\beta_{ij,k}(\omega)B_{ij,k}E_k^B(L_{i,k}))$,

$$\max_{\{\tilde{p}_{ij,k}(\omega)\}} \sum_{\omega,i} \left[\tilde{p}_{ij,k}(\omega) - \frac{\eta_{ij,k}w_i}{\alpha_{ij,k}(\omega)\beta_{ij,k}(\omega)E_k(L_{i,k})}\right] \delta_{ij,k}(\{\tilde{p}_{i'j,k}(\omega')\}_{i',\omega'}|\omega)X_{j,k}.$$

The solution to the previous problem must take the form

$$\tilde{p}_{ij,k}(\omega) = \frac{\eta_{ij,k}w_i}{\alpha_{ij,k}(\omega)\beta_{ij,k}(\omega)E_k(L_{i,k})}\mu_{ij,k}(\frac{\eta_{1j,k}w_1}{E_k(L_{1,k})}, ..., \frac{\eta_{Ij,k}w_I}{E_k(L_{I,k})}|\omega),$$

with $\mu_{ij,k}(\cdot|\omega)$ the markup on good ω as a function of the vector of cost shifters. Together with the observation that,

$$x_{ij,k} = \sum_{\omega} \tilde{p}_{ij,k}^k(\omega)\delta_{ij,k}(\{\tilde{p}_{i'j,k}(\omega')\}_{i',\omega'}|\omega),$$

this implies that

$$x_{ij,k} = \chi_{ij,k}(c_{1j,k}, ..., c_{Ij,k}),$$

with

$$\chi_{ij,k}(c_{1j,k},...,c_{Ij,k}) = \sum_{\omega}[\frac{c_{ij,k}}{\alpha_{ij,k}(\omega)\beta_{ij,k}(\omega)}\mu_{ij,k}(c_{1j,k},...,c_{Ij,k}|\omega), \\ \times \delta_{ij,k}(\{\frac{c_{i'j,k}}{\alpha_{i'j,k}(\omega')\beta_{i'j,k}(\omega')}\mu_{ij,k}(c_{1j,k},...,c_{Ij,k}|\omega')\}_{i',\omega'}|\omega),$$

as argued in the main text.

A.7 Exact Hat Algebra

General Case. Starting from equations (17)-(24), we can describe a competitive equilibrium with production subsidies, $\{s_{j,k}\}$, import tariffs, $\{t^m_{ij,k}\}$, export taxes, $\{t^x_{ij,k}\}$, and lump-sum transfers, $\{T_j\}$, as sector sizes, $\{L_{i,k}\}$, within-sector expenditure shares, $\{x_{ij,k}\}$, between-sector expenditure shares, $\{z_{j,k}\}$, sector price indices, $\{P_{j,k}\}$, and wages, $\{w_j\}$, such that

$$\frac{w_i L_{i,k}}{1+s_{i,k}} = \sum_j \left(1 - t^x_{ij,k}\right) \frac{x_{ij,k}}{1+t^m_{ij,k}} z_{j,k}\left(w_j L_j + T_j + D_j\right),$$

$$T_j = \sum_k \left[\sum_i \frac{t^m_{ij,k} x_{ij,k}}{1+t^m_{ij,k}} z_{j,k}\left(w_j L_j + T_j + D_j\right) + \sum_i \left[t^x_{ji,k}\left(1+s_{j,k}\right) - s_{j,k}\right] \frac{x_{ji,k}}{1+t^m_{ji,k}} z_{i,k}\left(w_i L_i + T_i + D_i\right)\right],$$

$$\sum_k L_{i,k} = L_i,$$

with

$$x_{ij,k} = \chi_{ij,k}\left(\left(1+t^m_{1j,k}\right)c_{1j,k},...,\left(1+t^m_{Ij,k}\right)c_{Ij,k}\right),$$

$$P_{j,k} = \Pi_{j,k}\left(\left(1+t^m_{1j,k}\right)c_{1j,k},...,\left(1+t^m_{Ij,k}\right)c_{Ij,K}\right),$$

$$c_{ij,k} = \frac{\eta_{ij,k} w_i}{(1-t^x_{ij,k})(1+s_{i,k})E_k(L_{i,k})},$$

$$z_{j,k} = \zeta_{j,k}\left(P_{j,1},...,P_{j,K}\right),$$

where $\Pi_{j,k}(c_{1j,k},...,c_{Ij,K}) \equiv \min_{\{\tilde{q}_{ij,k}(\omega)\}_{i,\omega}}\{\sum_{i,\omega}\frac{c_{ij,k}}{\alpha_{ij,k}(\omega)}\tilde{q}_{ij,k}(\omega)|U_{j,k}(\{\beta_{ij,k}(\omega)\tilde{q}_{ij,k}(\omega)\}_{i,\omega}) = 1\}$ corresponds to the unit cost of one aggregate unit of good k in country j as a function of effective price of labor from different countries. Here D_j is the trade deficit of country j and we assume that $\sum_j D_j = 0$.

We assume that the initial equilibrium has no taxes or subsidies. We are interested in a counterfactual equilibrium with taxes and subsidies:

$$t^x_{ij,k}, t^m_{ij,k}, s_{i,k}, T_j \neq 0 \text{ for some } i, j, k.$$

For any endogenous variable with value x in the initial equilibrium and x' in the counterfactual equilibrium, we let $\hat{x} = x'/x$ denote the change in this variable. We assume that D_j are fixed in terms of the numeraire, so they do not change as we move to the counterfactual equilibrium. Changes in prices and quantities the initial and counterfactual equilibria are given by the solution to

$$\frac{\hat{w}_i \hat{L}_{i,k} w_i L_{i,k}}{1+s_{i,k}} = \sum \left(1 - t^x_{ij,k}\right) \frac{\hat{x}_{ij,k} x_{ij,k}}{1+t^m_{ij,k}} \hat{z}_{j,k} z_{j,k} w_j L_j \left(\frac{\hat{w}_j w_j L_j + T_j + D_j}{w_j L_j}\right),$$

$$\hat{x}_{ij,k} x_{ij,k} = \chi_{ij,k}\left(\left(1+t^m_{1j,k}\right)\hat{c}_{1j,k} c_{1j,k}, ..., \left(1+t^m_{Ij,k}\right)\hat{c}_{Ij,k} c_{Ij,k}\right),$$

$$\hat{c}_{ij,k} c_{ij,k} = \frac{\eta_{ij,k} \hat{w}_i w_i}{(1-t^x_{ij,k})(1+s_{i,k}) E_k(\hat{L}_{i,k} L_{i,k})},$$

$$\hat{z}_{j,k} z_{j,k} = \zeta_{j,k}\left(\hat{P}_{j,1} P_{j,1}, ..., \hat{P}_{j,K} P_{j,K}\right),$$

$$\hat{P}_{j,k} P_{j,k} = \Pi_{j,k}\left(\left(1+t^m_{1j,k}\right)\hat{c}_{1j,k} c_{1j,k}, ..., \left(1+t^m_{Ij,k}\right)\hat{c}_{1j,k} c_{Ij,K}\right),$$

$$T_j = \sum_k \left[\sum_i \frac{t^m_{ij,k} \hat{x}_{ij,k} x_{ij,k}}{1+t^m_{ij,k}} \hat{z}_{j,k} z_{j,k}\left(\hat{w}_j w_j L_j + T_j + D_j\right) + \sum_i \left[t^x_{ji,k}\left(1+s_{j,k}\right) - s_{j,k}\right] \frac{\hat{x}_{ji,k} x_{ji,k}}{1+t^m_{ji,k}} \hat{z}_{i,k} z_{i,k}\left(\hat{w}_i w_i L_i + T_i + D_i\right)\right],$$

$$\sum_k \hat{L}_{i,k} L_{i,k} = L_i.$$

Expressed in terms of observables $X_{ij,k}$,$Y_{i,k} = \sum_j X_{ij,k}$, $Y_i = \sum_k Y_{i,k}$, and L_i, we therefore have

$$\frac{\hat{w}_i \hat{L}_{i,k}}{1+s_{i,k}} Y_{i,k} = \sum \left(1 - t^x_{ij,k}\right) \frac{\hat{x}_{ij,k}}{1+t^m_{ij,k}} \hat{z}_{j,k}\left(\hat{w}_j + \frac{T_j + D_j}{Y_j}\right) X_{ij,k},$$

$$\hat{x}_{ij,k} x_{ij,k} = \chi_{ij,k}\left(\left(1+t^m_{1j,k}\right)\hat{c}_{1j,k} c_{1j,k}, ..., \left(1+t^m_{Ij,k}\right)\hat{c}_{Ij,k} c_{Ij,k}\right),$$

$$\hat{c}_{ij,k} = \frac{\hat{w}_i}{(1-t^x_{ij,k})(1+s_{i,k}) E_k(\hat{L}_{i,k} Y_{i,k} L_i / Y_i) / E_k(Y_{i,k} L_i / Y_i)},$$

$$\hat{z}_{j,k} z_{j,k} = \zeta_{j,k}\left(\hat{P}_{j,1} P_{j,1}, ..., \hat{P}_{j,K} P_{j,K}\right),$$

$$\hat{P}_{j,k} P_{j,k} = \Pi_{j,k}\left(\left(1+t^m_{1j,k}\right)\hat{c}_{1j,k} c_{1j,k}, ..., \left(1+t^m_{Ij,k}\right)\hat{c}_{1j,k} c_{Ij,K}\right)$$

$$\frac{T_j}{Y_j} = \sum_k \left[\sum_i \frac{t^m_{ij,k} \hat{x}_{ij,k} \hat{z}_{j,k}}{1+t^m_{ij,k}}\left(\hat{w}_j + \frac{T_j + D_j}{Y_j}\right) X_{ij,k} + \sum_i \left[t^x_{ji,k}\left(1+s_{j,k}\right) - s_{j,k}\right] \frac{\hat{x}_{ji,k} \hat{z}_{i,k}}{1+t^m_{ji,k}}\left(\hat{w}_i + \frac{T_i + D_i}{Y_i}\right) X_{ji,k}\right],$$

$$\sum_k \hat{L}_{i,k} Y_{i,k} = Y_i.$$

where we have used the fact that in the initial equilibrium without production subsidies $L_{ik} =$

$Y_{i,k}/w_i = (Y_{i,k}/Y_i)L_i$. Both $c_{ij,k}$ and $P_{j,k}$ can also be expressed in terms of observables,

$$c_{ij,k} = \chi_{ij,k}^{-1}(x_{1j,k}, ..., x_{Ij,k}),$$

$$P_{j,k} = \zeta_{j,k}^{-1}\left(z_{j,1}, ..., z_{j,K}\right).$$

So, proportional changes in equilibrium variables solve

$$\frac{\hat{w}_i \hat{L}_{i,k}}{1+s_{i,k}} Y_{i,k} = \sum \left(1 - t_{ij,k}^{x}\right) \frac{\hat{x}_{ij,k}}{1+t_{ij,k}^{m}} \hat{z}_{j,k} \left(\hat{w}_j Y_j + T_j + D_j\right) \frac{X_{ij,k}}{Y_j}, \tag{28}$$

$$\hat{x}_{ij,k} x_{ij,k} = \chi_{ij,k}\left(\left(1+t_{1j,k}^{m}\right)\hat{c}_{1j,k}c_{1j,k}, ..., \left(1+t_{Ij,k}^{m}\right)\hat{c}_{Ij,k}c_{Ij,k}\right), \tag{29}$$

$$\hat{c}_{ij,k} = \frac{\hat{w}_i}{(1-t_{ij,k}^{x})(1+s_{i,k})E_k(\hat{L}_{i,k}Y_{i,k}L_i/Y_i)/E_k(Y_{i,k}L_i/Y_i)}, \tag{30}$$

$$\hat{z}_{j,k} z_{j,k} = \zeta_{j,k}\left(\hat{P}_{j,1}P_{j,1}, ..., \hat{P}_{j,K}P_{j,K}\right), \tag{31}$$

$$\hat{P}_{j,k} P_{j,k} = \Pi_{j,k}\left(\left(1+t_{1j,k}^{m}\right)\hat{c}_{1j,k}c_{1j,k}, ..., \left(1+t_{Ij,k}^{m}\right)\hat{c}_{1j,k}c_{Ij,K}\right), \tag{32}$$

$$T_j = \sum_k \left[\sum_i \frac{t_{ij,k}^{m}\hat{x}_{ij,k}\hat{z}_{j,k}}{1+t_{ij,k}^{m}}\left(\hat{w}_j Y_j + T_j + D_j\right)\frac{X_{ij,k}}{Y_j} + \sum_i \left[t_{ji,k}^{x}\left(1+s_{j,k}\right) - s_{j,k}\right]\frac{\hat{x}_{ji,k}\hat{z}_{i,k}}{1+t_{ji,k}^{m}}\left(\hat{w}_i Y_i + T_i + D_i\right)\frac{X_{ji,k}}{Y_i}\right], \tag{33}$$

$$\sum_k \hat{L}_{i,k} Y_{i,k} = Y_i, \tag{34}$$

with

$$c_{ij,k} = \chi_{ij,k}^{-1}(x_{1j,k}, ..., x_{Ij,k}) \tag{35}$$

$$P_{j,k} = \zeta_{j,k}^{-1}\left(z_{j,1}, ..., z_{j,K}\right) \tag{36}$$

CES Case. Consider the special case where $\chi_{ij,k}$, $\zeta_{j,k}$, and E_k are iso-elastic,

$$\chi_{ij,k}(c_{1j,k}, ..., c_{Ij,k}) = \frac{\left(c_{ij,k}\right)^{-\theta_k}}{\sum_{i'}\left(c_{i'j,k}\right)^{-\theta_k}},$$

$$\zeta_{j,k}(P_{j,1}, ..., P_{j,K}) = \frac{\left(P_{j,k}\right)^{1-\rho}}{\sum_{k'}\left(P_{j,k'}\right)^{1-\rho}},$$

$$E_k(L_{i,k}) = L_{i,k}^{\gamma_k}.$$

In this case we have

$$\hat{x}_{ij,k}x_{ij,k} = \frac{\left(\left(1+t^m_{ij,k}\right)\hat{c}_{ij,k}c_{ij,k}\right)^{-\theta_k}}{\sum_{i'}\left(\left(1+t^m_{i'j,k}\right)\hat{c}_{i'j,k}c_{i'j,k}\right)^{-\theta_k}} = \frac{\left(\left(1+t^m_{ij,k}\right)\hat{c}_{ij,k}\right)^{-\theta_k} c_{ij,k}^{-\theta_k}/\sum_l c_{lj,k}^{-\theta_k}}{\sum_{i'}\left(\left(1+t^m_{i'j,k}\right)\hat{c}_{i'j,k}\right)^{-\theta_k} c_{i'j,k}^{-\theta_k}/\sum_l c_{lj,k}^{-\theta_k}},$$

which implies

$$\hat{x}_{ij,k} = \frac{\left(\left(1+t^m_{ij,k}\right)\hat{c}_{ij,k}\right)^{-\theta_k}}{\sum_{i'}\left(\left(1+t^m_{i'j,k}\right)\hat{c}_{i'j,k}\right)^{-\theta_k} x_{i'j,k}}.$$

We also have

$$\hat{z}_{j,k}z_{j,k} = \frac{\left(\hat{P}_{j,k}P_{j,k}\right)^{1-\rho}}{\sum_{k'}\left(\hat{P}_{j,k'}P_{j,k'}\right)^{1-\rho}} = \frac{\left(\hat{P}_{j,k}\right)^{1-\rho} z_{j,k}}{\sum_{k'}\left(\hat{P}_{j,k'}\right)^{1-\rho} z_{j,k'}},$$

which implies

$$\hat{z}_{j,k} = \frac{\left(\hat{P}_{j,k}\right)^{1-\rho}}{\sum_{k'}\left(\hat{P}_{j,k'}\right)^{1-\rho} z_{j,k'}},$$

with

$$\hat{P}_{j,k}P_{j,k} = \left(\sum_i \left(\left(1+t^m_{ij,k}\right)\hat{c}_{ij,k}c_{ij,k}\right)^{-\theta_k}\right)^{-1/\theta_k} = \left(\sum_i \left(\left(1+t^m_{ij,k}\right)\hat{c}_{ij,k}\frac{c_{ij,k}}{P_{j,k}}\right)^{-\theta_k}\right)^{-1/\theta_k} P_{j,k},$$

and, in turn,

$$\hat{P}_{j,k} = \left(\sum_i \left(\left(1+t^m_{ij,k}\right)\hat{c}_{ij,k}\right)^{-\theta_k} x_{ij,k}\right)^{-1/\theta_k}.$$

Finally, we have

$$E_k(\hat{L}_{i,k}L_{i,k})/E_k(L_{i,k}) = \hat{L}_{i,k}^{\gamma_k}$$

Thus, the equilibrium system of equations (28)-(36) simplifies into

$$\frac{\hat{w}_i\hat{L}_{i,k}}{1+s_{i,k}}Y_{i,k} = \sum_j \left(1-t^x_{ij,k}\right)\frac{\hat{x}_{ij,k}}{1+t^m_{ij,k}}\hat{z}_{j,k}\left(\hat{w}_jY_j+T_j+D_j\right)\frac{X_{ij,k}}{Y_j}, \tag{37}$$

$$T_j = \sum_k\left[\sum_i \frac{t^m_{ij,k}\hat{x}_{ij,k}\hat{z}_{j,k}}{1+t^m_{ij,k}}\left(\hat{w}_jY_j+T_j+D_j\right)\frac{X_{ij,k}}{Y_j} + \sum_i\left[t^x_{ji,k}\left(1+s_{j,k}\right)-s_{j,k}\right]\frac{\hat{x}_{ji,k}\hat{z}_{i,k}}{1+t^m_{ji,k}}\left(\hat{w}_iY_i+T_i+D_i\right)\frac{X_{ji,k}}{Y_i}\right], \tag{38}$$

$$\sum_k \hat{L}_{i,k}Y_{i,k} = Y_i, \tag{39}$$

with

$$\hat{x}_{ij,k} = \frac{\left(\left(1+t^m_{ij,k}\right)\hat{c}_{ij,k}\right)^{-\theta_k}}{\sum_{i'}\left(\left(1+t^m_{i'j,k}\right)\hat{c}_{i'j,k}\right)^{-\theta_k} x_{i'j,k}}, \tag{40}$$

$$\hat{c}_{ij,k} = \frac{\hat{w}_i}{(1-t^x_{ij,k})(1+s_{i,k})\hat{L}^{\gamma_k}_{i,k}}, \tag{41}$$

$$\hat{z}_{j,k} = \frac{\left(\hat{P}_{j,k}\right)^{1-\rho}}{\sum_{k'}\left(\hat{P}_{j,k'}\right)^{1-\rho} z_{j,k'}}, \tag{42}$$

$$\hat{P}_{j,k} = \left(\sum_i \left(\left(1+t^m_{ij,k}\right)\hat{c}_{ij,k}\right)^{-\theta_k} x_{ij,k}\right)^{-1/\theta_k}. \tag{43}$$

Once changes in previous variables have been computed using equations (37)-(43), the welfare effect is given by

$$w_j$$

$$\hat{U}_j = \frac{\hat{w}_j w_j + T_j/L_j + D_j/L_j}{\hat{P}_j P_j}\frac{P_j}{w_j + D_j/L_j} = \frac{\hat{w}_j Y_j + T_j + D_j}{\hat{P}_j}\frac{1}{Y_j + D_j},$$

where

$$\hat{P}_j = \left(\sum_k \hat{P}^{1-\rho}_{j,k} z_{j,k}\right)^{1/(1-\rho)}.$$

For the case in which only country 1 imposes trade taxes and production subsidies, and country 1 is a small economy, the system is the one above but with $\hat{w}_i = \hat{P}_{i,k} = \hat{L}_{i,k} = 1$ for all k and all $i \neq 1$.

B Additional Empirical Results

Table B1: First Stage of Elasticity of Substitution Estimation

Sector	Coef.	Sector	Coef.
Food, Beverages and Tobacco	-0.05	Basic Metals	-0.00
	(0.01)		(0.01)
Textiles	-0.03	Fabricated Metals	-0.01
	(0.01)		(0.01)
Wood Products	-0.01	Machinery and Equipment	-0.01
	(0.01)		(0.01)
Paper Products	-0.01	Computers and Electronics	-0.00
	(0.01)		(0.01)
Coke/Petroleum Products	0.00	Electrical Machinery, NEC	-0.00
	(0.00)		(0.01)
Chemicals	-0.02	Motor Vehicles	-0.04
	(0.01)		(0.01)
Rubber and Plastics	-0.08	Other Transport Equipment	-0.02
	(0.01)		(0.01)
Mineral Products	-0.02		
	(0.01)		
Within R^2	0.16		
Observations	3,660		

Notes: This table reports the first stage coefficients corrresponding to the IV estimate of the upper tier elasticity of substitution $(1-\rho)$. It is an OLS regression of the log prices on log population interacted with sector dummies, with sector-time and country-time fixed effects. Coke and Petroleum produces are the omitted category. Standard errors clustered at the country-sector level.