The Textbook Case for Industrial Policy: Theory Meets Data*

Dominick Bartelme
University of Michigan

Arnaud Costinot
MIT

Dave Donaldson
MIT

Andres Rodriguez-Clare
UC Berkeley

November 2021

Abstract

The textbook case for industrial policy is well understood. If some sectors are subject to external economies of scale, whereas others are not, a government should subsidize the first group of sectors at the expense of the second. The empirical relevance of this argument, however, remains unclear. In this paper we develop a strategy to estimate sector-level economies of scale and evaluate the gains from such policy interventions in an open economy. Our results point towards significant and heterogeneous economies of scale across manufacturing sectors, but gains from industrial policy that are hardly transformative, even among the most open economies.

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

We are grateful to Susanto Basu, Jonathan Eaton, Giammario Impullitti, Yoichi Sugita, and David Weinstein for discussions, to Tishara Garg, Daniel Haanwinckel and Walker Ray for outstanding research assistance, and to many seminar and conference participants for comments. Costinot and Donaldson thank the NSF (grant 1559015) for financial support.

1 Introduction

The textbook case for industrial policy is well understood. In sectors subject to external economies of scale, private marginal costs of production are higher than social ones. This creates a rationale for Pigouvian subsidies equal to the difference between the two, with the associated welfare gains equal to the area of the Harberger triangle located between the demand and social marginal cost curves, as illustrated in Figure 1.

The empirical relevance of the previous considerations is another matter. In his original discussion of optimal industrial policy, Pigou (1920) already noted: "Attempts to develop and expand [these theoretical results] are sometimes frowned upon on the ground that they cannot be applied to practice. For, it is argued, though we may be able to say that [...] economic welfare would be increased by granting bounties to industries falling into one category and by imposing taxes on those falling into another category, we are not able to say to which of our categories the various industries of real life belong."

One hundred years later, the challenges that Pigou identified remain major obstacles to the pursuit of industrial policy. The goal of our paper is to narrow this gap between theory and data. We show how to estimate economies of scale across manufacturing sectors using data that is commonly available. Having identified the sectors that should be subsidized at the expense of others, we then explore the welfare gains from optimal industrial policies as well as how trade openness, and the access to trade policy instruments, may affect the design of industrial policy and its welfare implications.

Our main findings are twofold. First, there are sizable and heterogeneous economies of scale across manufacturing sectors, which opens up the possibility of substantial intersectoral misallocation. Second, despite these significant economies of scale in manufacturing, the gains from industrial policy that we estimate are hardly transformative, reflecting the fact that little reallocation across sectors actually takes place in response to the optimal industrial policy, even in the most open economies.

Section 2 presents our theoretical framework. Our baseline analysis focuses on 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—from the work of Marshall (1920) to Graham's (1923) famous argument for trade protection and the formal treatment of external economies of scale in Chipman (1970) and Ethier (1982)—as well as the recent emergence of the Ricardian model as a workhorse model for quantitative work. Within this environment, we provide sufficient statistics for the structure of, as well as the

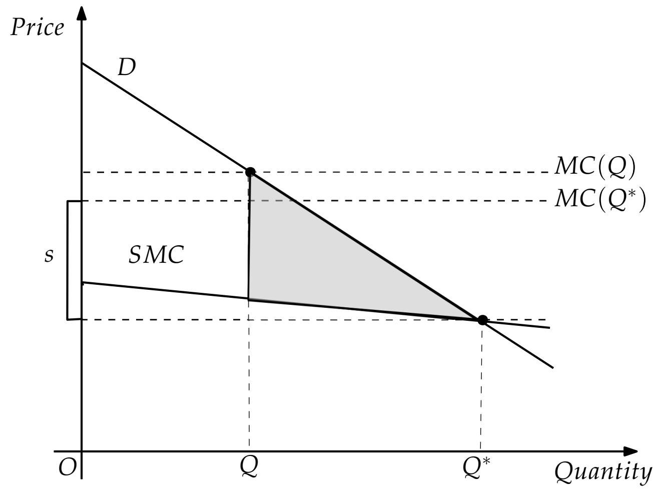


Figure 1: The Textbook Case for Industrial Policy

Notes: Due to external economies of scale in a sector, private marginal cost MC exceeds social marginal cost SMC and so output Q is less than the social optimum Q^*. The optimal industrial policy is a subsidy $s = MC(Q^*) - SMC(Q^*)$, which gives rise to gains equal to the area of the grey Harberger triangle.

welfare gains from, optimal policy.

Our first theoretical result establishes that, in the presence of optimal trade policy, the optimal industrial policy takes the form of an employment subsidy whose level only depends on the elasticity of productivity with respect to sector size, or "scale elasticity." Our second theoretical result shows that, up to a second-order approximation, the welfare gains from industrial policy are equal to half the product of the scale elasticity and the change in sector size that it generates, summed across all sectors and weighted by the share of each sector in GDP.

Section 3 describes our empirical strategy to estimate scale elasticities. It builds on two simple observations. First, if there are positive economies of scale, larger sectors should tend to sell their products at lower prices. Second, if prices are lower in those sectors, quantities demanded should tend to be higher. It follows that one can estimate scale elasticities by tracing out the impact of exogenous (i.e., demand driven) variation in sector size on equilibrium quantities.

We operationalize this general idea by assuming that within each sector: (i) productivity is a log-linear function of total sector employment, so that we have constant scale elasticities; and (ii) that the demand for output from different countries is a log-linear function of their prices, so that we have constant trade elasticities.[1] Under these restric-

[1]These parametric restrictions are satisfied by the multi-sector gravity models analyzed in Kucheryavyy et al. (2017), a set that includes models with perfect competition and external economies of scale, as in this paper, but also models with monopolistic competition and free entry, in which case scale effects arise from

tions, the (log of) export prices 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) export prices 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 idiosyncratic productivity differences across countries and sectors affect both sector size and bilateral exports, identification of scale elasticities requires demand-side instrumental variables (IVs) that are positively correlated with sector size yet uncorrelated with productivity shocks. To construct such instruments, we exploit variation in countries' population and preferences, by multiplying the former with estimates of structural demand residuals in each country-sector pair. The logic of our IV strategy is that, within each sector, employment should be higher in countries that are larger and/or have a stronger taste for goods from that sector. Under the assumption that population and demand residuals are uncorrelated with idiosyncratic productivity shocks, sector-by-sector, the previous procedure provides valid (and, in practice, strong) instruments for sector size at the country level. Importantly, our identification strategy deliberately draws on cross-sectional variation alone, so as to isolate the long-run notion of scale economies that animates the textbook case for industrial policy.

Section 4 presents our estimates of scale elasticities. Drawing on a cross-section of 61 of the world's largest countries in 2010, our results point to statistically significant scale elasticities in every 2-digit manufacturing sector, with an average of 0.17. There is also substantial heterogeneity, with sector-level estimates ranging from 0.08 to 0.42.[2] Three auxiliary findings lend support to the validity of these estimates. First, consistent with expected simultaneity bias in an open economy under elastic demand, in which sector size responds positively to productivity, in every sector we find that our demand-based IV estimate is lower than its corresponding OLS estimate. Second, while our baseline estimates are obtained from a single cross-section associated with data for the year 2010, we obtain very similar estimates from the other years in our dataset (1995, 2000, and 2005). And finally, consistent with the logic of our demand-based IV, these results are largely invariant to the inclusion of flexible supply-side controls.

Section 5 uses our empirical estimates to evaluate the gains from industrial policy.

product differentiation and love of variety within industries, as in Krugman (1980).

[2]Interestingly, the previous numbers are below the inverse of the trade elasticity in all sectors, implying that scale effects are weaker than those implicitly assumed in trade models with monopolistic competition à la Krugman (1980) or Melitz (2003), as discussed in Costinot and Rodríguez-Clare (2014) and Kucheryavyy, Lyn and Rodríguez-Clare (2017).

We find that gains from industrial policy in our baseline calibration range from 0.56% to 1.78% of GDP, with larger gains for more open economies. On average, gains from optimal industrial policy are equal to 0.98%. These modest gains from industrial policy do not reflect modest "wedges." According to our estimates of scale elasticities, if labor was to reallocate fully to the manufacturing sector with the largest scale elasticity, the average welfare gains predicted by the areas of Harberger triangles would be equal to 18.4%. Modest gains instead reflect the fact that only modest labor reallocations take place from "low-wedge" to "high-wedge" sectors, both because of inelastic domestic demand (due to a low estimated elasticity of substitution between sectors) and inelastic foreign demand (due to estimates of trade elasticities that are also small). Graphically, Harberger triangles have large heights, but small bases, even for the most open countries.

Section 6 extends our theoretical and empirical analysis to a general environment featuring multiple factors of production and input-output linkages across sectors. We show that the structure of optimal industrial policy is remarkably similar to that in our baseline environment, with a correlation between the optimal industrial policy estimated in the two environments equal to 0.99. In terms of the magnitude of the welfare gains, though, the introduction of input-output linkages raises the average gains from industrial policy from 0.98% to 3.11%. This derives both from a mechanical increase in the tax base, since gross output rather than value added is being subsidized, and from larger reallocations across sectors, since subsidizing a sector now also tends to lower the price of inputs in that sector.

Because of the prominence of industrial policy in accounts of development and underdevelopment, there is a large literature studying the rationale and potential consequences of industrial policy, as reviewed in Harrison and Rodríguez-Clare (2010) and Lane (2020). More recent work along these lines includes reduced-form empirical analysis on the consequences of the Napoleonic blockade (Juhasz, 2018) and South Korea's transition to a military dictatorship (Lane, 2017), as well as theoretical work on optimal industrial policy in the presence of financial frictions (Itskhoki and Moll, 2019 and Liu, 2019). There is, however, a dearth of work that has tried to combine both theory and empirics in order to estimate the benefits that textbook industrial policy could achieve in practice.

A notable exception is Lashkaripour and Lugovskyy (2018), which studies 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 is indirectly determined by the elasticity of substitution between domestic varieties, whereas the trade elasticity is determined by the elasticity of substitution between domestic and foreign varieties, so estimates of these two demand elasticities, obtained from monthly exchange rate variation in Colombia, can be used to calculate the effects of optimal policy. In contrast, our empirical strategy directly identifies scale elasticities from the responses of sector-level productivity, as revealed by exports, to changes in sector size caused by long-run variation in domestic demand.

Our paper also relates to a large literature that uses gravity models for counterfactual analysis, including Kucheryavyy et al. (2017) who develop the generalization of the Ricardian model with industry-level economies of scale used in our estimation of the gains from industrial policy. As discussed in Kucheryavyy et al. (2017), the quantitative predictions of gravity models hinge on two key elasticities: trade elasticities and scale elasticities. While the former have received significant attention in the empirical literature (see for instance 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, as emphasized by Costinot and Rodríguez-Clare (2014). One of the goals of our paper is to offer more direct and credible evidence about scale elasticities for use in quantitative multi-sector gravity models.

In line with the aforementioned quantitative trade literature, we focus on the estimation of scale effects that operate at the level of a country-sector pair. This matches the level at which industrial policy is most often enacted, as illustrated by the recent "Made in China 2025" initiative or its US counterpart, the "United States Innovation and Competition Act of 2021". A related literature in urban and regional economics estimates agglomeration effects at the sub-national level, as surveyed in Rosenthal and Strange (2004) and Combes and Gobillon (2015). Depending on whether these agglomeration effects are region- or region-and-sector-specific, this creates an additional rationale for place-based policies, as in Kline and Moretti (2014), or so called cluster policies, as in Duranton (2011).

Our empirical analysis at the country-sector level is most closely related to Caballero and Lyons (1992) and Basu and Fernald (1997) who estimate returns to scale for US manufacturing sectors at the two-or three-digit levels. A distinctive feature of our empirical strategy is that we do not rely on measures of real output, or price indices, collected by statistical agencies. Instead, we use existing estimates of trade elasticities to infer such prices. This provides a theoretically-grounded way to adjust for potential quality differences across origins within the same sector, as discussed in detail in Bartelme et al. (2019), as well as an approach that works symmetrically for a large set of countries around the world, albeit one that can only be applied to estimate scale elasticities in tradable sectors. Although tradable manufacturing sectors have been the traditional focus of industrial policy—from Japan, South Korea and Taiwan in the 20th century to China in the 21st—this is a non-trivial limitation of our approach (which we will have to deal with by considering a wide range of scale elasticities outside manufacturing in our counterfactual analysis).

Finally, 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. Head and Ries (2001), Davis and Weinstein (2003), and Costinot et al. (2019). Indeed, the 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, Donaldson and Komunjer (2012) and Levchenko and Zhang (2016). The starting point of these papers, like ours, is that trade flows contain information about relative costs of production, a point also emphasized by Antweiler and Trefler (2002).

2 Theory

2.1 Baseline Environment

Consider an economy with many countries, indexed by i or j, and many sectors, indexed by k. Labor is the only factor of production, with L_i the fixed labor supply in each country.

Technology. Production is subject to local external economies of scale that depend on total employment in each sector. Holding total employment fixed, firm-level production functions exhibit constant returns to scale.[3] If firms from an origin country i use $\ell_{ij,k}$ units of labor to produce good k for a destination country j, they can deliver $y_{ij,k}$ units of good k to consumers in that country with

$$y_{ij,k} = A_{ij,k} E_k(L_{i,k}) \ell_{ij,k}.$$

[3]In our working paper (Bartelme et al., 2019), we have allowed each sector to comprise multiple goods, each potentially produced by heterogeneous firms. In this more general environment, the critical assumption is constant returns to scale at the level of each good, but not at the level of firms. As is well understood, constant returns to scale at the good level may reflect the free entry of heterogeneous firms, each subject to decreasing returns to scale, as in Hopenhayn (1992).

Transportation costs, if any, are reflected in $A_{ij,k}$, while external economies of scale are reflected in $E_k(L_{i,k})$, with $L_{i,k}$ denoting total employment in sector k and country i,

$$L_{i,k} = \sum_j \ell_{ij,k}. \tag{1}$$

From now on, we simply refer to $L_{i,k}$ as sector size.

Preferences. In each destination country j, there is a representative agent with utility,

$$u_j(c_j),$$

where $c_j \equiv \{c_{ij,k}\}_{i,k}$ is the vector of consumption and $c_{ij,k}$ denotes consumption of good k from a particular origin i in a destination country j.[4]

Prices and Taxes. There are three types of ad-valorem taxes: import tariffs $\{t^m_{ij,k}\}$, export taxes $\{t^x_{ij,k}\}$, and employment subsidies $\{s_{j,k}\}$. These taxes create wedges between the prices and wages faced by consumers $\{p_{ij,k}, w_j\}$ and those faced by firms $\{q_{ij,k}, v_{j,k}\}$,

$$p_{ij,k} = (1 + t^m_{ij,k})\bar{p}_{ij,k}, \tag{2}$$

$$q_{ij,k} = (1 - t^x_{ij,k})\bar{p}_{ij,k}, \tag{3}$$

$$v_{j,k} = (1 - s_{j,k})w_j \tag{4}$$

where $\bar{p}_{ij,k}$ denotes the world (untaxed) price of a good k produced by country i and sold in country j. Net revenues from all taxes and subsidies imposed in any country j are rebated through a lump-sum transfer, T_j, to the representative agent in that country.

2.2 Competitive Equilibrium

In a competitive equilibrium, firms maximize profits; consumers maximize their utility; good and labor markets clear; and governments balance their budgets.

[4]For expositional purposes, we have restricted economies of scale to operate through physical productivity. As shown in our working paper (Bartelme et al., 2019) one can easily extend our analysis to allow sector size to affect quality. Formally, if we were to assume more generally that $y_{ij,k} = A_{ij,k}E^y_k(L_{i,k})\ell_{ij,k}$ and $u_j(\{E^c_k(L_{i,k})c_{ij,k}\})$, then external economies of scale would affect quality-adjusted productivity via $E_k(L_{i,k}) \equiv E^c_k(L_{i,k})E^y_k(L_{i,k})$, without any further implication for our analysis.

Profit Maximization. For any origin country i, any destination country j, and any sector k, profit maximization requires

$$(\ell_{ij,k}, y_{ij,k}) \in \text{argmax}_{(\ell,y)}\{q_{ij,k}y - v_{i,k}\ell | y = A_{ij,k}E_k(L_{i,k})\ell\}. \quad (5)$$

We let $\pi_{ij,k}(q_{ij,k}, v_{i,k}, L_{i,k})$ denote the value function associated with (5), i.e. the profit functions of firms from country i selling good k in country j.

Utility Maximization. For any destination country j, utility maximization requires

$$c_j \in \text{argmax}_c\{u_j(c)| \sum_{i,k} p_{ij,k}c_{ij,k} = w_jL_j + \sum_{i,k} \pi_{ji,k} + T_j\}. \quad (6)$$

We let $V_j(p_j, w_jL_j + \sum_{i,k} \pi_{ji,k} + T_j)$ denote the value function associated with (6), i.e. the indirect utility function in country j, with $p_j \equiv \{p_{ij,k}\}_{i,k}$ the vector of good prices faced by consumers in that country.

Market Clearing. All good and labor markets clear,

$$c_{ij,k} = y_{ij,k}, \quad (7)$$

$$\sum_{j,k} \ell_{ij,k} = L_i. \quad (8)$$

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

$$T_i = -\sum_k s_{i,k}w_iL_{i,k} + \sum_{j\neq i,k} t^m_{ji,k}\bar{p}_{ji,k}c_{ji,k} + \sum_{j\neq i} t^x_{ij,k}\bar{p}_{ij,k}y_{ij,k}. \quad (9)$$

Definition. A competitive equilibrium with employment subsidies, $\{s_{j,k}\}$, import tariffs, $\{t^m_{ij,k}\}$, and export taxes, $\{t^x_{ij,k}\}$ corresponds to an allocation, $\{c_{ij,k}, \ell_{ij,k}, y_{ij,k}\}$, with sector sizes, $\{L_{i,k}\}$, good prices, $\{p_{ij,k}, q_{ij,k}, \bar{p}_{ij,k}\}$, and wages, $\{w_i, v_{i,k}\}$, and lump-sum transfers, $\{T_j\}$, such that equations (1)-(9) hold.

2.3 Optimal Industrial and Trade Policy

The Government's Problem. Let $\tau \equiv \{s_{j,k}, t^m_{ij,k}, t^x_{ij,k}\}$ denote the full vector of employment subsidies and trade taxes around the world. Each competitive equilibrium with taxes τ induces equilibrium utility levels $\{U_j(\tau)\}$. The government's problem in any country j is to choose its policy $\tau_j \equiv \{s_{j,k}, t^m_{ij,k}, t^x_{ji,k}\}_{i\neq j,k}$ to maximize the utility of its representative agent,

$$\max_{\tau_j} U_j(\tau), \tag{10}$$

taking employment subsidies and trade taxes in other countries as given. In the rest of our analysis, an optimal industrial and trade policy $\tau^*_j \equiv \{s^*_{j,k}, t^{m*}_{ij,k}, t^{x*}_{ji,k}\}_{i\neq j,k}$ refers to an interior solution to (10).

To simplify the characterization of the optimal trade policy, we assume that country j is "small" in the sense that import prices, $\{\bar{p}_{ij,k}\}_{i\neq j,k}$, and other foreign equilibrium variables, $\{c_{im,k}, \ell_{im,k}, y_{im,k}, L_{i,k}, p_{im,k}, q_{im,k}, \bar{p}_{im,k}, w_i, v_{i,k}, T_m\}_{i\neq j,m\neq j,k}$, are treated as independent of τ_j, while export prices, $\{\bar{p}_{ji,k}\}_{i\neq j,k}$, are allowed to vary with its exports, $\bar{p}_{ji,k} = \tilde{p}_{ji,k}(y_{ji,k})$, reflecting the fact that goods may be differentiated by country of origin.[5]

The Structure of Optimal Taxes. To characterize the structure of optimal taxes in country j, it is convenient to start from the following identity,

$$U_j(\tau) = V_j(p_j(\tau), I_j(\tau)),$$

where both country j's consumer prices p_j and its income $I_j \equiv w_j L_j + \sum_{i,k} \pi_{ji,k} + T_j$ are expressed as functions of τ. For a small change in country j's employment subsidies and trade taxes $d\tau_j$ that generates changes in equilibrium prices and quantities, a standard envelope argument implies that

$$dU_j/\lambda_j = -\sum_k s_{j,k} w_j dL_{j,k} + \sum_{i\neq j,k} t^m_{ij,k}\bar{p}_{ij,k} dc_{ij,k} + \sum_{i\neq j,k} t^x_{ji,k}\bar{p}_{ji,k} dy_{ji,k} + \sum_k \epsilon^E_{j,k}(1-s_{j,k}) w_j dL_{j,k} + \sum_{i\neq j,k} \epsilon^p_{ji,k}\bar{p}_{ji,k} dy_{ji,k}. \tag{11}$$

where λ_j denotes the marginal utility of income in country j, $\epsilon^E_{j,k} \equiv d\ln E_k(L_{j,k})/d\ln L_{j,k}$ denotes the elasticity of productivity with respect to sector size, which we will simply refer to as the scale elasticity, and $\epsilon^p_{ji,k} \equiv d\ln \tilde{p}_{ji,k}(y_{ji,k})/d\ln y_{ji,k}$ denotes the elasticity of

[5]The small economy assumption is an implicit restriction on the primitives—preferences, technology, and labor endowments—described in Section 2.1. It holds if preferences are weakly separable across sectors and rest-of-the-world expenditure shares on goods from country j converge to zero within each sector, as we will assume in the quantitative analysis of Sections 5 and 6. In the context of a simple Armington model, Costinot and Rodríguez-Clare (2014) show that the optimal import tariff is very close to the one predicted by the small economy approximation, reflecting the fact that even the largest countries account for a very small fraction of expenditure in the rest of the world.

export prices with respect to export volumes.[6]

The first three terms in equation (11) represent the marginal increases in the deadweight loss from employment subsidies, import tariffs, and export taxes, respectively, whereas the final two terms represent the gains for country j of correcting the two sources of distortions: external economies of scale, as captured by $\sum_k \epsilon^E_{j,k}(1-s_{j,k})w_j dL_{j,k}$, and monopoly power in world markets, as captured by $\sum_{i\neq j,k}\epsilon^p_{ji,k}\bar{p}_{ji,k}dy_{ji,k}$. At the optimal policy mix, changes in utility associated with any tax variation should be zero, leading to the following necessary condition,

$$\sum_k s^*_{j,k}w^*_j dL_{j,k} - \sum_{i\neq j,k} t^{m*}_{ij,k}\bar{p}^*_{ij,k}dc_{ij,k} - \sum_{i\neq j,k} t^{x*}_{ji,k}\bar{p}^*_{ji,k}dy_{ji,k}$$
$$= \sum_k \epsilon^{E*}_{j,k}(1-s^*_{j,k})w^*_j dL_{j,k} + \sum_{i\neq j,k}\epsilon^{p*}_{ji,k}\bar{p}^*_{ji,k}dy_{ji,k}, \tag{12}$$

where asterisks denote variables in the equilibrium with optimal policy in country j. Condition (12) states that the marginal cost from taxation should be equal to its marginal benefit in terms of reducing distortions. It is immediate that (12) holds for (i) $s^*_{j,k}=\epsilon^{E*}_{j,k}/(1+\epsilon^{E*}_{j,k})$, (ii) $t^{m*}_{ij,k}=0$; and (iii) $t^{x*}_{ji,k}=-\epsilon^{p*}_{ji,k}$. Distortions caused by external economies of scale are best targeted by employment subsidies, whereas those caused by endogenous export prices are best targeted by export taxes. Since the prices of foreign goods are not manipulable, no import tariffs are necessary.

More generally, optimal industrial and trade policies can be characterized as follows:

Proposition 1. *In our baseline environment the optimal industrial and trade policy are such that:* (i) $s^*_{j,k}=(\bar{s}_j+\epsilon^{E*}_{j,k})/(1+\epsilon^{E*}_{j,k})$; (ii) $t^{m*}_{ij,k}=\bar{t}_j$; *and* (iii) $t^{x*}_{ji,k}=1-(1+\bar{t}_j)(1+\epsilon^{p*}_{ij,k})$ *for some* $\bar{s}_j$ *and* $\bar{t}_j$.

The formal proof can be found in Appendix A.2. The two shifters, $\bar{s}_j$ and $\bar{t}_j$, reflect two distinct sources of tax indeterminacy. First, since labor supply is perfectly inelastic, a uniform employment subsidy $\bar{s}_j$ only affects the level of wages in country j, but leaves the equilibrium allocation unchanged. Second, a uniform change 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 $\bar{s}_j$ and $\bar{t}_j$ to zero.

It is worth noting that many assumptions imposed previously can be relaxed without affecting Proposition 1. In terms of technology, our proof does not require the existence of a single factor and a linear production function. It can accommodate arbitrary

[6] Details are provided in Appendix A.1. Similar variational arguments can be found in Greenwald and Stiglitz (1986) and Costinot and Werning (2018), among many others.

convex technologies, a feature that we will use in Section 6 when introducing multiple factors of production and input-output linkages. Our proof can also accommodate external economies of scale $E_{j,k}(\cdot)$ that are both country-and-sector specific; the restriction that $E_{j,k}(\cdot) = E_k(\cdot)$ will only be used for identification in Section 3. Finally, our focus on a small economy is only relevant for the structure of optimal trade policy. For a large country, the optimal trade trade taxes would depend on the elasticities of export and import prices with respect to the entire vector of imports and exports, as described in Dixit (1985), not just exports within a given sector and destination, as described in Proposition 1. But regardless of whether country j is small or not, the optimal industrial policy, which is the main focus of our analysis, is given by $\epsilon_{j,k}^{E*}/(1+\epsilon_{j,k}^{E*})$.[7]

The Welfare Gains from Optimal Taxes. We define the welfare gains from optimal taxes, W_j^*, as the variation in country j's income under laissez-faire that would be equivalent to moving from laissez-faire, $\tau_j = 0$, to the optimal policy mix, $\tau_j = \tau_j^*$. Omitting foreign taxes for notational convenience, W_j^* is implicitly given by the solution to

$$V_j(p_j(\tau_j^*), I_j(\tau_j^*)) = V_j(p_j(0), I_j(0) + W_j^*). \tag{13}$$

Likewise, we define the gains from optimal industrial policy, W_j^I, as the gains from introducing optimal employment subsidies $s_j^* \equiv \{s_{j,k}^*\}$ conditional on having already imposed optimal trade taxes $t_j^* \equiv \{t_{ij,k}^{m*}, t_{ji,k}^{x*}\}_{i\neq j,k}$, and conversely, the welfare gains from optimal trade policy, W_j^T, as the gains from introducing the optimal trade taxes t_j^* conditional on having already imposed optimal employment subsidies s_j^*,

$$V_j(p_j(\tau_j^*), I_j(\tau_j^*)) = V_j(p_j(0, t_j^*), I_j(0, t_j^*) + W_j^I), \tag{14}$$

$$V_j(p_j(\tau_j^*), I_j(\tau_j^*)) = V_j(p_j(s_j^*, 0), I_j(s_j^*, 0) + W_j^T). \tag{15}$$

By definition, W_j^*, W_j^I, and W_j^T are all positive since the underlying changes in policy bring the economy towards full efficiency.[8]

[7]This derives from the fact that the welfare impact of marginal changes in sector sizes would still be equal to $\sum_k [\epsilon_{j,k}^{E*}(1-s_{j,k}^*) - s_{j,k}^*] w_j^* dL_{j,k}$. The critical assumption for the structure of industrial policy is that external economies of scale are country-and-sector specific. More generally, if productivity in a given country and sector were allowed to vary with the size of other sectors, then optimal industrial policy would depend on the elasticities of productivity with respect to the full vector of sector size, not just own-sector size. This is analogous to the issue that arises for trade policy as we go from a small to a large country.

[8]In contrast, starting from laissez-faire and only imposing the industrial or trade policy characterized in Proposition 1 may reduce welfare, as policies that only target a subset of distortions may aggravate the others. As an extreme example, imagine an economy that: (i) exports all of its output, so that employment subsidies and export taxes are perfect substitutes; and (ii) features $\epsilon_{ij,k}^{p*} = -\epsilon_{j,k}^{E*}/(1+\epsilon_{j,k}^{E*})$ for all i, j, and k

Building on the first-order necessary conditions used to characterize optimal policies, our next result offers a second-order approximation to these gains.

Proposition 2. *In our baseline environment, up to a second-order approximation, the welfare gains from optimal taxes satisfy*

$$\frac{W_j^*}{I_j} \simeq \frac{1}{2}\sum_k \frac{w_j L_{j,k}}{I_j} \cdot \frac{(\Delta L_{j,k})^*}{L_{j,k}} \cdot \epsilon_{j,k}^{E*} + \frac{1}{2}\sum_{i\neq j,k} \frac{\bar{p}_{ji,k} y_{ji,k}}{I_j} \cdot \frac{(\Delta y_{ji,k})^*}{y_{ji,k}} \cdot \epsilon_{ji,k}^{p*}, \tag{16}$$

where $(\Delta L_{j,k})^*$ and $(\Delta y_{ji,k})^*$ denote the employment and export changes associated with imposing τ_j^* and all other variables are evaluated under laissez-faire. Likewise, up to a second-order approximation, gains from industrial and trade policy satisfy

$$\frac{W_j^I}{I_j} \simeq \frac{1}{2}\sum_k \frac{w_j L_{j,k}}{I_j} \cdot \frac{(\Delta L_{j,k})^I}{L_{j,k}} \cdot \epsilon_{j,k}^{E*}, \tag{17}$$

$$\frac{W_j^T}{I_j} \simeq \frac{1}{2}\sum_{i\neq j,k} \frac{\bar{p}_{ji,k} y_{ji,k}}{I_j} \cdot \frac{(\Delta y_{ji,k})^T}{y_{ji,k}} \cdot \epsilon_{ji,k}^{p*}, \tag{18}$$

where $(\Delta L_{j,k})^I$ denotes the employment change associated with imposing s_j^*, conditional on having already imposed optimal trade taxes, and $(\Delta y_{ji,k})^T$ denotes the export change associated with imposing t_j^*, conditional on having already imposed optimal employment subsidies.

The formal proof can be found in Appendix A.3. Like the proof of Proposition 1, it remains valid for arbitrary convex technologies, including those featuring multiple factors and input-output linkages as in Section 6, and it only uses the restriction to a small economy in order to characterize the welfare gains from trade policy.

Proposition 2 can be viewed as the mathematical counterpart, in general equilibrium, of Figure 1. In each of the second-order approximations displayed in equations (16)-(18), the welfare gains correspond to a weighted sum of the areas of Harberger triangles, with weights given by employment shares, $w_j L_{j,k}/I_j$, for industrial policy and export shares, $\bar{p}_{ji,k} y_{ji,k}/I_j$, for trade policy. The height of each triangle is equal to the magnitude of the underlying distortion, either $\epsilon_{j,k}^{E*}$ or $\epsilon_{ji,k}^{p*}$, whereas the base of each triangle is equal to the changes in employment and exports generated by the policies targeting those distortions, $\Delta L_{j,k}/L_{j,k}$ and $\Delta y_{ji,k}/y_{ji,k}$.[9]

under laissez-faire. In this case, laissez-faire replicates the optimal allocation because the optimal employment subsidy exactly cancels out the export tax. Thus only imposing $s_{j,k}^* = \epsilon_{j,k}^{E*}/(1+\epsilon_{j,k}^{E*})$ in the absence of trade taxes or only imposing $t_{ji,k}^{x*} = -\epsilon_{ij,k}^{p*}$ in the absence of employment subsidies would necessarily lower welfare.

[9] Although we refer to our triangles as Harberger triangles, it is worth pointing one important concep-

For gains from industrial policy to be large, Proposition 2 highlights two conditions. First, production processes need to display large external economies of scale—such that a nation's productivity in a given sector is increasing in the size of that sector—and scale economies that differ in strength across sectors—such that any productivity-enhancing expansion of scale in one sector does not just lead to an equal and opposite decline in productivity elsewhere in the economy.[10] Second, countries need to produce highly substitutable and tradable goods—such that a country can simultaneously expand employment in one sector and find useful domestic or foreign substitutes for the sector that it chooses to shrink. These are the conditions that lead to large changes in employment, $\Delta L_{j,k}/L_{j,k}$. Likewise, for gains form trade policy to be large, world prices need to be highly manipulable, as reflected in high values of $\epsilon^{p*}_{ji,k}$, and trade flows need to be responsive to changes in trade policy, as reflected in high values of $\Delta y_{ji,k}/y_{ji,k}$.

3 Empirical Strategy

We aim now to quantify the benefits that a country can be expected to enjoy if it enacted the optimal industrial and trade policy described above. Doing so relies on estimates of scale and export price elasticities, $\{\epsilon^{E*}_{j,k}, \epsilon^{p*}_{ji,k}\}_{i\neq j,k}$. The latter can be directly recovered from existing estimates of so-called "gravity equations", as we discuss in detail below. In this section, we describe a procedure for obtaining estimates of $E_k(\cdot)$ and, in turn, $\{\epsilon^{E*}_{j,k}\}$.

3.1 General Idea

Our empirical strategy for estimating external economies of scale builds on two observations. First, the presence of external economies of scale should be reflected in prices. Everything else being equal, if there are positive external economies of scale, larger sectors should tend to sell their products at lower prices. Second, lower prices should themselves be reflected in higher quantities demanded for goods in that sector. Combining these two observations, one can therefore identify the presence of external economies by tracing out the impact of exogenous variation in sector size on quantities demanded.

tual distinction between Proposition 2 and the standard Harberger triangle analysis (see e.g. Hines, 1999). The standard analysis fixes the primitives of the economy and introduces small exogenous taxes in order to evaluate their welfare cost. Our exercise instead varies the primitives of the economy in order to generate small endogenous optimal taxes and evaluate their welfare gains. At a technical level, this distinction is reflected in the fact that the terms appearing in our Taylor expansion are $\epsilon^{E*}_{j,k}$ or $\epsilon^{p*}_{ji,k}$ and not $s^*_{j,k}$ or $t^{x*}_{ji,k}$.

[10] If $\epsilon^{E*}_{j,k} = \epsilon^{E*}_{j}$ for all k, then $\sum_k \frac{w_j L_{j,k}}{I_j} \cdot \frac{\Delta L_{j,k}}{L_{j,k}} \cdot \epsilon^{E*}_{j,k} = \frac{w_j}{I_j}\epsilon^{E*}_{j} \sum_k \Delta L_{j,k} = 0$.

Formally, the profit-maximization condition (5) implies the equality of producer price and marginal cost for any good k produced by country i for destination j,

$$q_{ij,k} = \frac{v_{i,k}}{A_{ij,k} E_k(L_{i,k})}. \tag{19}$$

Similarly, the utility-maximization condition (6) implies the equality of relative consumer price and the marginal rate of substitution for any good k sold by two countries in destination j. That is,

$$\frac{p_{ij,k}}{p_{1j,k}} = \frac{u_{ij,k}(c_j)}{u_{1j,k}(c_j)}, \tag{20}$$

where $u_{ij,k}(c_j)$ is the partial derivative of $u_j(\cdot)$ with respect to the quantity $c_{ij,k}$ evaluated at the consumption vector c_j and prices are expressed relative to those of goods from country 1. Combining these expressions, using the non-arbitrage conditions (2)-(4) relating producer and consumer prices, and averaging across destination countries implies

$$\frac{1}{J}\sum_j \mathrm{DD}_{i_0,k_0}\{\ln u_{ij,k}(c_j)\} = -\mathrm{DD}_{i_0,k_0}\{\ln E_k(L_{i,k})\} - \frac{1}{J}\sum_j \mathrm{DD}_{i_0,k_0}\{\ln \alpha_{ij,k}\}, \tag{21}$$

where $\alpha_{ij,k} \equiv A_{ij,k}(1-t^x_{ij,k})/[(1-s_{i,k})(1+t^m_{ij,k})]$ is the tax-adjusted productivity of firms from country i exporting good k to destination j, $\mathrm{DD}_{i_0,k_0}\{\cdot\}$ is the double difference relative to a reference country i_0 and a reference sector k_0, e.g., $\mathrm{DD}_{i_0,k_0}\{\ln u_{ij,k}(c_j)\} \equiv [\ln u_{ij,k}(c_j) - \ln u_{i_0j,k}(c_j)] - [\ln u_{ij,k_0}(c_j) - \ln u_{i_0j,k_0}(c_j)]$, and J is the number of destinations.[11]

In this cross-sectional difference-in-differences, the first difference, taken between country i and country i_0, reflects the fact that, as in equation (20), consumer choices reveal only relative prices, whereas the second difference, taken between sector k and sector k_0, derives from a desire to eliminate the endogenous wage w_i that is part of $v_{i,k}$ in equation (19). Lastly, the average over destinations j on the left-hand side is introduced to focus on the fact that, when buying from a given origin i and sector k, consumers in many locations j are facing potentially distinct prices and have distinct preferences, but they all face a price that is influenced by the origin-sector's size $L_{i,k}$, to the extent that external economies of scale are active.[12]

Equation (21) is the starting point of our empirical analysis. Given existing estimates

[11] It should be clear that the absence of internal decreasing returns to labor at the sector-level plays a key role in establishing equations (19) and (21). If such forces were active, our empirical strategy would only identify the combination of external and internal scale effects.

[12] While equation (21) focuses on averages across all destinations, it would be equally valid from a theoretical standpoint to use averages across any subsets of destinations with positive trade flows. For instance,

of consumer's preferences, $u_j(\cdot)$, this equation corresponds to a nonparametric regression of the form, $y = h(x) + \xi$. The left-hand side variable "y" can be measured using data on consumption choices $c_{ij,k}$ or, as we do below, data on expenditures, $X_{ij,k} \equiv p_{ij,k} c_{ij,k}$. The first term on the right-hand side, which depends on external economies of scale, $E_k(\cdot)$, is the unknown function "h" to be estimated; it is evaluated at observable industry sizes $L_{i,k}$, corresponding to "x". And the second term, which depends on $\frac{1}{J}\sum_j \ln \alpha_{ij,k}$, is the structural error term "ξ". While $L_{i,k}$ is endogenous and hence may be correlated with $\frac{1}{J}\sum_j \ln \alpha_{ij,k}$, suitable instrumental variables can be used to achieve nonparametric identification of $E_k(\cdot)$, as in Newey and Powell (2003).[13]

3.2 Parametric Model

In theory, one could point-identify external economies of scale, $E_k(\cdot)$, by tracing out nonparametrically the impact of exogenous changes in sector sizes $\{L_{i,k}\}$ on consumption choices $\{c_j\}$. In practice, our dataset will include only 61 countries. So, estimation inevitably needs to proceed parametrically, as we now describe.

Functional Form Assumptions. We assume that external economies of scale satisfy

$$E_k(L_{i,k}) = L_{i,k}^{\gamma_k}. \tag{22}$$

Hence, scale elasticities are constant within each sector, $\epsilon_{i,k}^{E*} = \gamma_k$, but may vary across sectors, as is critical for industrial policy considerations, as seen in Propositions 1 and 2. Similarly, we restrict preferences to be nested CES,

$$u_j = (\sum_k \beta_{j,k}^{\frac{1}{\rho+1}} C_{j,k}^{\frac{\rho}{\rho+1}})^{\frac{\rho+1}{\rho}}, \tag{23}$$

$$C_{j,k} = (\sum_i c_{ij,k}^{\frac{\theta_k}{1+\theta_k}})^{\frac{1+\theta_k}{\theta_k}}, \tag{24}$$

our model also implies

$$\frac{1}{J}\sum_{j\neq i} \mathrm{DD}_{i_0,k_0}\{\ln u_{ij,k}(c_j)\} = -\mathrm{DD}_{i_0,k_0}\{\ln E_k(L_{i,k})\} - \frac{1}{J}\sum_{j\neq i} \mathrm{DD}_{i_0,k_0}\{\ln \alpha_{ij,k}\}.$$

In practice, this "leave-out" specification, excluding domestics sales, leads to estimates of external economies of scale that are very similar to the baseline estimates reported in Table 1 below.

[13]The interested reader can find further details in our working paper, Bartelme et al. (2019).

with $1 + \theta_k$ the elasticity of substitution between goods from different origins, $1 + \rho$ the elasticity of substitution between goods from different sectors, and the taste shifters $\beta_{j,k}$ normalized such that $\sum_k \beta_{j,k} = 1$ for all j.[14] As is well known, such preferences give rise to a so-called "gravity equation" for within-sector trade flows, with θ_k the (sector-level) trade elasticity.[15]

Baseline Specification. As established in Appendix B.1, substituting for u_j and E_k using equations (22)-(24), we can rewrite equation (21) as

$$Y_{i,k} = \delta_i + \delta_k + \gamma_k \ln L_{i,k} + \varepsilon_{i,k}, \tag{25}$$

where $Y_{i,k} \equiv (\frac{1}{J}\sum_j \ln X_{ij,k})/\theta_k$ is the log expenditure $X_{ij,k}$ on goods from country i in sector k averaged across all destinations and adjusted by the trade elasticity θ_k, and δ_i and δ_k are country and sector fixed effects, respectively. The productivity shock, $\varepsilon_{i,k} \equiv \frac{1}{J}\sum_j \ln \alpha_{ij,k} - \mathbb{E}[\frac{1}{J}\sum_j \ln \alpha_{ij,k}|i] - \mathbb{E}[\frac{1}{J}\sum_j \ln \alpha_{ij,k}|k] + \mathbb{E}[\frac{1}{J}\sum_j \ln \alpha_{ij,k}]$, is demeaned so that $\mathbb{E}[\varepsilon_{i,k}|i] = 0$ for all i and $\mathbb{E}[\varepsilon_{i,k}|k] = 0$ for all k, where $\mathbb{E}[\varepsilon_{i,k}|i]$ refers to expectation holding i fixed (and hence the expectation is taken across k only), and analogously for $\mathbb{E}[\varepsilon_{i,k}|k]$.

To gain intuition about how equation (25) identifies scale elasticities, consider a hypothetical dataset with only two sectors, $k \in \{\text{Food,Textiles}\}$, three exporters, $i \in \{$Canada, Costa Rica, France$\}$, and one importer, $j \in \{$United States$\}$. Suppose that these three exporters have the same productivity in all sectors, $\varepsilon_{i,k} = 0$, but that because of demand-side considerations, they have different sector sizes: $L_{\text{Fra,Food}}/L_{\text{Can,Food}} > L_{\text{Fra,Text}}/L_{\text{Can,Text}} = 1$ and $L_{\text{Cos,Text}}/L_{\text{Can,Text}} > L_{\text{Cos,Food}}/L_{\text{Can,Food}} = 1$.[16] Differencing out the country and sector fixed effects in equation (25) then implies

$$\gamma_{\text{Food}} = \frac{(\ln X_{\text{FraUS,Food}} - \ln X_{\text{CanUS,Food}})/\theta_{\text{Food}} - (\ln X_{\text{FraUS,Text}} - \ln X_{\text{CanUS,Text}})/\theta_{\text{Text}}}{L_{\text{Fra,Food}} - \ln L_{\text{Can,Food}}},$$
$$\gamma_{\text{Text}} = \frac{(\ln X_{\text{CosUS,Text}} - \ln X_{\text{CanUS,Text}})/\theta_{\text{Text}} - (\ln X_{\text{CosUS,Food}} - \ln X_{\text{CanUS,Food}})/\theta_{\text{Food}}}{\ln L_{\text{Cos,Text}} - L_{\text{Can,Text}}}.$$

[14]The absence of bilateral taste shifters $\beta_{ij,k}$ in equation (24) is without loss of generality, as units of account can always be chosen such that it holds, with the productivity term $A_{ij,k}$ adjusted accordingly. We come back to this point briefly in Section 3.4.

[15]While the preferences in (23) and (24) correspond to an Armington model of trade, Costinot et al.'s (2012) multi-sector extension of Eaton and Kortum's (2002) Ricardian model delivers observationally equivalent "gravity equations". The same micro-foundations can be invoked in the presence of external economies of scale, as in Kucheryavyy et al. (2017), with identical implications for our analysis, as shown in our working paper, Bartelme et al. (2019).

[16]The assumption that France and Costa Rica have the same employment as Canada in the textile and food sectors, respectively, simplifies the algebra, but is not critical for identification. All that matters is that France and Costa Rica have different distributions of employment, $L_{\text{Fra,Food}}/L_{\text{Fra,Text}} \neq L_{\text{Cos,Food}}/L_{\text{Cos,Text}}$.

Intuitively, for a given difference in sector size, $\ln L_{\text{Fra,Food}} - \ln L_{\text{Can,Food}} > 0$, the scale elasticity in the food sector γ_{Food} is larger if: (i) the country with a large food sector exports more to the US than the country with a small food sector, i.e, $\ln X_{\text{FraUS,Food}} - \ln X_{\text{CanUS,Food}}$ is larger; (ii) US imports in the food sector are less responsive to changes in prices, i.e., θ_{Food} is lower; and (iii) the country with a large food sector has a higher wage, as revealed by its relative exports in the textile sector, i.e., $(\ln X_{\text{FraUS,Text}} - \ln X_{\text{CanUS,Text}})/\theta_{\text{Text}}$ is lower. The scale elasticity in the textiles sector γ_{Text} is identified in a similar manner by comparing the exports of Costa Rica and Canada to the US.

This example operationalizes the general idea, described in Section 3.1, that one can identify scale elasticities, i.e., by how much increases in employment lowers unit costs and, in turn, prices, by tracing out how changes in employment affect consumption choices and, in turn, expenditures. All that is required to go from the latter to the former is an estimate of demand that maps changes in expenditures into changes prices. Under our parametric assumptions, this boils down to knowledge of the trade elasticities θ_k.

In practice, of course, bilateral trade flows may depend on productivity shocks, $\varepsilon_{i,k} \neq 0$, and as noted at the end of Section 3.1, sector size $\ln L_{i,k}$ may respond endogenously to these shocks, $\mathbb{E}[\ln L_{i,k} \times \varepsilon_{i,k}|k] \neq 0$ for all k. Estimation of the vector of supply-side parameters $\{\gamma_k\}$ thus requires at least as many demand-side instrumental variables. We now describe a procedure for constructing such variables.

3.3 Instrumental Variables

In our model, sector size $L_{i,k}$ is an endogenous object determined by countries' factor supply, preferences, technology and taxes. In addition, the adjusted productivity residual $\varepsilon_{i,k}$ in our baseline specification (25) depends on technology and taxes. To construct instrumental variables, we propose to exploit variation in countries' factor supply and preferences, which we will assume to be orthogonal to the unobserved variation in technology and taxes.

To fix ideas, consider first the special case where upper-level preferences are Cobb-Douglas ($\rho \to 0$). In the absence of trade, the level of sectoral employment $L_{i,k}$ predicted by our model would then be $\beta_{i,k} \times L_i$, where $\beta_{i,k}$ can be directly obtained as the share of expenditure by country i on goods from sector k across all origins, $x_{i,k} \equiv \sum_j X_{ji,k} / \sum_{j,l} X_{ji,l}$, and L_i can be proxied by country i's population, $\hat{L}_i$. Given the magnitude of trade costs in practice, we would expect $x_{i,k} \times \hat{L}_i$ to remain a good predictor of $L_{i,k}$ in the presence of trade in this Cobb-Douglas special case.

Our instrumental variables generalize the previous idea to the case where ρ may be different from zero. The key difference between the Cobb-Douglas and CES case is that the demand shifters $\beta_{i,k}$ used to predict sectoral employment no longer satisfy $\beta_{i,k} = x_{i,k}$. Under CES, equations (23) and (24) instead imply

$$\beta_{i,k} = \frac{x_{i,k}/(P_{i,k})^{-\rho}}{\sum_l x_{i,l}/(P_{i,l})^{-\rho}}, \tag{26}$$

where $P_{i,k} \equiv (\sum_j p_{ji,k}^{-\theta_k})^{-1/\theta_k}$ is sector k's price index in country i. Provided that ρ is different from zero, the adjustment by $1/(P_{i,k})^{-\rho}$ is required to purge expenditure shares from prices and, in turn, the technological and tax considerations that may affect them. Section 4.2 describes in detail how we obtain estimates of sector-level price indices, $\hat{P}_{i,k}$, and the elasticity of substitution, $\hat{\rho}$, in order to recover estimates of the demand residuals $\hat{\beta}_{i,k}$ via equation (26). Given such estimates, we then construct a measure of demand-predicted sector size as $\hat{L}_{i,k} \equiv \hat{\beta}_{i,k} \times \hat{L}_i$, which will serve as the basis for constructing our IVs as the interaction of sector indicator variables and $\hat{L}_{i,k}$. This will deliver consistent estimates of the parameters $\{\gamma_k\}$ under the exclusion restriction,

$$\mathbb{E}[\ln \hat{L}_{i,k} \times \varepsilon_{i,k}|k] = 0 \text{ for all } k. \tag{27}$$

This requires that there be no systematic relationship, within any sector k, between a country's demand-predicted sector size $\ln \hat{L}_{i,k}$ and the productivity shock $\varepsilon_{i,k}$.[17]

3.4 Threats to the Exclusion Restriction

By construction, the (log) demand-predicted sector size variable consists of two parts, $\ln \hat{L}_{i,k} = \ln \hat{L}_i + \ln \hat{\beta}_{i,k}$. There are therefore two potential sources of violation for the above exclusion restriction,

$$\mathbb{E}[\ln \hat{L}_i \times \varepsilon_{i,k}|k] \neq 0, \text{ for some } k, \tag{28}$$
$$\mathbb{E}[\ln \hat{\beta}_{i,k} \times \varepsilon_{i,k}|k] \neq 0, \text{ for some } k. \tag{29}$$

Though condition (27) may, in principle, hold when both violations exactly compensate each other, it is hard to imagine such a situation being relevant in practice. We therefore discuss each of these two sources of violation separately.

The first violation (28) arises, for instance, if there are country-wide scale effects that

[17] Appendix B.1 explicitly describes the two stages of our IV strategy using dummy variable notation.

are heterogeneous across sectors, as in $A_{ij,k} \propto (\hat{L}_i)^{\Gamma_k}$. As demonstrated in Appendix B.2, this leads to an overestimate of the extent of external economies of scale γ_k in any sector k with higher than average values of country-wide scale effects, $\Gamma_k > \mathbb{E}[\Gamma_k]$. It should be clear, however, that the source of bias in this scenario is not the existence of country-wide scale effects, but their heterogeneity across sectors.

The second violation arises if countries that have a higher propensity to buy in some sectors, because of preference considerations, also tend to have a higher propensity to sell in those same sectors, because of technological or tax considerations. More specifically, suppose that $\hat{\beta}_{i,k} \propto (\bar{\alpha}_{i,k})^{\phi}$, with $\ln \bar{\alpha}_{i,k} \equiv \frac{1}{J}\sum_j \ln \alpha_{ij,k} - \mathbb{E}[\frac{1}{J}\sum_j \ln \alpha_{ij,k}|i]$. As demonstrated in Appendix B.2, if $\phi > 0$, this leads to an overestimate of the extent of economies of scale, and more so in sectors where $\mathrm{Var}(\varepsilon_{i,k}|k)$ is higher. It is worth noting that the issue here is not the standard concern about supply and demand residuals being correlated because of quality considerations. Such a correlation would affect the structural interpretation of our tax-adjusted productivity measures $\alpha_{ij,k}$, but leave the rest of our analysis unchanged.[18] The source of bias here is more subtle. It requires, for instance, countries with lower costs in a given sector k to develop tastes that are tilted towards goods from that sector (which may be important for cultural goods subject to habit formation, like particular food items in Atkin (2013), but is not something we expect to be prevalent among the 15 manufacturing sectors that we consider) or countries with stronger tastes for consumption in sector k to impose policies set to favor production in that sector (despite optimal policy in our model being only a function of the elasticities θ_k and γ_k, as discussed in Section 5).

The previous discussion implicitly focuses on biases arising from the structural parameters, L_i and $\beta_{i,k}$, being systematically correlated with the productivity shocks, $\varepsilon_{i,k}$. Even if they are not, our exclusion restriction may also fail if the difference between the true parameters, L_i and $\beta_{i,k}$, and our proxies for these parameters, $\hat{L}_i$ and $\hat{\beta}_{i,k}$, is itself correlated with $\varepsilon_{i,k}$, perhaps because of misspecification of preferences in equations (23) and (24), bias in the estimation of ρ, or some other form of measurement error. To alleviate concerns about these potential sources of biases in our estimates of scale elasticities, Section 4.4 reports the robustness of our results to the inclusion of controls for various sources of productivity differences across countries and sectors. Section 6 further introduces controls for cost differences deriving from multiple factors of production and tradable inputs.

[18]Formally, suppose that preferences take the form $u_j(\{B_{ij,k}c_{ij,k}\}_{i,k})$, with $B_{ij,k}$ a measure of the quality of good k sold by country i in country j. The standard concern is that $A_{ij,k}$ and $B_{ij,k}$ may be positively correlated. To see that this is irrelevant for our empirical analysis, note that starting from this more general model, one can always define quality-and-tax-adjusted measures of productivity, $\tilde{\alpha}_{ij,k} \equiv A_{ij,k}B_{ij,k}(1-t^x_{ij,k})/[(1-s_{i,k})(1+t^m_{ij,k})]$, without any further implication.

4 Empirical Results

4.1 Data

To estimate scale elasticities γ_k in equation (25) and construct our instrumental variables we need measures of bilateral trade flows $X_{ij,k}$, population $\hat{L}_i$, and sector size $L_{i,k}$. We discuss each of these in turn.

Bilateral Trade Flows. We obtain data on bilateral trade flows $X_{ij,k}$ from the OECD's Inter-Country Input-Output (ICIO) tables. This source documents bilateral trade among 61 major exporters i and importers j, listed in Table C.1. ICIO tables report all bilateral flows, including domestic sales $X_{ii,k}$, in each sector, a feature that we use below to construct aggregate measures of expenditure, $X_{j,k} \equiv \sum_i X_{ij,k}$ and $X_j \equiv \sum_{i,k} X_{ij,k}$, as well as sales, $S_{i,k} \equiv \sum_j X_{ij,k}$ and $S_i \equiv \sum_{j,k} X_{ij,k}$. We focus our empirical analysis on 15 manufacturing sectors k defined at a similar level to the 2-digit SIC and listed in Table 1.[19] Our baseline estimates use ICIO data from 2010 but we report similar estimates from the other available cross-sections (1995, 2000, and 2005).

Population. We take our measure of population $\hat{L}_i$ 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.

Sector Size. According to the baseline model developed here, the total wage bill in a sector is equal to total sales across all destinations, $w_i L_{i,k} = S_{i,k}$. In turn, the share of total employment allocated to a given sector k is equal to the share of its sales, $S_{i,k}/S_i$. Using country i's population $\hat{L}_i$ as a measure of its labor supply L_i, we can therefore construct sector size $L_{i,k}$ as $(S_{i,k}/S_i)\hat{L}_i$.[20]

4.2 Estimates of Auxiliary Elasticities

Our estimation of scale elasticities γ_k via equation (25) requires estimates of two auxiliary parameters: (i) trade elasticities θ_k in each sector, to construct the dependent variable;

[19] We omit a 16th manufacturing sector, "Recycling and manufacturing not elsewhere classified", from the estimation, as this label covers a small amount of output in a range of highly heterogeneous activities.

[20] Given this measure of sector size, any discrepancy between population and labor supply in efficiency units is therefore also implicitly part of the error term $\varepsilon_{i,k}$ in equation (25). That is, if $L_i = (\exp \chi_i)\hat{L}_i$, then $\varepsilon_{i,k}$ also includes $\chi_i \gamma_k - \mathbb{E}[\chi_i \gamma_k | i] - \mathbb{E}[\chi_i \gamma_k | k] + \mathbb{E}[\chi_i \gamma_k]$, with our exclusion restriction (27) applying to this term as well—though, as we describe in Section 4.3, our estimates are robust to the inclusion of controls for per-capita GDP interacted with sector indicators, which lends support to this assumption.

and (ii) the elasticity of substitution between manufacturing sectors $\rho + 1$, to construct our demand-side instruments. We begin with a discussion of these auxiliary estimates.

Trade Elasticities. Nested CES preferences in (23) and (24) imply the following log-linear relationship between expenditure shares $X_{ij,k}/X_{j,k}$ and relative prices $p_{ij,k}/P_{j,k}$ within any sector k:

$$\ln(X_{ij,k}/X_{j,k}) = -\theta_k \ln(p_{ij,k}/P_{j,k}). \tag{30}$$

Equation (30) is at the core of a vast "gravity" literature, reviewed in Head and Mayer (2013), that provides estimates of trade elasticities θ_k by finding exogenous sources of variation in $p_{ij,k}$ and using fixed effects or other strategies to control for $P_{j,k}$. Rather than reestimating trade elasticities θ_k ourselves by combining the data on bilateral trade flows $X_{ij,k}$ presented in Section 4.1 with additional data on prices $p_{ij,k}$ or shifters of these prices, such as import tariffs and shipping costs, we use existing estimates of θ_k from three recent studies: Caliendo and Parro (2015), Shapiro (2016), and Giri, Yi and Yilmazkuday (2021). Although none of the structural models considered in these papers feature external economies of scale, the empirical procedures that they develop starting from equation (30) are fully consistent with our model. This is because productivity levels, regardless of whether they are exogenous (as in the three previous papers) or endogenous (as in ours), are either differenced out or absorbed by fixed effects in their empirical procedures. As such, the trade elasticities that they estimate are fully portable into our analysis, as shown in Appendix B.3. Table B.1 reports the estimates of trade elasticities from these prior studies. For our baseline analysis, we set θ_k equal to the median of their estimates, within each sector, as reported in column (4). We explore the sensitivity of our results to using alternative trade elasticities in Section 5.4.

Elasticity of Substitution Between Sectors. Nested CES preferences in (23) and (24) further imply that the share of expenditure in country j on sector k is given by

$$\ln(X_{j,k}/X_j) = -\rho \ln(P_{j,k}/P_j) + \ln \beta_{j,k}, \tag{31}$$

where $P_j = (\sum_s \beta_{j,s} P_{j,s}^{-\rho})^{-1/\rho}$ is the overall price index for country j. We estimate ρ via the following specification

$$\ln X_{j,k} = \phi_j + \phi_k - \rho \ln \hat{P}_{j,k} + \phi_{j,k}, \tag{32}$$

where ϕ_j is treated as a country fixed effect and ϕ_k is treated as a sector fixed effect. In the absence of data on the price index $P_{j,k}$, we use a proxy obtained by averaging equation (30) across all origins, $\ln \hat{P}_{j,k} \equiv \frac{1}{J} \sum_i \ln(X_{ij,k}/X_{j,k})/\theta_k$. This means that, after controlling for ϕ_j and ϕ_k, the structural error term, $\phi_{j,k}$, comprises both the preference shock from equation (31), $\ln \beta_{j,k}$, and the average of the log-productivity across all origin countries for that sector and destination, $\ln \hat{\alpha}_{j,k} \equiv \frac{1}{J} \sum_i \ln \alpha_{ij,k}$, each demeaned across countries and sectors, as further described in Appendix B.4. Our price proxy $\hat{P}_{j,k}$ captures the fact that, everything else being equal, an exporter i should tend to have a larger share of sales in destination j and sector k when the price index $P_{j,k}$ in that same destination and sector is high, with the trade elasticity θ_k controlling the steepness of this relationship, $P_{j,k} \propto (X_{ij,k}/X_{j,k})^{1/\theta_k}$. The logic is the same as in the formula for the welfare gains from trade in Arkolakis et al. (2012), with the market share of domestic producers and the trade elasticity revealing the CES price index in a destination, $P_{j,k} \propto (X_{jj,k}/X_{j,k})^{1/\theta_k}$. Since the previous logic can be applied to any exporter, our proxy instead uses the simple average across all i in order to minimize the role of idiosyncratic productivity differences.[21]

For standard reasons, OLS estimates of the demand elasticity ρ based on this expression would suffer from simultaneity bias. We therefore use an IV procedure in which the instruments for $\ln \hat{P}_{j,k}$ in (32) are formed from the product of country j's log population $\ln \hat{L}_j$ and a full set of sector indicators. The relevance of such instruments draws on the supply-side logic of our economies of scale model. In particular, we expect large countries to be relatively productive (and hence have relatively low price indices $P_{j,k}$) in sectors with relatively large scale elasticities, γ_k; this, in turn, means that the impact of population on prices should vary in distinct ways across sectors. In line with the previous logic, we will later confirm that there is a strong (inverse) correlation between the first-stage coefficients obtained here, on each sector interaction variable, and the sector's corresponding estimate of γ_k. The validity of this IV procedure relies on the exclusion restriction: $\mathbb{E}[\ln \hat{L}_j \times \phi_{j,k}|k] = 0$. Recalling that the error term includes both $\ln \beta_{j,k}$ and $\ln \hat{\alpha}_{j,k}$, it requires no systematic tendency for larger countries (those with higher values of $\ln \hat{L}_j$) to have stronger preferences for a given sector k (higher values of $\ln \beta_{j,k}$) or access to better technologies to source good k (higher values of $\ln \hat{\alpha}_{j,k}$).[22]

Our estimate of ρ is reported in Table B.3, with the corresponding first-stage results in Table B.2. The IV estimate is $\hat{\rho} = 0.28$, implying an elasticity of substitution (i.e. $\hat{\rho} + 1 = 1.28$) in the substitutes range. This IV estimate is lower than the OLS estimate (of

[21] Summing across all i also allows us to control for the impact of wages and sector sizes on expenditure shares through the sector fixed effect ϕ_k, as shown in Appendix B.4. We come back to the choice of our price proxies in the sensitivity analysis of Section 4.4.

[22] This second condition mirrors the orthogonality condition between population and productivity shocks, $\mathbb{E}[\ln \hat{L}_i \times \varepsilon_{i,k}|k] = 0$, discussed in Section 3.4. There we were ruling out systematic advantages of larger countries as sellers in specific sectors, i.e. higher values of $\frac{1}{J} \sum_j \ln \alpha_{ij,k}$. Here we are also ruling out systematic advantages of larger countries as buyers in those same sectors, i.e. higher values of $\frac{1}{J} \sum_i \ln \alpha_{ij,k}$.

$\hat{\rho} = 1.65$), as is consistent with the presence of increasing returns at the sector level. That is, when supply curves slope downwards, positive demand shocks lead to reductions in prices. Hence, the OLS estimate of the impact of prices on expenditure shares, which confounds a truly downward-sloping demand curve with the negative correlation between demand shocks and prices, will be an underestimate of $-\rho$, leading to an overestimate of ρ.

Reassuringly, our IV estimate of $\hat{\rho} = 0.28$ is similar to that seen in prior work using alternative sources of variation. For example, Redding and Weinstein (2018) estimate $\rho = 0.36$ using 4-digit U.S. import data and Oberfield and Raval (2021) obtain estimates of ρ based on 2-digit production data that range from -0.14 to 0.27 depending on the specification.[23] Sections 4.4 and 5.4 demonstrate how our broad conclusions about the effects of industrial policy are insensitive to using, both for constructing IVs and conducting counterfactual analysis, values of ρ throughout our estimated 95% confidence interval (-0.12 to 0.68), a range that also spans the estimates from other work discussed here.

4.3 Estimates of Scale Elasticities

We now return to the estimation of scale elasticities γ_k. We begin by reporting OLS estimates of γ_k in column (1) of Table 1.

All of these estimates imply precisely-estimated economies of scale, i.e. $\gamma_k > 0$, but as previously discussed, we expect OLS to deliver biased estimates of true economies of scale. For this reason we turn to the IV procedure described in Section 3.3, which relies on the demand-side instruments constructed from $\hat{L}_{i,k} \equiv \hat{\beta}_{i,k} \times \hat{L}_i$, with $\hat{\beta}_{i,k} = [x_{i,k}/(\hat{P}_{i,k})^{-\hat{\rho}}]/[\sum_l x_{i,l}/(\hat{P}_{i,l})^{-\hat{\rho}}]$ and $\hat{\rho} = 0.28$ (see Table B.3).

The IV estimates of γ_k are reported in column (2) of Table 1, with summary statistics indicating the strength of the instruments in columns (4) and (5) and the first-stage coefficients themselves summarized in Table B.4.[24] These are our preferred estimates of the strength of economies of scale within each of the 15 manufacturing sectors in our sample. The results point to substantial economies of scale—with an average scale elasticity of

[23] At higher levels of aggregation (that is, across the broad categories of agriculture, manufacturing, and services), Herrendorf et al. (2013) obtain -0.19 using US time series consumption data and Comin et al. (2021) estimate -0.50 using cross-country panel data.

[24] For each first-stage regression, Table B.4 reveals that the demand-based IV for any given sector has a strong correlation with its own sector size and a far weaker correlation with any other sector's size. Consequently, as seen in column (4) of Table 1, the conventional F-statistic from the 15 instruments in each first-stage equation is large and the Sanderson and Windmeijer (2016) F-statistic in column (5), which assesses the extent to which each first-stage is affected by independent variation in the instruments from that in the other 14 first-stages, is considerably larger. Figure B.1 provides a visualization of the first-stage relationships for each sector and Figure B.2 does the same for the corresponding reduced-form relationships.

Table 1: **Estimates of Sector-Level 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	0.24	0.22	0.25	172.8	1176.6
	(0.01)	(0.02)	(0.03)		
Textiles	0.14	0.12	0.14	105.6	1999.7
	(0.01)	(0.01)	(0.02)		
Wood Products	0.15	0.13	0.15	156.9	1126.9
	(0.02)	(0.02)	(0.02)		
Paper Products	0.18	0.15	0.17	155.9	1671.5
	(0.01)	(0.01)	(0.02)		
Coke/Petroleum Products	0.09	0.09	0.13	15.4	383.0
	(0.01)	(0.01)	(0.02)		
Chemicals	0.27	0.24	0.29	63.3	902.6
	(0.01)	(0.02)	(0.03)		
Rubber and Plastics	0.45	0.42	0.50	94.8	1426.4
	(0.03)	(0.04)	(0.06)		
Mineral Products	0.19	0.17	0.20	279.1	1091.7
	(0.02)	(0.01)	(0.02)		
Basic Metals	0.11	0.09	0.11	184.7	1234.1
	(0.01)	(0.01)	(0.02)		
Fabricated Metals	0.14	0.12	0.13	179.7	2020.7
	(0.02)	(0.02)	(0.02)		
Machinery and Equipment	0.28	0.24	0.29	122.2	1334.4
	(0.02)	(0.02)	(0.03)		
Computers and Electronics	0.09	0.08	0.09	77.6	608.3
	(0.01)	(0.01)	(0.02)		
Electrical Machinery, NEC	0.10	0.08	0.10	113.9	1193.0
	(0.01)	(0.01)	(0.02)		
Motor Vehicles	0.20	0.18	0.24	86.6	1055.1
	(0.01)	(0.01)	(0.02)		
Other Transport Equipment	0.20	0.18	0.22	26.4	745.1
	(0.01)	(0.02)	(0.03)		

Notes: Column (1) reports the OLS estimate, and column (2) the IV estimate, of equation (25). Column (3) reports the reduced form coefficients. The instruments are the log of (country population × sectoral demand shifter), interacted with sector indicators. Column (4) reports the conventional F-statistic, and column (5) the Sanderson-Windmeijer F-statistic, from the first-stage regression corresponding to the endogenous regressor formed by interacting the log of sector size with an indicator for the sector named in each row. All regressions include exporter and sector fixed effects. Standard errors in parentheses are clustered at the exporter level.

0.167—that are statistically significantly different from zero in every sector. At the same time, there is widespread heterogeneity, with estimates ranging from $\gamma_k = 0.08$ in the Computers and Electronics and Electrical Machinery, NEC sectors to $\gamma_k = 0.42$ in the Rubber and Plastics sector.[25] We can easily reject the hypothesis of coefficient equality at the 1% level. As highlighted by Proposition 2, this heterogeneity is important for the scope for industrial policy.

To understand what moments in the data lead us to the previous conclusion, note that estimating γ_k in our baseline specification is equivalent to estimating $\xi_k \equiv \theta_k \gamma_k$ in the following alternative specification,

$$\frac{1}{J}\sum_j \ln X_{ij,k} = \theta_k \delta_i + \theta_k \delta_k + \xi_k \ln L_{i,k} + \theta_k \varepsilon_{i,k},$$

where we have multiplied both the left- and right-hand side variables of equation (25) by the trade elasticity θ_k. Our IV estimates in Rubber and Electronics are $\xi_{\text{Rubber}} = 0.714$ and $\xi_{\text{Electronics}} = 0.864$. Since the (median of) existing estimates of trade elasticities in these two sectors are $\theta_{\text{Rubber}} = 1.7$ and $\theta_{\text{Electronics}} = 10.8$ (see Table B.1) we conclude that the scale elasticities must be significantly smaller in Electronics than Rubber: $\gamma_{\text{Electronics}} = 0.864/10.8 = 0.08 \ll \gamma_{\text{Rubber}} = 0.714/1.7 = 0.42$. Intuitively, if changes in employment have similar effects on bilateral trade flows in the two sectors ($\xi_{\text{Rubber}} \simeq \xi_{\text{Electronics}}$), but bilateral trade flows are much less sensitive to changes in costs in Rubber than in Electronics ($\theta_{\text{Rubber}} \ll \theta_{\text{Electronics}}$), then the impact of changes in employment on costs must be much smaller in Electronics ($\gamma_{\text{Electronics}} \ll \gamma_{\text{Rubber}}$).[26]

Another feature of these estimates is that, in each sector, the OLS estimate of γ_k is larger than its corresponding IV estimate. This upward OLS bias is to be expected in an open economy in which countries specialize in sectors where they have a comparative advantage; further, it is consistent with our finding that the elasticity of substitution between sectors, $\rho + 1$, is greater than one, which magnifies the impact of productivity

[25] This finding of significant economies of scale in manufacturing is consistent with prior estimates using alternative empirical strategies. For example, Caballero and Lyons (1992) estimate a scale elasticity for (pooled) US manufacturing sectors of 0.07-0.29 depending on the instrument used, Basu and Fernald (1997) estimate that the equivalent of γ_k for (a weighted average of) US manufacturing sectors is 0.06, Antweiler and Trefler (2002) use international trade data and a Heckscher-Ohlin-Vanek approach to estimate scale elasticities ranging from 0-0.40 among the set of manufacturing sectors in which precise estimates are obtained, and Costinot et al. (2019) estimate a value of $\gamma_k = 0.25$ for the global pharmaceutical sector.

[26] Across sectors, our empirical results imply less variation in the elasticities ξ_k of average bilateral trade flows with respect to employment, which range from 0.71 to 0.97, than in the existing estimates of trade elasticities θ_k, which range from 1.7 to 10.8 in Table B.1. As a result, the correlation between scale elasticities γ_k and trade elasticities θ_k is -0.87.

differences on specialization.[27]

Finally, column (3) reports the reduced-form parameter estimates of the impact of predicted (log) demand $\ln \hat{L}_{i,k}$ in each sector on the dependent variable, $Y_{i,k} \equiv (\frac{1}{J}\sum_j \ln X_{ij,k})/\theta_k$. We see that countries with higher predicted demand in a sector tend to have larger values of $Y_{i,k}$ in that sector. This is again consistent with the existence of increasing returns at the sector level, which implies that positive shocks to domestic demand cause lower prices and, in turn, greater exports. This is a manifestation of the home-market effect.[28]

4.4 Sensitivity to Alternative Samples and Specifications

As characterized by Propositions 1 and 2, the estimates of scale economies γ_k in Table 1 shape both the structure of and the gains from optimal industrial policy. Before turning to such implications of our γ_k estimates, we describe three exercises that probe their sensitivity to alternative samples and specifications.

Heterogeneity over time. The estimates of γ_k in Table 1 are obtained from a single cross-section (2010). As a first robustness check, we re-estimate the scale elasticity parameters γ_k in each of the other cross-sections that are available to us (from 1995, 2000, and 2005), as well as a specification that pools across all years, 1995-2010. These estimates, displayed in Table B.5, are very similar to each other and to those reported in Table 1—indeed, in all three years the correlation with our 2010 baseline is 0.99. This finding highlights how our conclusions are not specific to one particular year. More broadly, it is also consistent with our focus on long-run scale elasticities that we do not expect to vary over time.

Alternative Instrumental Variables. To construct our instrumental variables, we use demand residuals $\hat{\beta}_{i,k} = [x_{i,k}/(\hat{P}_{i,k})^{-\hat{\rho}}]/[\sum_l x_{i,l}/(\hat{P}_{i,l})^{-\hat{\rho}}]$ whose value depends both on our estimate of the upper-level elasticity $\hat{\rho}$ as well as our price proxies, $\hat{P}_{i,k}$. Our second robustness exercise explores the sensitivity of our estimates of γ_k to using alternative demand residuals. We start by recomputing $\hat{\beta}_{i,k}$ under alternative values of $\hat{\rho} = -0.9$, 0, and 2, a range that goes well beyond the 95th confidence interval around our point estimate $\hat{\rho} = 0.28$ in Table B.3. It includes the Cobb-Douglas special case, $\hat{\rho} = 0$, for which no price

[27] An estimate of ρ that is close to 0 and trade volumes that are low for the average country, however, imply that the OLS bias is not large, ranging between 0 and 0.04 across sectors in Table 1.

[28] Echoing that view, the first-stage results (reported in Table B.2) corresponding to the estimation of the elasticity substitution ρ show that the impact of country size on sector-level prices is negative in all sectors but one. Further, the correlation between these sector-specific first-stage coefficients and the estimates of $\widehat{\gamma}_k$ is -0.90, 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.

proxies are required. Even if there were no measurement error in our price proxies, $\hat{P}_{i,k} = P_{i,k}$, using a value of $\hat{\rho}$ that is distinct from its true value ρ creates a wedge between the estimated and the true demand residuals, $\hat{\beta}_{i,k}/\beta_{i,k} = (P_{i,k})^{\hat{\rho}-\rho}/[\sum_l \beta_{i,l}(P_{i,l})^{\hat{\rho}-\rho}]$. Thus, even if the true demand residuals satisfy the exclusion restriction, $\mathbb{E}[\ln \beta_{i,k} \times \varepsilon_{i,k}|k] = 0$, this restriction may not be satisfied for the estimated demand residuals, $\mathbb{E}[\ln \hat{\beta}_{i,k} \times \varepsilon_{i,k}|k] \neq 0$, leading to a bias in our IV estimates of scale elasticities γ_k whose magnitude would scale up with the difference between $\hat{\rho}$ and ρ. The estimates of γ_k displayed in columns (2)-(4) of Table B.6 conform with that observation, though they are all close to our baseline estimates, also reported in column (1) of Table B.6. These small differences are consistent with the small difference between the OLS and IV estimates of γ_k documented in Table 1.

The fact that we use price proxies $\hat{P}_{i,k}$ rather than the true prices $P_{i,k}$ is another source of measurement error in the construction of the demand residuals $\hat{\beta}_{i,k}$ and a potential source of bias in the IV estimation of γ_k. Equation (30) implies that the log difference between the two prices is equal to the average of the log prices across exporters in a given sector and destination, $\ln \hat{P}_{i,k} - \ln P_{i,k} = -\frac{1}{J}\sum_j \ln p_{ji,k}$. Hence, even if we knew the true upper-level elasticity, $\hat{\rho} = \rho$, there would still be a wedge between the estimated and the true demand residuals, $\hat{\beta}_{i,k}/\beta_{i,k} = \exp(-\frac{1}{J}\sum_j \ln p_{ji,k})^{\rho}/[\sum \beta_{i,l}\exp(-\frac{1}{J}\sum_j \ln p_{jl,k})^{\rho}]$. This would lead to upward bias in the estimates of γ_k if consumers in more productive countries also tend to face lower average prices (for example because country i's local productivity in sector k, $A_{ii,k}$, mechanically lowers domestic prices $p_{ii,k}$) or downward bias if the opposite pattern prevails (for example because country i's higher productivity in sector k tends to lower its neighbors' employment, $L_{j,k}$, and in turn, to raise their export prices, $p_{ji,k}$). To address this endogeneity concern, we consider an alternative price proxy $\hat{P}^{alt}_{i,k} \equiv \frac{1}{J}\sum_{j\in D(i)} \ln(X_{ji,k}/X_{i,k})/\theta_k$ that builds on the same strategy as in Section 4.2 of inferring price indices from exporters' markets shares, but only incorporates the exporters located in a "doughnut" $D(i)$ that excludes all countries whose distance to country i is less than half the median distance between countries in our sample. The estimates of γ_k using these alternative price proxies are reported in column (5) of Table B.6. Reassuringly, they are very similar to the baseline estimates reported in column (1).

Additional Controls. As discussed above, our estimation procedure requires that each country's comparative advantage—deriving from cost differences across sectors k and exporting countries i, which are absorbed in the error term $\varepsilon_{i,k}$ of equation (25)—is orthogonal to our demand-side instruments. One test of this requirement is that adding observable proxies for comparative advantage should not substantially change our estimates of γ_k. We report here several variants of this idea that are based on Ricardian comparative advantage, deriving from productivity differences alone, consistent with the model of Section 2. Section 6 presents additional robustness checks based on other sources of cost differences (coming from multiple factors of production and tradable intermediate inputs) across countries and sectors.

A prominent source of Ricardian comparative advantage stems from heterogeneity in institutions across countries and the differential implications that those institutions have for productivity across sectors. Following the empirical literature surveyed by Nunn and Trefler (2014), we model this as an interaction between two components: (i) country-level proxies for institutional quality and (ii) sector-level characteristics. For the country component, we use both a measure of contract enforcement and a measure of financial development, thereby encompassing the sources of comparative advantage stressed in Levchenko (2007) and Nunn (2007) as well as Beck (2002) and Manova (2008).[29] For the sector component, we use a full set of sector indicator variables, which has the benefit of controlling for any form of systematic Ricardian comparative advantage based on contract enforcement or financial development. As an extension to this procedure we add regressors formed from interactions between the exporter's per capita GDP and sector indicators, which hence controls for any potential reason for relatively rich countries to be differentially productive in certain sectors.[30]

The results from both of these exercises are reported in Table B.7. In each case, we see only minor changes in the estimates of γ_k as compared to our baseline (a correlation of 0.99 in both extensions), suggesting that our IV strategy utilizes variation in sector size driven by factors that are orthogonal to observable sources of Ricardian comparative advantage. This lends credence to the view that our estimates draw only on demand-side variation in order to identify the supply-side scale economies γ_k.

5 The Textbook Gains from Industrial Policy

5.1 The Calibrated Economy

We focus on a world economy that comprises the 61 countries from the ICIO dataset, the 15 manufacturing sectors for which we have estimated scale economies, as well as the 17 non-manufacturing sectors also included in the ICIO dataset. Throughout our

[29] Following Nunn (2007), we measure contract enforcement by the "rule of law" variable due to Kaufmann, Kraay and Mastruzzi (2003) (in 1997-98) and financial development by the (log of the) ratio of private bank credit to GDP (in 1997).

[30] We obtain data on per-capita GDP by dividing the real output variable, "RGDPO", in the Penn World Tables by $\hat{L}_i$.

quantitative exercise, we maintain the parametric restrictions imposed in equations (22)-(24). Thus, scale elasticities are given by γ_k, whereas trade elasticities are given by θ_k and the elasticity of substitution between sectors is given by $1+\rho$.

To calibrate γ_k, we use the IV estimates reported in column (2) of Table 1 for all manufacturing sectors and set $\gamma_k = 0$ for all non-manufacturing sectors. 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 some positive value. We consider alternative cases in the sensitivity analysis of Section 5.3.

To calibrate θ_k, we use the median values of the trade elasticities from recent studies reported in column (5) of Table B.1 for each manufacturing sector, in line with the empirical analysis of Section (4), and we set $\theta_k = 5.8$ for all non-manufacturing sectors, which is the median elasticity throughout manufacturing.

Finally, we set $\rho = 0.28$, in line with the IV estimate for manufacturing sectors reported in column (2) of Table B.3. This implies the same elasticity of substitution between manufacturing and non-manufacturing sectors as within the set of manufacturing sectors. As we discuss further below, this leads to higher gains from industrial policy than if we had assumed an elasticity of substitution between manufacturing and non-manufacturing below one, as estimated in some recent studies, e.g., Comin, Lashkari and Mestieri (2021) and Cravino and Sotelo (2019). Overall, our calibration choices are on the more aggressive side, to give a chance for industrial policy to yield high gains—yet as we discuss below, even with these choices, the gains are on the low side.

Following the theoretical analysis of Section 2.3, we treat each country as small. Under the assumption of nested CES preferences, as a given country j becomes arbitrarily small relative to other countries, sector-level prices and expenditures, $P_{j,k}$ and $X_{j,k}$, become independent of the prices of its exports, $p_{ji,k}$, and the price elasticities $\epsilon^{p*}_{ij,k}$ converge to $-1/(1+\theta_k)$. Under the normalization $\bar{s}_j = \bar{t}_j = 0$, the optimal industrial and trade policy described in Proposition 1 are therefore given by

$$s^*_{j,k} = \frac{\gamma_k}{1+\gamma_k}, \text{ for all } k, \tag{33}$$

$$t^{x*}_{ji,k} = \frac{1}{1+\theta_k}, \text{ for all } k \text{ and } i \neq j, \tag{34}$$

$$t^{m*}_{ij,k} = 0, \text{ for all } k \text{ and } i. \tag{35}$$

Given estimates of the optimal policies, we compute counterfactual equilibria under those policies using exact hat algebra, as in Dekle, Eaton and Kortum (2008), under the

Table 2: Gains from Optimal Policies, Selected Countries

Country	Fully Optimal Policy (1)	Trade Policy (2)	Industrial Policy (3)
United States	0.72%	0.50%	0.58%
China	0.97%	0.43%	0.83%
Germany	1.16%	1.86%	1.10%
Ireland	1.86%	3.96%	1.78%
Vietnam	1.74%	1.80%	1.13%
Avg., Unweighted	**1.34%**	**1.49%**	**0.98%**
Avg., GDP-weighted	**0.94%**	**0.87%**	**0.76%**

Notes: Column (1) reports the gains associated with fully optimal policies, as defined in equation (13); column (2) reports the gains associated with optimal trade policy, as defined in equation (15); and column (3) reports the gains associated with optimal industrial policy, as defined in equation (14). All gains are reported as a share of initial income.

assumption that the initial equilibrium observed in the data (corresponding to 2010, the final year in the ICIO dataset) features neither taxes nor subsidies and that trade deficits D_i correspond to transfers across countries which we hold fixed in the counterfactual equilibria. The full non-linear system of equations that determines these counterfactual changes can be found in Appendix C.1.

5.2 Baseline Results

Table 2 shows the gains from optimal trade and industrial policy for a subset of countries as well as the average across countries, both unweighted and weighted by GDP. Results for each of the 61 countries in our dataset can be found in Table C.1 in Appendix C.2.

Column (1) reports the gains from the fully optimal policy, defined in equation (13) and expressed as a share of initial income. This corresponds to the welfare gains of going from laissez-faire to export taxes equal to $1/(1+\theta_k)$ and employment subsidies equal to $\gamma_k/(1+\gamma_k)$. The separate gains from optimal industrial and trade policy, defined in equations (14) and (15) and again expressed as a share of initial income, are reported in columns (2) and (3) respectively.

The results in Table 2 reveal that the gains from fully optimal policy (column 1) are on average 1.34%. The gains from optimal industrial policy (column 3) are smaller than those from optimal trade policy (column 2): focusing on the unweighted average across countries, these gains are 0.98% and 1.49%, respectively. Interestingly, the gains from fully optimal policy are often lower than the gains from optimal trade policy, reflecting the fact

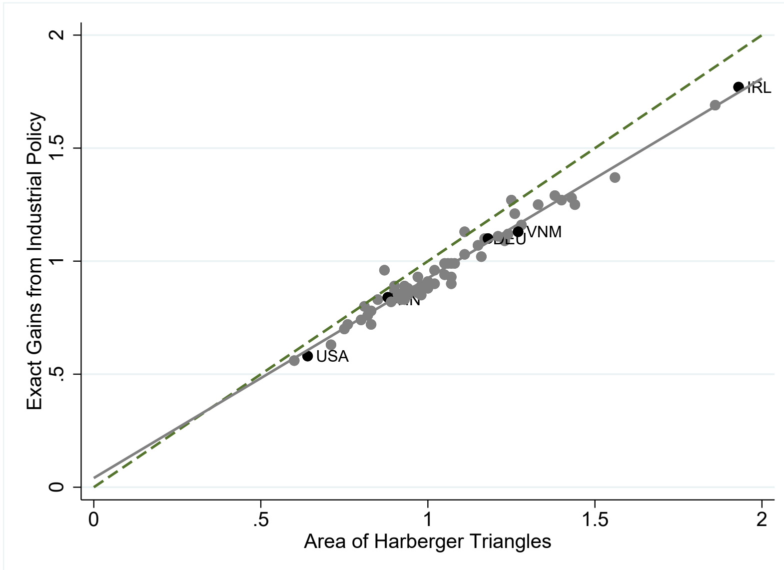


Figure 2: Exact Policy Gains versus Areas of Harberger Triangles

Notes: Figure 2 reports the exact gains from industrial policy (as a share of initial income) on the y-axis, as defined in equation (14), against the second-order approximation from Proposition 2 on the x-axis, as described in equation (17). The solid line is the line of best fit and the dashed line is the 45° line.

that using only industrial policy ignoring their terms-of-trade implications leads to large welfare losses in some countries.

The gains from industrial policy are closely aligned with those predicted by the areas of Harberger triangles in Proposition 2, as can be seen in Figure 2, which reports the exact gains from industrial policy generated by the model plotted against the second-order approximation in equation (17). A linear regression of the exact gains on our second-order approximation yields a slope equal to 0.88 and an R^2 equal to 0.97.[31]

Modest gains from industrial policy do not reflect modest "wedges." According to our estimates of scale elasticities, if all labor were to reallocate to the manufacturing sector with the largest scale elasticity $\gamma_k = 0.42$, the average welfare gains predicted by the area of the Harberger triangles would be equal to 18.4%.[32] Modest gains instead reflect the fact that only modest labor reallocations take place from "low-wedge" to "high-wedge" sectors. This is due to a relatively low elasticity of substitution between sectors and to

[31] Harberger triangles also provide good approximations for the gains from fully optimal policy and optimal trade policy, with corresponding slope and R^2 given by 1.40 and 0.90 for the fully optimal policy and 0.80 and 0.98 for optimal trade policy.

[32] Specifically, the sector with the largest elasticity $\gamma_k = 0.42$ is Rubber and Plastic. If labor were to fully reallocate to that sector, the welfare change in a given country j predicted by equation (17) would be $\frac{1}{2} \times [(1 - L_{j,\text{Rubber}}/L_j) \times 0.42 - \sum_{k \neq \text{Rubber}} (L_{j,k}/L_j)\gamma_k]$. The average of the previous number across all countries in our dataset is equal to 0.184.

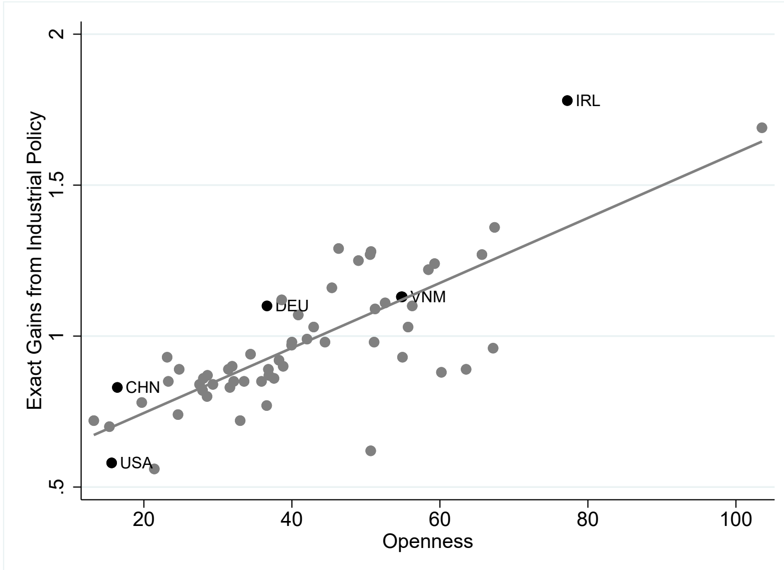


Figure 3: Industrial Policy Gains versus Openness

Notes: Figure 3 reports the exact gains from industrial policy (as a share of initial income) on the y-axis, as defined in equation (14), against openness on the x-axis, measured as exports plus imports over gross output.

trade elasticities that are far from infinity. Going back to the textbook case for industrial policy illustrated in Figure 1, our empirical analysis points to Harberger triangles that have non-trivial heights, but small bases.

Across countries, our quantitative results nevertheless exhibit substantial heterogeneity. Smaller countries, in particular, tend to gain more from industrial policy than larger ones. This is revealed by the fact that the simple average in column (3) is higher than the corresponding GDP-weighted average. As an example, Ireland has gains from industrial policy that are roughly three times higher than those of the United States (1.78% vs 0.58%).[33]

A key difference between small and large countries is their degree of openness to trade. Smaller countries are more open and more open countries gain more from optimal industrial policy. This is illustrated in Figure 3, 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 their Harberger

[33]Smaller countries also gain more from trade policy and the fully optimal policy, as can also be seen by comparing the simple averages in column (1) and (2) to the corresponding GDP-weighted averages. The reason why smaller countries gain more from optimal trade policy is simple: such a policy improves a country's terms-of-trade, and since small countries tend to trade more, they benefit more from that improvement.

triangles have larger bases, leading to bigger gains from industrial policy. In line with the previous intuition, a cross-country regression of $\sum_k |(\Delta L_{j,k})^I / L_j|$ on openness yields a strong positive relationship with an R^2 of 0.84.

An additional implication of the Harberger formula described in equation (17) is that, for given elasticities $(\Delta L_{j,k})^I / L_{j,k}$, countries with higher employment in sectors with stronger scale economies should benefit more from industrial policy. Our quantitative results are also consistent with this observation: countries with a higher correlation between L_k/L and γ_k do indeed have higher gains from optimal industrial policy. However, this channel accounts for a smaller fraction of the cross-country variation in the gains compared to the openness channel discussed above: a regression of the gains from industrial policy on openness yields $R^2 = 0.62$, while adding the variable $\sum_k L_k/L \cdot \gamma_k$ as a predictor increases the R^2 to 0.77.

5.3 Gains from Industrial Policy in the Presence of Trade Agreements

The previous quantitative results assume that countries are free to pursue their unilaterally optimal trade policies. In practice, explicit trade agreements or implicit threats of foreign retaliation may prevent countries from doing so. How would such considerations affect the gains from industrial policy?

We address this issue under two benchmark scenarios. In the first case, countries are forced to set zero trade taxes, but still face incentives to manipulate their terms-of-trade using industrial policy, as in Lashkaripour and Lugovskyy (2018). In the second, we assume that, despite the availability of other policy instruments, trade agreements have been designed to internalize terms-of-trade externalities and restore global efficiency, as in Bagwell and Staiger (2001).[34]

Under the first scenario, we numerically find the employment subsidies that maximize utility in a given country conditional on zero trade taxes. As mentioned above, these constrained-optimal employment subsidies involve a compromise between the textbook Pigouvian motive of internalizing production externalities and the goal of improving the country's terms-of-trade.[35] Column (2) of Table 3 reports the gains from industrial pol-

[34] Another potentially interesting scenario would be to consider the gains from industrial and trade policy in an environment where all countries simultaneously impose policies non-cooperatively. Although we stop short of analyzing such Nash equilibria, we expect countries to be less open when they all choose trade taxes in a non-cooperative manner and, in turn, the gains from industrial policy to be smaller, consistent with the results of Figure 3. Indeed, when considering counterfactual autarkic economies, we find gains from industrial policy that are 0.46% on average.

[35] As an example, consider an economy that exports all its output. In this case employment subsidies equal to $1 - (\frac{1}{1+\gamma_k})(\frac{1+\theta_k}{\theta_k})$ would perfectly replicate the effect of both the employment subsidies and export

Table 3: Gains from Constrained and Globally Efficient Industrial Policies, Selected Countries

Country	Baseline Industrial Policy (1)	Constrained Industrial Policy (2)	Globally Efficient Industrial Policy (3)
United States	0.58%	0.41%	0.55%
China	0.83%	0.65%	0.46%
Germany	1.10%	0.48%	-0.47%
Ireland	1.78%	1.01%	-2.13%
Vietnam	1.13%	1.00%	1.87%
Avg., Unweighted	**0.98%**	**0.73%**	**0.49%**
Avg., GDP-Weighted	**0.76%**	**0.53%**	**0.35%**

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.

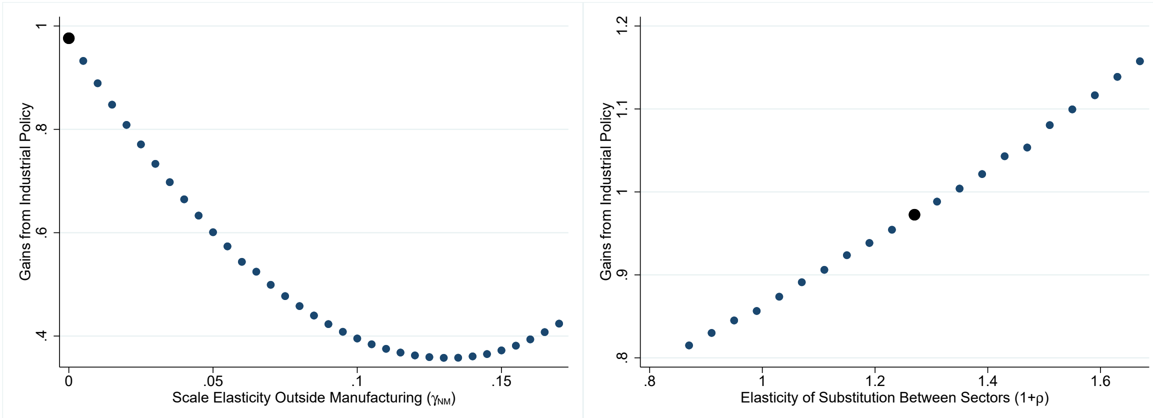
icy under this first scenario.[36] For convenience, column (1) reports again the gains from industrial policy when trade policy is unconstrained as well, i.e., column (2) of Table 2. On average, the gains from this type of constrained industrial policy are a bit lower than those from the optimal industrial policy reported in our baseline, with the unweighted world average decreasing from 0.98% to 0.73%.

Under the second scenario, we assume that employment subsidies are chosen in a Pareto-efficient manner, with lump-sum transfers between countries available if necessary. Hence, only the Pigouvian motive remains and all countries set employment subsidies satisfying $s_{j,k} = \gamma_k/(1+\gamma_k)$ for all j,k. The gains from industrial policy (gross of lump-sum transfers between countries, if any) in this case are reported in column (3) of Table 3. The GDP-weighted average of the gains associated with this policy is 0.49%, but with (gross of transfer) gains of 1.87% in Vietnam and losses of 2.13% in Ireland. Such welfare losses derive from adverse terms-of-trade effects: larger employment subsidies in sectors with high scale elasticities cause an expansion of these sectors and a deterioration of the terms-of-trade of countries specializing in them.[37]

taxes in the unconstrained policy case. As long as there is some sector in which part of domestic production is sold at home, however, the constrained-optimal employment subsidies will deviate from those and the corresponding gains will be strictly lower than those when policy is unconstrained.

[36] Results for all countries in our dataset can be found in Table C.2.

[37] The correlation between the gains from the coordinated industrial policy considered in column (3) of Table 3 and the comparative advantage of countries in high scale elasticity sectors (which we measure simply as the correlation, within each country, between its sector-level net exports and sector scale elasticities), is equal to -0.50.



(a) Scale Elasticity Outside Manufacturing **(b) Elasticity of Substitution Between Sectors**

Figure 4: Gains from Industrial Policy, Alternative Parameter Values

Notes: Figure 4a reports the (global, unweighted average) gains from industrial policy, as defined in equation (14), for different values of the scale elasticity outside manufacturing, γ_{NM}. Figure 4b does the same for different values of the elasticity of substitution between sectors, $1+\rho$. Large circles indicate baseline parameter values used in Table 2.

5.4 Sensitivity to Calibrated Parameters

We conclude this section by exploring the sensitivity of our findings with respect to the main parameters used in our quantitative exercise: scale elasticities γ_k, trade elasticities θ_k, and the elasticity of substitution between sectors ρ.

Scale Economies in Non-Manufacturing. We begin with the implications of the value of scale economies in non-manufacturing sectors, a parameter that we denote γ_{NM}. Our baseline assumption of $\gamma_{NM} = 0$—in which only manufacturing sectors exhibit economies of scale—fits the traditional view behind industrial policy. But in the absence of strong evidence to suggest that industrial sectors have superior economies of scale, it is important to explore the quantitative implications of this assumption.

To shed light on this issue, Figure 4a plots the average of the estimated gains from optimal industrial policy (across all countries in the world) as a function of γ_{NM}.[38] Not surprisingly, as we start raising γ_{NM} from zero, the relative size of manufacturing and non-manufacturing sectors in the competitive equilibrium gets closer to its optimal value and the gains from industrial policy fall. Indeed, the gains are minimized when γ_{NM} is equal to around 0.13, close to the average of the manufacturing sector γ_k values of 0.167.

[38] We continue to calculate the gains from optimal policy as described in equations (33)-(35); that is, now that $\gamma_{NM} > 0$ the optimal policy includes an employment in non-manufacturing sectors.

Our baseline finding that the gains from optimal industrial policy are relatively modest therefore appears to hold across a wide range of reasonable values for scale elasticities outside manufacturing, in line with the fact that these modest gains derive from Harberger triangles with modestly sized bases rather than heights.

Elasticity of Substitution Across Sectors. We next consider the role played by the elasticity of substitution $1 + \rho$. In our baseline analysis, we have used our IV estimate $\rho = 0.28$. Table B.3 implies a 95% confidence interval of $[-0.12, 0.68]$ for this estimate. In Figure 4b, we therefore plot the average gains from optimal industrial policy throughout this range. As discussed above, we expect that a higher ρ would lead to larger reallocations in response to optimal industrial policy and, in turn, larger welfare gains. Qualitatively, this is confirmed in Figure 4b. Quantitatively, however, we see that even for $\rho = 0.68$, the average gains from industrial policy are only 1.18%.

Trade Elasticities. Trade elasticities play a dual role in our analysis: first, they affect the demand functions that we use to estimate scale elasticities γ_k in equation (25); and second, they shape the gains from optimal policy for any given value of γ_k. Our baseline estimates used θ_k from prior estimates in the literature, taking the median estimate of θ_k, within each sector, across three recent studies in the literature. We now instead estimate the gains from industrial and trade policy when drawing trade elasticities from each of those three studies in turn. In each case we repeat the entire two-step estimation procedure from Section 4 for the new vector of θ_k values under consideration.[39]

The results of this exercise are shown in Table C.3 in Appendix C.4. Although there is a wide range of trade elasticity estimates across these three studies (as seen in Table B.1), the average estimated gains from industrial policy remain fairly similar for two out of three of these studies, about 1.15% as compared to the baseline average gains of 0.98%. The outlier is column (3) of Table C.3, which draws on estimates from Caliendo and Parro (2015) that contain a much broader range of trade elasticities across sectors, from 0.4 to 64.9, than the other studies. Using these parameter estimates leads, in turn, to significantly larger gains from industrial policy, equal to 3.28% on average. Our baseline analysis, which focuses on the median of trade elasticities across studies, is purposefully designed to attenuate the potential effects of outliers.

[39] In some cases this procedure results in $\gamma_k\theta_k \geq 1$, which leads to multiplicity of equilibria, as pointed out by Kucheryavyy et al. (2017). Since the solution of the counterfactual equilibrium under optimal policy becomes numerically unstable as $\gamma_k\theta_k$ approaches one from below, we cap the value of γ_k such that $\gamma_k\theta_k$ is never higher than 0.98.

6 Beyond Ricardian Economies

In previous sections, we have explored the structure and welfare consequences of optimal industrial and trade policy in a Ricardian environment in the tradition of Graham (1923), Chipman (1970), and Ethier (1982). In this final section, we extend our earlier analysis to richer environments featuring multiple factors of production and input-output linkages.

6.1 General Environment

Consider a generalized version of the environment of Section 2.1. Each country is now endowed with both labor and physical capital, L_i and K_i, and all production functions now use labor, physical capital, and intermediate goods from other countries and sectors,

$$y_{ij,k} = A_{ij,k} E_k(Z_{i,k}) z_{ij,k},$$
$$z_{ij,k} = f_k(\ell_{ij,k}, k_{ij,k}, m_{ij,k}),$$

where $m_{ij,k} \equiv \{m_{oij,sk}\}$ denotes the vector of intermediate inputs, $m_{oij,sk}$, from a given origin country o and source sector s used by firms producing good k in country i and exporting it to country j, $k_{ij,k}$ denotes the amount of physical capital that they use, and f_k is homogeneous of degree one in all inputs.[40] The relevant measure of sector size is now given by the total use of the composite input in country i and sector k,

$$Z_{i,k} = \sum_j z_{ij,k}. \tag{36}$$

In terms of policy, each country j's government can still tax imports and exports, at rates $t^m_{ij,k}$ and $t^x_{ji,k}$, as well as subsidize purchases of the composite input at rate $s_{j,k}$, leading to the arbitrage condition

$$v_{j,k} = (1 - s_{j,k}) w_{j,k}, \tag{37}$$

where $w_{j,k}$ denotes the price received by firms from country j producing sector k's composite input and $v_{j,k}$ denotes the price paid by the firms purchasing it. The rest of the

[40]We allow for Hicks-neutral differences across countries and sectors, but restrict sector-level production functions, f_k, to be constant across countries. In our theoretical analysis, this restriction can be dispensed with. In our quantitative analysis, the assumption of a common f_k across countries helps to avoid computational problems arising from the combination of economies of scale and very imbalanced input-output tables in some countries, while still allowing differences in factor endowments to act as a source of comparative advantage.

equilibrium conditions can be found in Appendix A.4.

6.2 Revisiting the Structure of Optimal Policy

Within this environment, we can use the same strategy as in Section 2.3 to characterize the optimal production subsidies, $s^*_{j,k}$, and trade taxes, $t^{x*}_{ji,k}$ and $t^{m*}_{ij,k}$. The same Pigouvian logic leads to the following generalization of Proposition 1.

Proposition 3. *In the general environment of Section 6.1 the optimal industrial and trade policy are such that:* (i) $s^*_{j,k} = \epsilon^{E*}_{j,k}/(1+\epsilon^{E*}_{j,k})$*;* (ii) $t^{m*}_{ij,k} = \bar{t}_j$*; and* (iii) $t^{x*}_{ji,k} = 1-(1+\bar{t}_j)(1+\epsilon^{p*}_{ij,k})$ *for some* $\bar{t}_j$*.*

The formal proof can be found in Appendix A.4.2. The only difference between Propositions 1 and 3 comes from the level of industrial policy. Since the total supply of the composite input generating the externality, $\sum_k Z_{i,k}$, is elastic, the absolute level of the scale elasticity is no longer irrelevant and there is no longer an indeterminacy in the level of the input subsidy, $\bar{s}_j$: optimal subsidies now require $\bar{s}_j = 0$.

To go from theory to data, we maintain the same parametric restrictions as in Section 3.2—that is, E_k and u_j satisfy equations (22)-(24)—and assume that f_k is nested CES,

$$f_k(\ell_{ij,k}, k_{ij,k}, m_{ij,k}) = [(\ell_{ij,k})^{a_k}(k_{ij,k})^{1-a_k}]^{b_k}\prod_s [M_s(\{m_{oij,sk}\}_o)]^{b_{sk}}, \tag{38}$$

$$M_s(\{m_{oij,sk}\}_o) = (\sum_o m_{oij,sk}^{\frac{\theta_s}{1+\theta_s}})^{\frac{1+\theta_s}{\theta_s}}, \tag{39}$$

with $b_k + \sum_s b_{sk} = 1$ and $a_k \in [0,1]$, with average shares consistent with the OECD's ICIO tables. The estimation of scale elasticities γ_k can then proceed following the same steps as in Section 3 given knowledge of the sector-level production function f_k that maps labor, physical capital, and intermediate goods into a country-and-sector specific composite factor. The counterpart of our baseline specification (25) is

$$Y^{\text{IO}}_{i,k} = \delta_i + \delta_k + \gamma_k \ln Z_{i,k} + \varepsilon_{i,k}, \tag{40}$$

where $Y^{\text{IO}}_{i,k} \equiv \frac{1}{J}\sum_j(\ln X_{ij,k}/\theta_k) + (1-a_k)b_k\ln(r_i/w_i) + \sum_s b_{sk}\ln(P_{i,s}/w_i)$ now also controls for differences in relative factor prices, r_i/w_i, and relative intermediate good prices, $P_{i,s}/w_i$, as described in Appendix B.5.1. This adjustment is necessary to isolate the impact of external economies of scale on quantities demanded.

Figure 5 plots the optimal industrial policy $s^*_{j,k} = \gamma_k/(1+\gamma_k)$ estimated in the general environment of Section 6.1 against the optimal industrial policy $s^*_{j,k} = \gamma_k/(1+\gamma_k)$ esti-

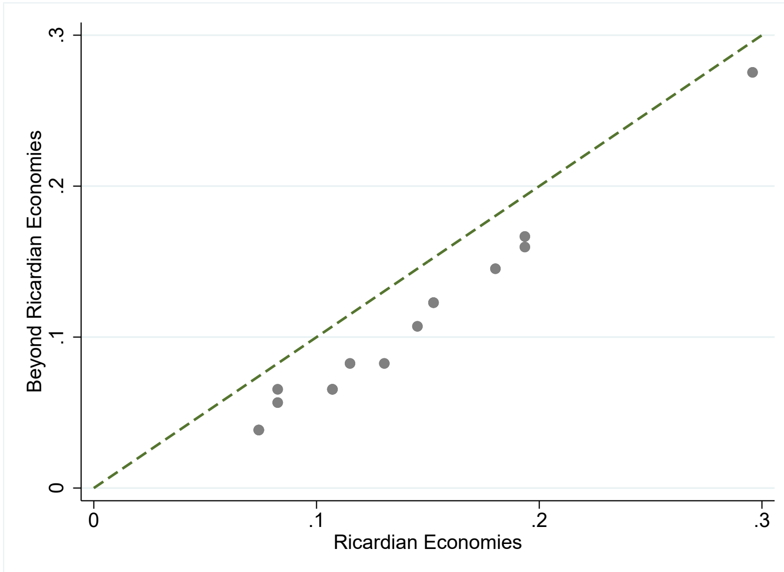


Figure 5: Optimal Industrial Policy, Beyond Ricardian Economies

Notes: The y-axis reports the optimal industrial policy $s^*_{j,k} = \gamma_k/(1+\gamma_k)$ estimated for 15 manufacturing sectors in the environment with capital and input-output linkages, as described in Section 6.1, whereas the x-axis reports the optimal industrial policy $s^*_{j,k} = \gamma_k/(1+\gamma_k)$ estimated in the baseline environment, as described in Section 2.1.

mated in the Ricardian environment of Section 2.1 for the same 15 manufacturing sectors. All observations are close to but below the 45-degree line (illustrated in green).[41] The average optimal subsidy in manufacturing sectors is about 28% lower in the presence of physical capital and input-output linkages. This reflects the fact that our baseline estimation does not control for variation in the price of intermediate goods. Since each sector and country sources a large fraction of its inputs from itself, countries with larger sectors also tend to have lower input prices. Controlling for input prices in our modified estimating equation (40) removes this source of upward bias in our baseline estimation. Remarkably, the identity of the sectors that should be subsidized remains unchanged, with a correlation between the optimal policies estimated in the two models equal to 0.99.[42]

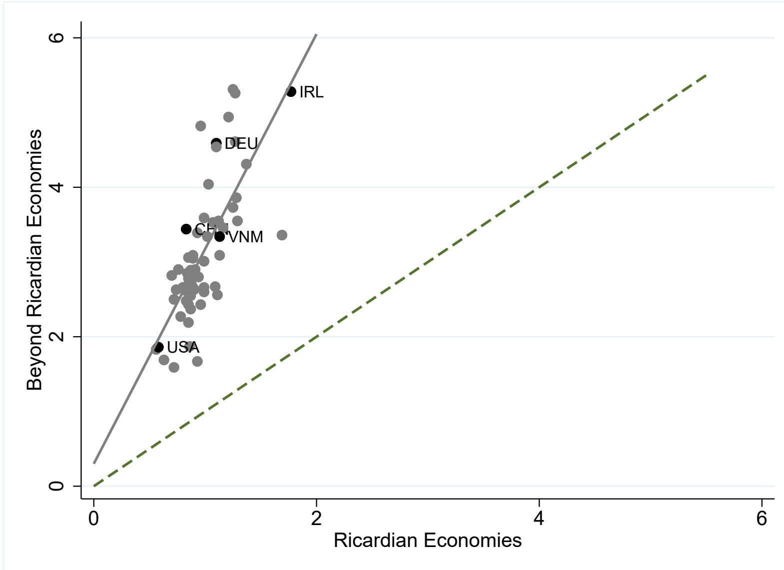


Figure 6: Gains from Industrial Policy, Beyond Ricardian Economies

Notes: Figure 6 reports the gains from industrial policy, as defined in equation (14), in the environment with capital and input-output linkages, as described in Section 6.1, on the y-axis, against the gains from industrial policy in the baseline environment, as described in Section 2.1, on the x-axis.

6.3 Revisiting the Gains from Optimal Policy

While the structure of optimal policy is very similar to that in our baseline environment, the consequences of such policy are not. As can be seen from Figure 6, the gains from optimal industrial policy are significantly larger in the presence of physical capital and input-output linkages. The average gains estimated here are equal to 3.11%, about three times larger than in a Ricardian environment.[43] To understand why the gains from optimal policy are higher in this more general environment, it is useful to turn to the following generalization of Proposition 2.

Proposition 4. *In the general environment of Section 6.1, up to a second-order approximation,*

[41] All estimates of γ_k remain statistically significant at the 5% level and the instruments remain strong by conventional standards (with the lowest first-stage SW F-statistic equal to 132.0). Full estimates are reported in Table B.8.

[42] Since none of the estimates of trade elasticities discussed in Appendix B.3 are affected by the presence of physical capital and input-output linkages, the structure of optimal trade policy is trivially unchanged.

[43] The counterpart of the exact hat algebra of Section 5 used for the counterfactual analysis in this section can be found in Appendix C.5, whereas the gains from industrial policy for all countries can be found in Table C.4 in Appendix C.5.2, along with the gains from trade policy and the fully optimal policy. For completeness, we also report the sensitivity of these numbers to the values of the scale elasticity outside manufacturing (γ_{NM}), the elasticity of substitution between sectors (ρ), and the trade elasticity (θ_k) in Appendix C.5.4.

the welfare gains from optimal taxes satisfy

$$\frac{W_j^*}{I_j} \simeq \frac{1}{2}\sum_k \frac{w_{j,k}Z_{j,k}}{I_j} \cdot \frac{(\Delta Z_{j,k})^*}{Z_{j,k}} \cdot \epsilon_{j,k}^{E*} + \frac{1}{2}\sum_{i\neq j,k} \frac{\bar{p}_{ji,k}y_{ji,k}}{I_j} \cdot \frac{(\Delta y_{ji,k})^*}{y_{ji,k}} \cdot \epsilon_{ji,k}^{p*}, \tag{41}$$

where $(\Delta Z_{j,k})^*$ *and* $(\Delta y_{ji,k})^*$ *denote the input and export changes associated with imposing* τ_j^* *and all other variables are evaluated under laissez-faire. Likewise, up to a second-order approximation, gains from industrial and trade policy satisfy*

$$\frac{W_j^I}{I_j} \simeq \frac{1}{2}\sum_k \frac{w_j Z_{j,k}}{I_j} \cdot \frac{(\Delta Z_{j,k})^I}{Z_{j,k}} \cdot \epsilon_{j,k}^{E*}, \tag{42}$$

$$\frac{W_j^T}{I_j} \simeq \frac{1}{2}\sum_{i\neq j,k} \frac{\bar{p}_{ji,k}y_{ji,k}}{I_j} \cdot \frac{(\Delta y_{ji,k})^T}{y_{ji,k}} \cdot \epsilon_{ji,k}^{p*}, \tag{43}$$

where $(\Delta Z_{j,k})^I$ *denotes the input change associated with imposing* s_j^**, conditional on having already imposed optimal trade taxes, and* $(\Delta y_{ji,k})^T$ *denotes the export change associated with imposing* t_j^**, conditional on having already imposed optimal employment subsidies.*

The formal proof can be found in Appendix A.4.3. The key distinction between Propositions 2 and 4 comes from the "shares", $w_{j,k}Z_{j,k}/I_j$ and $\bar{p}_{ji,k}y_{ji,k}/I_j$, appearing in equations (41)-(43). The numerators, $w_{j,k}Z_{j,k}$ and $\bar{p}_{ji,k}y_{ji,k}/I_j$, are gross flows, since gross output is what is being subsidized by industrial policy and taxed by trade policy. In contrast, the denominator, I_j, still measures total value added in country j. Thus "shares" now add up to numbers that are strictly greater than one, which mechanically raises the gains associated with both industrial and trade policy.

Quantitatively, the mechanical adjustment from value added to gross flows plays an important role in generating these larger gains. Starting from the second-order approximations in Proposition 2 and substituting the "shares" from Proposition 4, while keeping everything else constant, reduces the difference between the gains from industrial policy in the two environments by about one half.[44]

The other force behind the larger industrial policy gains in the environment with physical capital and input-output linkages are the larger "elasticities" of the quantities being taxed, $\Delta Z_{j,k}/Z_{j,k}$. In the Ricardian environment of Section 2.1, the average elasticity of sector size $|(\Delta L_{i,k})^I/L_{i,k}|$ associated with optimal industrial policy is 0.44; in the general environment of this section the average elasticity of sector size $|(\Delta Z_{j,k})^I/Z_{j,k}|$ is equal to

[44] Like in the baseline Ricardian environment, our second-order formulas provide good approximations in the general environment of Section 6.1. This can be seen from Figure C.1 in Appendix C.5.3.

0.61. Intuitively, since each sector and country sources a large fraction of its inputs from itself, input-output loops magnify the impact of sector-specific subsidies by further lowering the (pre-subsidy) average input costs faced by firms in the subsidized sector, the same way they magnify the effects of productivity shocks in Jones (2011).

Overall, though, gains from industrial policy remain significantly smaller than the potential gains that one might have expected given the magnitude of scale elasticities that we have estimated (18.4% in the simple back-of-the-envelope calculations of Section 5.2). This reflects the fact that while the wedges that we have estimated empirically are substantial, the cross-sectoral reallocations taking place under the optimal policy are fairly modest in all our quantitative exercises.

7 Concluding Remarks

A major source of skepticism about industrial policy is that governments simply do not know which sectors should be subsidized at the expense of others. In this paper, we have focused on the textbook case for industrial policy, in which the rationale for such policy arises from the existence of external economies of scale; we have shown how, in this environment, one can use commonly available trade and production data to estimate economies of scale at the sector-level; and, in turn, we have characterized the structure and consequences of optimal industrial policy.

Empirically, we find that sector-level economies of scale indeed exist and do differ substantially across manufacturing sectors. Yet, even under our optimistic assumption that governments maximize welfare and have full knowledge of the underlying economy, our baseline analysis points towards gains from optimal industrial policy that are hardly transformative, ranging from an average across countries of 0.98% of GDP in our baseline analysis to 3.11% in the general environment with physical capital and input-output linkages.

To put these numbers in perspective, recall that South Korea, a country often presented as an industrial policy success story, experienced gains in real GDP per capita of 6.82% *per year* from 1960 to 1989, as documented in Rodrik (1995). A 3.11% long-run welfare increase is nothing to spit at, but a miracle it is not.

Intuitively, for the textbook gains from industrial policy to be large, two sets of conditions should hold. First, production processes should exhibit significant external economies of scale that differ in strength across sectors—conditions that would give rise to a Harberger triangle with a large height in Figure 1. Second, high elasticities of substitution and trade openness should allow significant reallocations of resources across sectors—conditions that would give rise to a Harberger triangle with a large base in Figure 1. Across our quantitative exercises, given low elasticities of demand, the non-trivial gap between social and marginal costs that we have inferred remains too small to generate large gains from industrial policy, even for the most open economies.

There are, of course, many other market failures that industrial policy may target. Alleviating financial distortions, as in Itskhoki and Moll (2019) and Liu (2019), solving coordination failures, as in Murphy et al. (1989), Rodrik (1996) and Buera et al. (2021), and fostering sectors with positive spillovers to the rest of the economy, as in Greenwald and Stiglitz (2006), may all, in theory, generate transformative gains from industrial policy. Our results, however, offer little empirical support for the notion that these gains can arise from the textbook case based on external economies of scale, at least at the level of aggregation considered in our analysis. An interesting open question is whether allowing for more granular economies of scale, within both more narrowly defined sectors and regions, may magnify the gains from policy intervention by combining features of optimal industrial and trade policy, as emphasized in this paper, with those of optimal place-based policy, as emphasized in Fajgelbaum and Gaubert (2020) and Rossi-Hansberg et al. (2019). We hope that the combination of theory and empirics developed in this paper will prove useful in making progress on this and other related questions.

References

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

Arkolakis, Costas, Arnaud Costinot, and Andres Rodríguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, 2012, *102* (1), 94–130.

Atkin, David, "Trade, Tastes and Nutrition in India," *American Economic Review*, 2013, *103* (5).

Bagwell, Kyle and Robert W. Staiger, "Domestic Policies, National Sovereignty, and International Economic Institutions," *The Quarterly Journal of Economics*, 2001, *116* (2), 519–562.

Bartelme, Dominick, Arnaud Costinot, Dave Donaldson, and Andres Rodríguez-Clare, "The Textbook Case for Industrial Policy: Theory Meets Data," *NBER working paper 26193*, 2019.

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

Beck, Thorsten, "Financial development and international trade: is there a link?," *Journal of International Economics*, 2002, *57* (1), 107–131.

Buera, Francisco J, Hugo Hopenhayn, Yongseok Shin, and Nicholas Trachter, "Big Push in Distorted Economies," Technical Report, National Bureau of Economic Research 2021.

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.

Chipman, John S., "External Economies of Scale and Competitive Equilibrium," *The Quarterly Journal of Economics*, 1970, *84* (3), 347–385.

Combes, Pierre-Philippe and Laurent Gobillon, "The empirics of agglomeration economies," in Gilles Duranton, J. Vernon Henderson, and William C. Strange, eds., *Handbook of regional and urban economics*, Vol. 5, Elsevier, 2015, pp. 247–348.

Comin, Diego, Danial Lashkari, and Marti Mestieri, "Structural Change with Long-Run Income and Price Effects," *Econometrica*, 2021, *89* (1), 311–374.

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, 2014.

_ and Ivan Werning, "Robots, Trade, and Luddism: A Sufficient Statistics to Optimal Technology Regulation," *NBER working paper 25103*, 2018.

_ , 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," *Quarterly Journal of Economics*, 2019, *134* (2), 843–894.

Cravino, Javier and Sebastian Sotelo, "Trade Induced Structural Change and the Skill Premium," *American Economic Journal: Macroeconomics*, 2019, *11* (3), 289–326.

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, "Global Rebalancing with Gravity: Measuring the Burden of Adjustment," *IMF Staff Papers*, 2008, *55* (3), 511–540.

Dixit, Avinash K., "Tax Policy in Open Economies," in Martin Feldstein, ed., *Handbook of Public Economics*, Vol. 1, Elsevier, 1985.

Duranton, Gilles, "California Dreamin': The Feeble Case for Cluster Policies," *Review of Economic Analysis*, 2011, *3*, 3–45.

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

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

Fajgelbaum, Pablo D. and Cecile Gaubert, "Optimal Spatial Policies, Geography, and Sorting," *Quarterly Journal of Economics*, 2020, *135* (2), 959–1036.

Giri, Rahul, Kei-Mu Yi, and Hakan Yilmazkuday, "Gains from Trade: Does Sectoral Heterogeneity Matter?," *Journal of International Economics*, 2021, *129*, 103429.

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

Greenwald, Bruce and Joseph E Stiglitz, "Helping infant economies grow: Foundations of trade policies for developing countries," *American Economic Review*, 2006, *96* (2), 141–146.

Greenwald, Bruce C. and Joseph E. Stiglitz, "Externalities in Economies with Imperfect Information and Incomplete Markets," *Quarterly Journal of Economics*, 1986, *101* (2), 229–264.

Harrison, Ann E. and Andres Rodríguez-Clare, "Trade, Foreign Investment, and Industrial Policy for Developing Countries," in Dani Rodrik and Mark Rosenzweig, eds., *Handbook of Development Economics*, Vol. 5, 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.

Herrendorf, Berthold, Richard Rogerson, and Akos Valentinyi, "Two Perspectives on Preferences and Structural Transformation," *American Economic Review*, 2013, *103* (7), 2752–89.

Hines, James R., "Three Sides of Harberger Triangles," *Journal of Economic Perspectives*, 1999.

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

Itskhoki, Oleg and Benjamin Moll, "Optimal Development Policies with Financial Frictions," *Econometrica*, 2019, *87* (1), 139–173.

Jones, Charles I., "Intermediate Goods and Weak Links in the Theory of Economic Development," *American Economic Journal: Macroeconomics*, 2011, *3* (2), 1–28.

Juhasz, Reka, "Temporary Protection and Technology Adoption: Evidence from the Napoleonic Blockade," *American Economic Review*, 2018, *108* (11), 3339–3376.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi, "Governance Matters III: Governance Indicators for 1996–2002," 2003. World Bank Working Paper No. 3106.

Kline, Patrick and Enrico Moretti, "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs," *Annual Review of Economics*, 2014, *6*, 629–662.

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

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

Lane, Nathan, "Manufacturing Revolutions: Industrial Policy and Networks in South Korea," *mimeo IIES*, 2017.

_ , "The New Empirics of Industrial Policy," *Journal of Industry, Competition and Trade*, 2020, *20* (209-234).

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

Levchenko, Andrei A., "Institutional quality and international trade," *Review of Economic Studies*, 2007, *74* (3), 791–819.

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

Liu, Ernest, "Industrial Policies in Production Networks," *Quarterly Journal of Economics*, 2019, *134* (4), 1883–1948.

Manova, Kalina, "Credit constraints, equity market liberalizations and international trade," *Journal of International Economics*, 2008, *76* (1), 33–47.

Marshall, Alfred, *Principles of Economics, An Introductory Volume*, London: Mcmillan, eighth edition, 1920.

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

Murphy, Kevin M, Andrei Shleifer, and Robert W Vishny, "Industrialization and the big push," *Journal of political economy*, 1989, *97* (5), 1003–1026.

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

Nunn, Nathan, "Relationship-specificity, incomplete contracts, and the pattern of trade," *Quarterly Journal of Economics*, 2007, *122* (2), 569–600.

_ **and Daniel Trefler**, "Domestic Institutions as a Source of Comparative Advantage," in G. Gopinath, E. Helpman, and K. Rogoff, eds., *Handbook of International Economics*, Vol. 4, Elsevier, 2014.

Oberfield, Ezra and Devesh Raval, "Micro Data and Macro Technology," *Econometrica*, 2021, *89* (2), 703–732.

Pigou, A. C., *The Economics of Welfare*, Macmillan, 1920.

Redding, Stephen J. and David Weinstein, "Accounting for Trade Patterns," 2018. NBER working paper 24051.

Rodrik, Dani, "Getting interventions right: how South Korea and Taiwan grew rich," *Economic Policy*, 1995, *10* (20), 53–107.

_ , "Coordination failures and government policy: A model with applications to East Asia and Eastern Europe," *Journal of international economics*, 1996, *40* (1-2), 1–22.

Rosenthal, Stuart S. and William C. Strange, "Evidence on the Nature and Sources of Agglomeration Economies," in J.V. Henderson and J.F. Thisse, eds., *Handbook of Regional and Urban Economics*, Vol. 4, Elsevier, 2004.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Felipe Schwartzman, "Cognitive Hubs and Spatial Redistribution," *NBER Working Paper 26267*, 2019.

Sanderson, Eleanor and Frank Windmeijer, "A weak instrument F-test in linear IV models with multiple endogenous variables," *Journal of Econometrics*, 2016, *190* (2), 212–221.

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

Simonovska, Ina and Michael Waugh, "Different Trade Models, Different Trade Elasticities?," *mimeo*, 2012.

A Online Appendix: Theory

A.1 Derivation of Equation (11)

Proof. Start from the identity,

$$U_j(\tau) = V_j(p_j(\tau), I_j(\tau)), \tag{A.1}$$

with the total income of country j's representative agent equal to

$$I_j(\tau) = w_j L_j + \sum_{i,k} \pi_{ji,k} - \sum_k s_{j,k} w_j L_{j,k} + \sum_{j \neq i,k} t^m_{ij,k} \bar{p}_{ij,k} c_{ij,k} + \sum_{j \neq i} t^x_{ji,k} \bar{p}_{ji,k} y_{ji,k}.$$

Totally differentiating (A.1) implies

$$\begin{aligned} dU_j = \sum_{i,k} \frac{\partial V_j}{\partial p_{ij,k}} \times dp_{ij,k} + \frac{\partial V_j}{\partial I_j} \times \Bigg\{ L_j dw_j + \sum_{i,k} \left[\frac{\partial \pi_{ji,k}}{\partial q_{ji,k}} dq_{ji,k} + \frac{\partial \pi_{ji,k}}{\partial v_{j,k}} dv_{j,k} + \frac{\partial \pi_{ji,k}}{\partial L_{j,k}} dL_{j,k} \right] \\ + \sum_k (dv_{j,k} - dw_j) L_{i,k} + \sum_k (v_{j,k} - w_j) dL_{i,k} + \sum_{i,k} (dp_{ij,k} - d\bar{p}_{ij,k}) c_{ij,k} \\ + \sum_{i,k} (p_{ij,k} - \bar{p}_{ij,k}) dc_{ij,k} + \sum_{i,k} (d\bar{p}_{ji,k} - dq_{ji,k}) y_{ji,k} + \sum_{i,k} (\bar{p}_{ji,k} - q_{ji,k}) dy_{ji,k} \Bigg\}. \end{aligned} \tag{A.2}$$

Next, apply the Envelope Theorem to the utility maximization problem of country j's representative agent and the profit maximization problem of its firms. This implies

$$\frac{\partial V_j}{\partial p_{ij,k}} = -\frac{\partial V_j}{\partial I_j} c_{ij,k}, \text{ for all } i \text{ and } k, \tag{A.3}$$

$$\frac{\partial \pi_{ji,k}}{\partial q_{ji,k}} = y_{ji,k}, \text{ for all } i \text{ and } k, \tag{A.4}$$

$$\frac{\partial \pi_{ji,k}}{\partial v_{j,k}} = -\ell_{ji,k} \text{ for all } i \text{ and } k. \tag{A.5}$$

Substituting (A.3)-(A.5) into (A.2) gives

$$\begin{aligned} dU_j = \frac{\partial V_j}{\partial I_j} \times \Bigg\{ \sum_k (v_{j,k} - w_j) dL_{j,k} + \sum_{i,k} (p_{ij,k} - \bar{p}_{ij,k}) dc_{ij,k} + \sum_{i,k} (\bar{p}_{ji,k} - q_{ji,k}) dy_{ji,k} \\ + \sum_{i,k} \frac{\partial \pi_{ji,k}}{\partial L_{j,k}} dL_{j,k} - \sum_{i \neq j,k} c_{ij,k} d\bar{p}_{ij,k} + \sum_{i \neq j,k} y_{ji,k} d\bar{p}_{ji,k} \Bigg\}, \end{aligned} \tag{A.6}$$

where we have also used the good market clearing condition (7) for local goods, $c_{jj,k} = y_{jj,k}$, as well as country j's labor market clearing condition (8). To conclude, note that

$$\frac{\partial \pi_{ji,k}}{\partial L_{j,k}} = v_{j,k} \frac{E'_k(L_{j,k})}{E_k(L_{j,k})} \ell_{ji,k} \text{ for all } i \text{ and } k, \tag{A.7}$$

$$d\bar{p}_{ij,k} = 0 \text{ for all } i \neq j \text{ and } k, \tag{A.8}$$

$$d\bar{p}_{ji,k} = \tilde{p}'_{ji,k}(y_{ji,k}) dy_{ji,k} \text{ for all } i \neq j \text{ and } k. \tag{A.9}$$

Combining (2)-(4) with (A.6)-(A.9) and letting $\lambda_j \equiv \partial V_j / \partial I_j$, $\epsilon^E_{j,k} \equiv d \ln E_k(L_{j,k}) / d \ln L_{j,k}$, and $\epsilon^p_{ji,k} \equiv d \ln \tilde{p}_{ji,k}(y_{ji,k}) / d \ln y_{ji,k}$, (A.6) simplifies into (11). □

A.2 Proof of Proposition 1

Proof. Start from equation (12),

$$\sum_k s^*_{j,k} w^*_j dL_{j,k} - \sum_{i \neq j,k} t^{m*}_{ij,k} \bar{p}^*_{ij,k} dc_{ij,k} - \sum_{i \neq j,k} t^{x*}_{ji,k} \bar{p}_{ji,k} dy_{ji,k}$$
$$= \sum_k \epsilon^{E*}_{j,k} w^*_j (1 - s^*_{j,k}) dL_{j,k} + \sum_{i \neq j,k} \epsilon^{p*}_{ji,k} \bar{p}^*_{ji,k} dy_{ji,k}.$$

Since the previous condition must hold for any feasible variation, it must hold, in particular, for variations that: (i) decrease domestic sales in sector 1 and increase domestic sales in sector k so that $dL_{j,k} = -dL_{j,1}$, holding all other equilibrium variables fixed; (ii) increase the imports of good 1 from country 1 and increase the imports of good k from country i, $dc_{ij,k} = -\frac{\bar{p}^*_{1j,1}}{\bar{p}^*_{ij,k}} dc_{1j,1}$, holding all other equilibrium variables fixed; or (iii) decrease the imports of good 1 from country 1, increase the exports of good k to country i, $dy_{ji,k} = \frac{\bar{p}^*_{1j,1}}{\bar{p}^*_{ji,k}\left(1+\epsilon^{p*}_{ji,k}\right)} dc_{1j,1}$, and decrease its domestic sales of good k, $dy_{jj,k} = -\frac{A^*_{jj,k}}{A^*_{ji,k}} dy_{ji,k}$, holding all other equilibrium variables fixed. Specializing (12) to variations (i)-(iii), respectively, imply

$$s^*_{j,k} = [s^*_{j,1} - \epsilon^{E*}_{j,1}(1 - s^*_{j,1}) + \epsilon^{E*}_{j,k}] / (1 + \epsilon^{E*}_{j,k}),$$
$$t^{m*}_{ij,k} = t^{m*}_{1j,1},$$
$$t^{x*}_{ji,k} = 1 - (1 + t^{m*}_{1j,1})(1 + \epsilon^{p*}_{ji,k}).$$

Setting $\bar{s}_j = s^*_{j,1} - \epsilon^{E*}_{j,1}(1 - s^*_{j,1})$ and $\bar{t}_j = t^{m*}_{1j,1}$ concludes the proof of Proposition 1. □

A.3 Proof of Proposition 2

Proof. We first approximate the gains from fully optimal policies, W^*_j, and then show how a similar argument can be used to approximate the gains from industrial and trade policy, W^I_j and W^T_j.

Part I: Derivation of Equation (16).

Proof. Consider a set of world economies, each associated with distinct external economies of scale, $\{E_k(\cdot)\}$, and inverse import demand functions, $\{\tilde{p}_{ji,k}(\cdot)\}_{i\neq j,k}$, that generate distinct optimal policies $\tau_j^* \equiv \{s_{j,k}^*, t_{ij,k}^{m*}, t_{ji,k}^{x*}\}_{i\neq j,k}$ in country j as well as distinct elasticities $\epsilon_j^* \equiv \{\epsilon_{j,k}^{E*}, \epsilon_{ji,k}^{p*}\}_{i\neq j,k}$ in the equilibrium associated with these optimal policies. Let $\hat{W}_j^*(\epsilon_j^*)$ denote country j's welfare gains from optimal policy for each of these world economies. It varies with ϵ_j^* for two reasons: (i) through the direct effect of the underlying changes in $\{E_k(\cdot)\}$ and $\{\tilde{p}_{ji,k}(\cdot)\}_{i\neq j,k}$ on equilibrium prices and quantities, and (ii) the indirect effect of these changes on the value of the optimal taxes, $\tau_j^* \equiv \{\hat{s}_{j,k}(\epsilon_j^*), \hat{t}_{ij,k}^m(\epsilon_j^*), \hat{t}_{ji,k}^x(\epsilon_j^*)\} \equiv \hat{\tau}_j(\epsilon_j^*)$, with $\hat{s}_{j,k}(\epsilon_j^*) \equiv \epsilon_{j,k}^{E*}/(1+\epsilon_{j,k}^{E*})$, $\hat{t}_{ij,k}^m(\epsilon_j^*) \equiv 0$, and $\hat{t}_{ji,k}^x(\epsilon_j^*) \equiv -\epsilon_{ji,k}^{p*}$, as shown in Proposition 1 under the normalization $\bar{s}_j = \bar{t}_j = 0$. To keep track of these direct and indirect effects, rewrite equation (13) as

$$V_j(\hat{p}_j(\hat{\tau}_j(\epsilon_j^*), \epsilon_j^*), \hat{I}_j(\hat{\tau}_j(\epsilon_j^*), \epsilon_j^*)) = V_j(\hat{p}_j(0, \epsilon_j^*), \hat{I}_j(0, \epsilon_j^*) + \hat{W}_j^*(\epsilon_j^*)), \tag{A.10}$$

with the second argument of $\hat{p}_j(\hat{\tau}_j(\epsilon_j^*), \epsilon_j^*)$ and $\hat{I}_j(\hat{\tau}_j(\epsilon_j^*), \epsilon_j^*)$ capturing the direct effect on prices and income of $\{E_k(\cdot)\}$ and $\{\tilde{p}_{ji,k}(\cdot)\}_{i\neq j,k}$. The goal is to show that as $\epsilon_j^* \to 0$, the difference between $\hat{W}_j^*(\epsilon_j^*)$ and $\frac{1}{2}\sum_k w_j \Delta L_{j,k}\epsilon_{j,k}^{E*} + \frac{1}{2}\sum_{i\neq j,k} \bar{p}_{ji,k}\Delta y_{ji,k}\epsilon_{ji,kn}^{p*}$ goes to zero at a rate no slower than $|\epsilon_j^*|^3$,

$$\hat{W}_j^*(\epsilon_j^*) = \frac{1}{2}\sum_k w_j(\Delta L_{j,k})^*\epsilon_{j,k}^{E*} + \frac{1}{2}\sum_{i\neq j,k} \bar{p}_{ji,k}(\Delta y_{ji,k})^*\epsilon_{ji,k}^{p*} + O(|\epsilon_j^*|^3), \tag{A.11}$$

which is the formal counterpart of the approximation displayed in equation (16).

To establish this result, start from the second-order Taylor expansion,

$$\hat{W}_j^*(\epsilon_j^*) = \hat{W}_j^*(0) + [D_\epsilon \hat{W}_j^*]'_{\epsilon_j^*=0}\epsilon_j^* + \frac{1}{2}(\epsilon_j^*)'[H_\epsilon \hat{W}_j^*]_{\epsilon_j^*=0}\epsilon_j^* + O(|\epsilon_j^*|^3), \tag{A.12}$$

with $D_\epsilon \hat{W}_j^* \equiv \{d\hat{W}_j^*/d\epsilon_{j,n}^*\}$ the vector of first derivatives with respect to all elasticities, $\epsilon_j^* \equiv \{\epsilon_{j,n}^*\}$, with the convention $\epsilon_{j,n}^* = \epsilon_{j,k}^{E*}$ if $n = (k, E)$ and $\epsilon_{j,n}^* = \epsilon_{ji,k}^{p*}$ if $n = (i, k, p)$, and $H_\epsilon \hat{W}_j^* \equiv \{d^2\hat{W}_j^*/(d\epsilon_{j,n}^* d\epsilon_{j,m}^*)\}$ the associated Hessian. First, note that gains from optimal policy are zero if $\epsilon_j^* = 0$,

$$\hat{W}_j^*(0) = 0. \tag{A.13}$$

Second, differentiate (A.10) with respect $\epsilon_{j,n}^*$,

$$\frac{\partial \hat{V}_j(\hat{\tau}_j(\epsilon_j^*), \epsilon_j^*, 0)}{\partial \tau_{j,n}}\frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*} + \frac{\partial \hat{V}_j(\hat{\tau}_j(\epsilon_j^*), \epsilon_j^*, 0)}{\partial \epsilon_{j,n}^*} = \frac{\partial \hat{V}_j(0, \epsilon_j^*, \hat{W}_j^*(\epsilon_j^*))}{\partial \epsilon_{j,n}^*} + \frac{\partial \hat{V}_j(0, \epsilon_j^*, \hat{W}_j^*(\epsilon_j^*))}{\partial W_j}\frac{d\hat{W}_j^*}{d\epsilon_{j,n}^*}, \tag{A.14}$$

with $\hat{V}_j(\tau_j, \epsilon_j^*, W_j) \equiv V_j(\hat{p}_j(\tau_j, \epsilon_j^*), \hat{I}_j(\tau_j, \epsilon_j^*) + W_j)$ and $\hat{\tau}_{j,n}$ is the policy associated with the elasticity $\epsilon_{j,n}^*$, i.e., $\hat{\tau}_{j,n}(\epsilon_j^*) = \hat{s}_{j,k}(\epsilon_j^*)$ if $n = (k, E)$ and $\hat{\tau}_{j,n}(\epsilon_j^*) = \hat{t}_{ij,k}^x(\epsilon_j^*)$ if $n = (i, k, p)$. At the optimal policy mix, $\tau_j^* = \hat{\tau}_j(\epsilon_j^*)$, we know that $\partial \hat{V}_j(\hat{\tau}_j(\epsilon_j^*), \epsilon_j^*, 0)/\partial \tau_{j,n} = 0$. Noting that $(\partial \hat{V}_j/\partial W_j)_{(\tau_j, \epsilon_j^*, W_j)=0} =$

$\lambda_j \neq 0$, where λ_j denotes the marginal utility of income in the laissez-faire equilibrium with $\epsilon_j^* = 0$, equations (A.13) and (A.14) then imply

$$[D_\epsilon \hat{W}_j^*]_{\epsilon_j^*=0} = 0. \tag{A.15}$$

Third, differentiate (A.14) with respect to $\epsilon_{j,m}^*$,

$$\begin{aligned}
&\frac{\partial \hat{V}_j^2(\hat{\tau}_j(\epsilon_j^*), \epsilon_j^*, 0)}{\partial \tau_{j,m}\partial \tau_{j,n}} \frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*} \frac{d\hat{\tau}_{j,m}}{d\epsilon_{j,m}^*} + \frac{\partial \hat{V}_j^2(\hat{\tau}_j(\epsilon_j^*), \epsilon_j^*, 0)}{\partial \epsilon_{j,m}^*\partial \tau_{j,n}} \frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*} + \frac{\partial^2 \hat{V}_j(\hat{\tau}_j(\epsilon_j^*), \epsilon_j^*, 0)}{\partial \tau_{j,m}\partial \epsilon_{j,n}^*} \frac{d\hat{\tau}_{j,m}}{d\epsilon_{j,m}^*} + \frac{\partial^2 \hat{V}_j(\hat{\tau}_j(\epsilon_j^*), \epsilon_j^*, 0)}{\partial \epsilon_{j,m}^*\partial \epsilon_{j,n}^*} \\
&\quad = \frac{\partial^2 \hat{V}_j(0, \epsilon_j^*, \hat{W}_j^*(\epsilon_j^*))}{\partial \epsilon_{j,m}^*\partial \epsilon_{j,n}^*} + \frac{\partial^2 \hat{V}_j(0, \epsilon_j^*, \hat{W}_j^*(\epsilon_j^*))}{\partial W_j \partial \epsilon_{j,n}^*} \frac{d\hat{W}_j^*}{d\epsilon_{j,m}^*} + \frac{\partial^2 \hat{V}_j(0, \epsilon_j^*, \hat{W}_j^*(\epsilon_j^*))}{\partial \epsilon_{j,m}^* \partial W_j} \frac{d\hat{W}_j^*}{d\epsilon_{j,n}^*} \\
&\qquad + \frac{\partial^2 \hat{V}_j(0, \epsilon_j^*, \hat{W}_j^*(\epsilon_j^*))}{\partial W_j^2} \frac{d\hat{W}_j^*}{d\epsilon_{j,m}^*} \frac{d\hat{W}_j^*}{d\epsilon_{j,n}^*} + \frac{\partial \hat{V}_j(0, \epsilon_j^*, \hat{W}_j^*(\epsilon_j^*))}{\partial W_j} \frac{d^2\hat{W}_j^*}{d\epsilon_{j,m}^* d\epsilon_{j,n}^*}.
\end{aligned}$$

Evaluated at $\epsilon_j^* = 0$, the previous expression implies

$$\frac{d^2\hat{W}_j^*(0)}{d\epsilon_{j,m}^* d\epsilon_{j,n}^*} = \frac{1}{\lambda_j} \left[\frac{\partial \hat{V}_j^2}{\partial \tau_{j,m}\partial \tau_{j,n}} \frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*} \frac{d\hat{\tau}_{j,m}}{d\epsilon_{j,m}^*} + \frac{\partial \hat{V}_j^2}{\partial \epsilon_{j,m}^* \partial \tau_{j,n}} \frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*} + \frac{\partial^2 \hat{V}_j}{\partial \tau_{j,m}\partial \epsilon_{j,n}^*} \frac{d\hat{\tau}_{j,m}}{d\epsilon_{j,m}^*} \right]_{(\tau_j, \epsilon_j^*, W_j)=0}, \tag{A.16}$$

where we have used (A.15) and $(\partial \hat{V}_j / \partial W_j)_{(\tau_j, \epsilon_j^*, W_j)=0} = \lambda_j$. Next, compute the cross-derivatives of $\hat{V}_j$. We start from equation (11). Using the present notation and the fact that optimal tariffs are zero, it can be rearranged as

$$\frac{1}{\lambda_j^*} \frac{\partial \hat{V}_j(\hat{\tau}_j(\epsilon_j^*), \epsilon_j^*, 0)}{\partial \tau_{j,n}} = -\sum_k s_{j,k}^* \hat{w}_j \frac{\partial \hat{L}_{j,k}}{\partial \tau_{j,n}} + \sum_{i\neq j,k} t_{ji,k}^{x*} \hat{\tilde{p}}_{ji,k} \frac{\partial \hat{y}_{ji,k}}{\partial \tau_{j,n}} + \sum_k \epsilon_{j,k}^{E*}(1 - s_{j,k}^*) \hat{w}_j \frac{\partial \hat{L}_{j,k}}{\partial \tau_{j,n}} + \sum_{i\neq j,k} \epsilon_{ji,k}^{p*} \hat{\tilde{p}}_{ji,k} \frac{\partial \hat{y}_{ji,k}}{\partial \tau_{j,n}},$$

where λ_j^* is the marginal utility of income under the optimal policy mix and all equilibrium prices and quantities appearing on the right-hand side with hats are treated as functions of $(\hat{\tau}_j(\epsilon_j^*), \epsilon_j^*)$. Differentiating a second time with respect to $\omega_{j,m} \in \{\tau_{j,m}, \epsilon_{j,m}\}$ and using $\tau_j^* = \hat{\tau}_j(\epsilon_j^*)$, we obtain

$$\begin{aligned}
\frac{\partial \left(1/\lambda_j^*\right)}{\partial \omega_{j,m}} \frac{\partial \hat{V}_j}{\partial \tau_{j,n}} + \frac{1}{\lambda_j^*} \frac{\partial^2 \hat{V}_j}{\partial \omega_{j,m}\partial \tau_{j,n}} = \sum_k \left(-\frac{\partial s_{j,k}^*}{\partial \omega_{j,m}} + \frac{\partial [\epsilon_{j,k}^{E*}(1 - s_{j,k}^*)]}{\partial \omega_{j,m}} \right) \left(\hat{w}_j \frac{\partial \hat{L}_{j,k}}{\partial \tau_{j,n}} \right) \\
+ \sum_{i\neq j,k} \left(\frac{\partial t_{ji,k}^{x*}}{\partial \omega_{j,m}} - \frac{\partial(-\epsilon_{ji,k}^{p*})}{\partial \omega_{j,m}} \right) \left(\hat{\tilde{p}}_{ji,k} \frac{\partial \hat{y}_{ji,k}}{\partial \tau_{j,n}} \right).
\end{aligned}$$

Since $\partial \hat{V}_j(\hat{\tau}_j(\epsilon_j^*), \epsilon_j^*, 0)/\partial \tau_{j,n} = 0$ at the optimal policy mix, this yields

$$\frac{1}{\lambda_j}\left(\frac{\partial^2 \hat{V}_j}{\partial s_{j,k}\partial \tau_{j,n}}\right)_{(\tau_j,\epsilon_j^*,W_j)=0} = \left(-\hat{w}_j(1+\epsilon_{j,k}^{E*})\frac{\partial \hat{L}_{j,k}}{\partial \tau_{j,n}}\right)_{(\tau_j,\epsilon_j^*)=0}, \tag{A.17}$$

$$\frac{1}{\lambda_j}\left(\frac{\partial^2 \hat{V}_j}{\partial \epsilon_{j,k}^{E*}\partial \tau_{j,n}}\right)_{(\tau_j,\epsilon_j^*,W_j)=0} = \left(\hat{w}_j(1-s_{j,k}^*)\frac{\partial \hat{L}_{j,k}}{\partial \tau_{j,n}}\right)_{(\tau_j,\epsilon_j^*)=0}, \tag{A.18}$$

$$\frac{1}{\lambda_j}\left(\frac{\partial^2 \hat{V}_j}{\partial t_{ji,k}^{x}\partial \tau_{j,n}}\right)_{(\tau_j,\epsilon_j^*,W_j)=0} = \left(\hat{\bar{p}}_{ji,k}\frac{\partial \hat{y}_{ji,k}}{\partial \tau_{j,n}}\right)_{(\tau_j,\epsilon_j^*)=0}, \tag{A.19}$$

$$\frac{1}{\lambda_j}\left(\frac{\partial^2 \hat{V}_j}{\partial \epsilon_{ji,k}^{p*}\partial \tau_{j,n}}\right)_{(\tau_j,\epsilon_j^*,W_j)=0} = \left(\hat{\bar{p}}_{ji,k}\frac{\partial \hat{y}_{ji,k}}{\partial \tau_{j,n}}\right)_{(\tau_j,\epsilon_j^*)=0}. \tag{A.20}$$

By definition, $\hat{\tau}_j(\epsilon_j^*) \equiv \{\hat{s}_{j,k}(\epsilon_j^*), \hat{t}_{ij,k}^{m}(\epsilon_j^*), \hat{t}_{ji,k}^{x}(\epsilon_j^*)\}$ satisfies

$$\frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*} = \frac{d\hat{s}_{j,k}(\epsilon_j^*)}{d\epsilon_{j,k}^{E*}} = \frac{1}{(1+\epsilon_{j,k}^{E*})^2}, \text{ if } n = (k, E), \tag{A.21}$$

$$\frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*} = \frac{d\hat{t}_{ji,k}^{x}(\epsilon_j^*)}{d\epsilon_{ji,k}^{p*}} = -1, \text{ if } n = (i, k, p). \tag{A.22}$$

Equations (A.12)-(A.22) imply

$$\hat{W}_j^*(\epsilon_j^*) = \frac{1}{2}\sum_{k,n}\left(\hat{w}_j(1-\hat{s}_{j,k})\frac{\partial \hat{L}_{j,k}}{\partial \tau_{j,n}}\frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*}\right)_{(\tau_j,\epsilon_j^*)=0}\epsilon_{j,n}^*\epsilon_{j,k}^{E*}$$
$$+\frac{1}{2}\sum_{i\neq j,k,n}\left(\hat{\bar{p}}_{ji,k}\frac{\partial \hat{y}_{ji,k}}{\partial \tau_{j,n}}\frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*}\right)_{(\tau_j,\epsilon_j^*)=0}\epsilon_{j,n}^*\epsilon_{ji,k}^{p*} + O(|\epsilon_j^*|^3). \tag{A.23}$$

To simplify the previous expression, note that a zero-order Taylor approximation implies

$$w_j \equiv \hat{w}_j(0,\epsilon_j^*) = \hat{w}_j(0,0) + O(|\epsilon_j^*|), \tag{A.24}$$

$$\bar{p}_{ji,k} \equiv \hat{\bar{p}}_{ji,k}(0,\epsilon_j^*) = \hat{\bar{p}}_{ji,k}(0,0) + O(|\epsilon_j^*|). \tag{A.25}$$

Note also that

$$(\Delta L_{j,k})^* \equiv \hat{L}_{j,k}(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*) - \hat{L}_{j,k}(0,\epsilon_j^*) = \hat{L}_{j,k}(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*) - \hat{L}_{j,k}(0,0) + \hat{L}_{j,k}(0,0) - \hat{L}_{j,k}(0,\epsilon_j^*), \tag{A.26}$$

$$(\Delta y_{ji,k})^* \equiv \hat{y}_{ji,k}(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*) - \hat{y}_{ji,k}(0,\epsilon_j^*) = \hat{y}_{ji,k}(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*) - \hat{y}_{ji,k}(0,0) + \hat{y}_{ji,k}(0,0) - \hat{y}_{ji,k}(0,\epsilon_j^*). \tag{A.27}$$

First-order Taylor approximations therefore imply

$$\hat{L}_{j,k}(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*) - \hat{L}_{j,k}(0,0) = \sum_n \left(\frac{\partial \hat{L}_{j,k}}{\partial \tau_{j,n}} \frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*} + \frac{\partial \hat{L}_{j,k}}{\partial \epsilon_{j,n}^*} \right)_{(\tau_j,\epsilon_j^*)=0} \epsilon_{j,n}^* + O(|\epsilon_j^*|^2), \tag{A.28}$$

$$\hat{L}_{j,k}(0,\epsilon_j^*) - \hat{L}_{j,k}(0,0) = \sum_n \left(\frac{\partial \hat{L}_{j,k}}{\partial \epsilon_{j,n}^*} \right)_{(\tau_j,\epsilon_j^*)=0} \epsilon_{j,n}^* + O(|\epsilon_j^*|^2), \tag{A.29}$$

$$\hat{y}_{ji,k}(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*) - \hat{y}_{ji,k}(0,0) = \sum_n \left(\frac{\partial \hat{y}_{ji,k}}{\partial \tau_{j,n}} \frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*} + \frac{\partial \hat{y}_{ji,k}}{\partial \epsilon_{j,n}^*} \right)_{(\tau_j,\epsilon_j^*)=0} \epsilon_{j,n}^* + O(|\epsilon_j^*|^2), \tag{A.30}$$

$$\hat{y}_{ji,k}(0,\epsilon_j^*) - \hat{y}_{ji,k}(0,0) = \sum_n \left(\frac{\partial \hat{y}_{ji,k}}{\partial \epsilon_{j,n}^*} \right)_{(\tau_j,\epsilon_j^*)=0} \epsilon_{j,n}^* + O(|\epsilon_j^*|^2). \tag{A.31}$$

Combining equations (A.23)-(A.31) as well as noting that $\hat{s}_{j,k}, (\Delta L_{j,k})^*, (\Delta y_{ji,k})^*, \epsilon_{j,k}^{E*}, \epsilon_{ji,k}^{p*} = O(|\epsilon_j^*|)$ and $aO(|\epsilon_j^*|^b)O(|\epsilon_j^*|^c) = O(|\epsilon_j^*|^{b+c})$ for any a, b, and c, we obtain (A.11). □

Part II: Derivation of Equation (17).

Proof. Like in Part I, to keep track of the direct and indirect effects associated with changes in ϵ_j^*, rewrite equation (14) as

$$V_j(\hat{p}_j(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*), \hat{I}_j(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*)) = V_j(\hat{p}_j((0,\hat{t}_j(\epsilon_j^*)),\epsilon_j^*), \hat{I}_j((0,\hat{t}_j(\epsilon_j^*)),\epsilon_j^*) + \hat{W}_j^I(\epsilon_j^*)), \tag{A.32}$$

with $\hat{t}_j(\epsilon_j^*) \equiv \{\hat{t}_{ij,k}^m(\epsilon_j^*), \hat{t}_{ji,k}^x(\epsilon_j^*)\}$. The goal is now to show that as $\epsilon_j^* \to 0$, the difference between $\hat{W}_j^I(\epsilon_j^*)$ and $\frac{1}{2}\sum_k w_j(\Delta L_{j,k})^I \epsilon_{j,k}^{E*}$ goes to zero at a rate no slower than $|\epsilon_j^*|^3$,

$$\hat{W}_j^I(\epsilon_j^*) = \frac{1}{2}\sum_k w_j(\Delta L_{j,k})^I \epsilon_{j,k}^{E*} + O(|\epsilon_j^*|^3), \tag{A.33}$$

which is the formal counterpart of the approximation displayed in equation (17).

The first part of the proof is unchanged with

$$\hat{W}_j^I(\epsilon_j^*) = \hat{W}_j^I(0) + [D_\epsilon \hat{W}_j^I]'_{\epsilon_j^*=0}\epsilon_j^* + \frac{1}{2}(\epsilon_j^*)'[H_\epsilon \hat{W}_j^I]_{\epsilon_j^*=0}\epsilon_j^* + O(|\epsilon_j^*|^3), \tag{A.34}$$

$$\hat{W}_j^I(0) = 0. \tag{A.35}$$

Compared to Part I, differentiating (A.32) with respect $\epsilon^*_{j,n}$ now implies

$$\frac{\partial \hat{V}_j(\hat{\tau}_j(\epsilon^*_j),\epsilon^*_j,0)}{\partial \tau_{j,n}}\frac{d\hat{\tau}_{j,n}}{d\epsilon^*_{j,n}} + \frac{\partial \hat{V}_j(\hat{\tau}_j(\epsilon^*_j),\epsilon^*_j,0)}{\partial \epsilon^*_{j,n}}$$
$$= \frac{\partial \hat{V}_j((0,\hat{t}_j(\epsilon^*_j)),\epsilon^*_j,\hat{W}^I_j(\epsilon^*_j))}{\partial \epsilon^*_{j,n}} + \frac{\partial \hat{V}_j((0,\hat{t}_j(\epsilon^*_j)),\epsilon^*_j,\hat{W}^I_j(\epsilon^*_j))}{\partial W_j}\frac{d\hat{W}^I_j}{d\epsilon^*_{j,n}}$$
$$+ \sum_m \frac{\partial \hat{V}_j((0,\hat{t}_j(\epsilon^*_j)),\epsilon^*_j,\hat{W}^I_j(\epsilon^*_j))}{\partial t_{j,m}}\frac{d\hat{t}_{j,m}}{d\epsilon^*_{j,n}}, \tag{A.36}$$

which still yields

$$[D_\epsilon \hat{W}^I_j]_{\epsilon^*_j=0} = 0. \tag{A.37}$$

Compared to Part I, differentiating (A.36) and evaluating it at $\epsilon^*_j = 0$ then implies

$$\frac{d^2\hat{W}^I_j(0)}{d\epsilon^*_{j,m}d\epsilon^*_{j,n}} = \begin{cases} \frac{1}{\lambda_j}\left[\frac{\partial \hat{V}^2_j}{\partial \tau_{j,m}\partial \tau_{j,n}}\frac{d\hat{\tau}_{j,n}}{d\epsilon^*_{j,n}}\frac{d\hat{\tau}_{j,m}}{d\epsilon^*_{j,m}} + \frac{\partial \hat{V}^2_j}{\partial \epsilon^*_{j,m}\partial \tau_{j,n}}\frac{d\hat{\tau}_{j,n}}{d\epsilon^*_{j,n}} + \frac{\partial^2 \hat{V}_j}{\partial \tau_{j,m}\partial \epsilon^*_{j,n}}\frac{d\hat{\tau}_{j,m}}{d\epsilon^*_{j,m}}\right]_{(\tau_j,\epsilon^*_j,W_j)=0} & \text{if } \epsilon^*_{j,n} = \epsilon^{E*}_{j,k}, \epsilon^*_{j,m} = \epsilon^{E*}_{j,l}, \\ \frac{1}{\lambda_j}\left[\frac{\partial \hat{V}^2_j}{\partial \tau_{j,m}\partial \tau_{j,n}}\frac{d\hat{\tau}_{j,n}}{d\epsilon^*_{j,n}}\frac{d\hat{\tau}_{j,m}}{d\epsilon^*_{j,m}} + \frac{\partial \hat{V}^2_j}{\partial \epsilon^*_{j,m}\partial \tau_{j,n}}\frac{d\hat{\tau}_{j,n}}{d\epsilon^*_{j,n}}\right] & \text{if } \epsilon^*_{j,n} = \epsilon^{E*}_{j,k}, \epsilon^*_{j,m} = \epsilon^{p*}_{ji,l}, \\ \frac{1}{\lambda_j}\left[\frac{\partial \hat{V}^2_j}{\partial \tau_{j,m}\partial \tau_{j,n}}\frac{d\hat{\tau}_{j,n}}{d\epsilon^*_{j,n}}\frac{d\hat{\tau}_{j,m}}{d\epsilon^*_{j,m}} + \frac{\partial^2 \hat{V}_j}{\partial \tau_{j,m}\partial \epsilon^*_{j,n}}\frac{d\hat{\tau}_{j,m}}{d\epsilon^*_{j,m}}\right]_{(\tau_j,\epsilon^*_j,W_j)=0} & \text{if } \epsilon^*_{j,n} = \epsilon^{p*}_{ji,k}, \epsilon^*_{j,m} = \epsilon^{E*}_{j,l}, \\ 0 & \text{if } \epsilon^*_{j,n} = \epsilon^{p*}_{ji,k}, \epsilon^*_{j,m} = \epsilon^{p*}_{jd,l}. \end{cases}$$

Since the cross-derivatives of $\hat{V}_j$ and the derivatives of $\hat{\tau}_j$ displayed in equations (A.17)-(A.22) are unchanged, the previous expression yields

$$\frac{d^2\hat{W}^I_j(0)}{d\epsilon^*_{j,m}d\epsilon^*_{j,n}} = \begin{cases} \left(\hat{w}_j(1-s^*_{j,k})\frac{\partial \hat{L}_{j,k}}{\partial \tau_{j,m}}\frac{d\hat{\tau}_{j,m}}{d\epsilon^*_{j,m}}\right)_{(\tau_j,\epsilon^*_j)=0} & \text{if } \epsilon^*_{j,n} = \epsilon^{E*}_{j,k}, \epsilon^*_{j,m} = \epsilon^{E*}_{j,l}, \\ 0 & \text{otherwise.} \end{cases} \tag{A.38}$$

Combining equations (A.34), (A.35), (A.37), and (A.38), we obtain

$$\hat{W}^I_j(\epsilon^*_j) = \frac{1}{2}\sum_{k,l}\left(\hat{w}_j(1-\hat{s}_{j,k})\frac{\partial \hat{L}_{j,k}}{\partial s_{j,l}}\frac{d\hat{s}_{j,l}}{d\epsilon^{E*}_{j,l}}\right)_{(\tau_j,\epsilon^*_j)=0}\epsilon^{E*}_{j,l}\epsilon^{E*}_{j,k} + O(|\epsilon^*_j|^3).$$

The final part of the argument is the same as in Part I, now applied to

$$(\Delta L_{j,k})^I \equiv \hat{L}_{j,k}(\hat{\tau}_j(\epsilon^*_j),\epsilon^*_j) - \hat{L}_{j,k}((0,\hat{t}_j(\epsilon^*_j)),\epsilon^*_j),$$

using the first-order Taylor approximations,

$$\hat{L}_{j,k}(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*) - \hat{L}_{j,k}(0,0) = \sum_n \left(\frac{\partial \hat{L}_{j,k}}{\partial \tau_{j,n}} \frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*} + \frac{\partial \hat{L}_{j,k}}{\partial \epsilon_{j,n}^*} \right)_{(\tau_j,\epsilon_j^*)=0} \epsilon_{j,n}^* + O(|\epsilon_j^*|^2),$$

$$\hat{L}_{j,k}((0,\hat{t}_j(\epsilon_j^*)),\epsilon_j^*) - \hat{L}_{j,k}(0,0) = \sum_n \left(\sum_m \frac{\partial \hat{L}_{j,k}}{\partial t_{j,m}} \frac{d\hat{t}_{j,m}}{d\epsilon_{j,n}^*} + \frac{\partial \hat{L}_{j,k}}{\partial \epsilon_{j,n}^*} \right)_{(\tau_j,\epsilon_j^*)=0} \epsilon_{j,n}^* + O(|\epsilon_j^*|^2).$$

□

Part III: Derivation of Equation (15).

Proof. Like in Parts I and II, rewrite equation (15) as

$$V_j(\hat{p}_j(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*), \hat{I}_j(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*)) = V_j(\hat{p}_j((\hat{s}_j(\epsilon_j^*),0),\epsilon_j^*), \hat{I}_j((\hat{s}_j(\epsilon_j^*),0),\epsilon_j^*) + \hat{W}_j^T(\epsilon_j^*)), \tag{A.39}$$

with $\hat{s}_j(\epsilon_j^*) \equiv \{\hat{s}_{j,k}(\epsilon_j^*)\}$. The goal is now to show that as $\epsilon_j^* \to 0$, the difference between $\hat{W}_j^T(\epsilon_j^*)$ and $\frac{1}{2}\sum_{i\neq j,k} \bar{p}_{ji,k}(\Delta y_{ji,k})^T \epsilon_{ji,k}^{p*}$ goes to zero at a rate no slower than $|\epsilon_j^*|^3$,

$$\hat{W}_j^T(\epsilon_j^*) = \frac{1}{2}\sum_{i\neq j,k} \bar{p}_{ji,k}(\Delta y_{ji,k})^T \epsilon_{ji,k}^{p*} + O(|\epsilon_j^*|^3), \tag{A.40}$$

which is the formal counterpart of the approximation displayed in equation (18).

The first part of the proof is again unchanged with

$$\hat{W}_j^T(\epsilon_j^*) = \hat{W}_j^T(0) + [D_\epsilon \hat{W}_j^T]'_{\epsilon_j^*=0}\epsilon_j^* + \frac{1}{2}(\epsilon_j^*)'[H_\epsilon \hat{W}_j^T]_{\epsilon_j^*=0}\epsilon_j^* + O(|\epsilon_j^*|^3), \tag{A.41}$$

$$\hat{W}_j^T(0) = 0. \tag{A.42}$$

Compared to Parts I and II, differentiating (A.39) with respect $\epsilon_{j,n}^*$ implies

$$\begin{aligned}
&\frac{\partial \hat{V}_j(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*,0)}{\partial \tau_{j,n}} \frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*} + \frac{\partial \hat{V}_j(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*,0)}{\partial \epsilon_{j,n}^*} \\
&\qquad = \frac{\partial \hat{V}_j((\hat{s}_j(\epsilon_j^*),0),\epsilon_j^*,\hat{W}_j^T(\epsilon_j^*))}{\partial \epsilon_{j,n}^*} + \frac{\partial \hat{V}_j((\hat{s}_j(\epsilon_j^*),0),\epsilon_j^*,\hat{W}_j^T(\epsilon_j^*))}{\partial W_j} \frac{d\hat{W}_j^T}{d\epsilon_{j,n}^*} \\
&\qquad\qquad + \sum_m \frac{\partial \hat{V}_j((\hat{s}_j(\epsilon_j^*),0),\epsilon_j^*,\hat{W}_j^T(\epsilon_j^*))}{\partial s_{j,m}} \frac{d\hat{s}_{j,m}}{d\epsilon_{j,n}^*},
\end{aligned} \tag{A.43}$$

which still yields

$$[D_\epsilon \hat{W}_j^T]_{\epsilon_j^*=0} = 0. \tag{A.44}$$

Compared to Parts I and II, differentiating (A.43) and evaluating it at $\epsilon_j^* = 0$ then implies

$$\frac{d^2\hat{W}_j^T(0)}{d\epsilon_{j,m}^* d\epsilon_{j,n}^*} = \begin{cases} 0 & \text{if } \epsilon_{j,n}^* = \epsilon_{j,k}^{E*}, \epsilon_{j,m}^* = \epsilon_{j,l}^{E*}, \\ \frac{1}{\lambda_j}\left[\frac{\partial \hat{V}_j^2}{\partial \tau_{j,m}\partial \tau_{j,n}} \frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*}\frac{d\hat{\tau}_{j,m}}{d\epsilon_{j,m}^*} + \frac{\partial^2 \hat{V}_j}{\partial \tau_{j,m}\partial \epsilon_{j,n}^*}\frac{d\hat{\tau}_{j,m}}{d\epsilon_{j,m}^*}\right]_{(\tau_j,\epsilon_j^*,W_j)=0} & \text{if } \epsilon_{j,n}^* = \epsilon_{j,k}^{E*}, \epsilon_{j,m}^* = \epsilon_{ji,l}^{p*}, \\ \frac{1}{\lambda_j}\left[\frac{\partial \hat{V}_j^2}{\partial \tau_{j,m}\partial \tau_{j,n}} \frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*}\frac{d\hat{\tau}_{j,m}}{d\epsilon_{j,m}^*} + \frac{\partial \hat{V}_j^2}{\partial \epsilon_{j,m}^*\partial \tau_{j,n}}\frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*}\right]_{(\tau_j,\epsilon_j^*,W_j)=0} & \text{if } \epsilon_{j,n}^* = \epsilon_{ji,k}^{p*}, \epsilon_{j,m}^* = \epsilon_{j,l}^{E*}, \\ \frac{1}{\lambda_j}\left[\frac{\partial \hat{V}_j^2}{\partial \tau_{j,m}\partial \tau_{j,n}} \frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*}\frac{d\hat{\tau}_{j,m}}{d\epsilon_{j,m}^*} + \frac{\partial \hat{V}_j^2}{\partial \epsilon_{j,m}^*\partial \tau_{j,n}}\frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*} + \frac{\partial^2 \hat{V}_j}{\partial \tau_{j,m}\partial \epsilon_{j,n}^*}\frac{d\hat{\tau}_{j,m}}{d\epsilon_{j,m}^*}\right]_{(\tau_j,\epsilon_j^*,W_j)=0} & \text{if } \epsilon_{j,n}^* = \epsilon_{ji,k}^{p*}, \epsilon_{j,m}^* = \epsilon_{jd,l}^{p*}. \end{cases}$$

Using again the fact that the cross-derivatives of $\hat{V}_j$ and the derivatives of $\hat{\tau}_j$ displayed in equations (A.17)-(A.22) are unchanged, this yields

$$\frac{d^2\hat{W}_j^T(0)}{d\epsilon_{j,m}^* d\epsilon_{j,n}^*} = \begin{cases} \left(\hat{\bar{p}}_{ji,k}\frac{\partial \hat{y}_{ji,k}}{\partial \tau_{j,m}}\frac{d\hat{\tau}_{j,m}}{d\epsilon_{j,m}^*}\right)_{(\tau_j,\epsilon_j^*)=0} & \text{if } \epsilon_{j,n}^* = \epsilon_{ji,k}^{p*}, \epsilon_{j,m}^* = \epsilon_{jd,l}^{p*}, \\ 0 & \text{otherwise.} \end{cases} \tag{A.45}$$

Combining equations (A.41), (A.42), (A.44), and (A.45), we obtain

$$\hat{W}_j^T(\epsilon_j^*) = \frac{1}{2}\sum_{d\neq j, i\neq j,k,l}\left(\hat{\bar{p}}_{ji,k}\frac{\partial \hat{y}_{ji,k}}{\partial t_{jd,l}^x}\frac{d\hat{t}_{jd,l}^x}{d\epsilon_{jd,l}^{p*}}\right)_{(\tau_j,\epsilon_j^*)=0}\epsilon_{jd,l}^{p*}\epsilon_{ji,k}^{p*} + O(|\epsilon_j^*|^3).$$

The final part of the argument is the same as in Parts I and II, now applied to

$$(\Delta y_{ji,k})^T \equiv \hat{y}_{ji,k}(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*) - \hat{y}_{ji,k}((\hat{s}_j(\epsilon_j^*),0),\epsilon_j^*),$$

using the first-order Taylor approximations,

$$\hat{y}_{ji,k}(\hat{\tau}_j(\epsilon_j^*),\epsilon_j^*) - \hat{y}_{ji,k}(0,0) = \sum_n\left(\frac{\partial \hat{y}_{ji,k}}{\partial \tau_{j,n}}\frac{d\hat{\tau}_{j,n}}{d\epsilon_{j,n}^*} + \frac{\partial \hat{y}_{ji,k}}{\partial \epsilon_{j,n}^*}\right)_{(\tau_j,\epsilon_j^*)=0}\epsilon_{j,n}^* + O(|\epsilon_j^*|^2),$$

$$\hat{y}_{ji,k}(0,\epsilon_j^*) - \hat{y}_{ji,k}(0,0) = \sum_n\left(\sum_m\frac{\partial \hat{y}_{ji,k}}{\partial s_{j,m}}\frac{d\hat{s}_{j,m}}{d\epsilon_{j,n}^*} + \frac{\partial \hat{y}_{ji,k}}{\partial \epsilon_{j,n}^*}\right)_{(\tau_j,\epsilon_j^*)=0}\epsilon_{j,n}^* + O(|\epsilon_j^*|^2).$$

□

A.4 Beyond Ricardian Economies

A.4.1 Competitive Equilibrium with Physical Capital and Input-Output Linkages.

For any origin i, any destination j, and any sector k, the profit maximization condition (5) generalizes to

$$(\ell_{ij,k}, k_{ij,k}, m_{ij,k}, z^s_{ij,k}) \in \text{argmax}_{(\ell,k,m,z)}\{w_{i,k}z - w_i\ell - r_ik - \sum_{o,s} p_{oi,s}m_{o,s} | z = f_k(\ell,k,m)\}, \quad (A.46)$$

$$(z^d_{ij,k}, y_{ij,k}) \in \text{argmax}_{(z,y)}\{q_{ij,k}y - v_{i,k}z | y = A_{ij,k}E_k(Z_{i,k})z\}, \quad (A.47)$$

with r_i the rental rate of capital in country i. Below we let $\pi^{\text{input}}_{ij,k}(w_{i,k}, w_i, r_i, \{p_{oi,s}\}_{o,s})$ denote the value function associated with (A.46) and $\pi^{\text{output}}_{ij,k}(q_{ij,k}, v_{i,k}, Z_{i,k})$ denote the value function associated with (A.47), with $\pi_{ij,k} \equiv \pi^{\text{input}}_{ij,k} + \pi^{\text{output}}_{ij,k}$. The market clearing conditions generalize to

$$c_{ij,k} + \sum_{d,l} m_{ijd,kl} = y_{ij,k}, \quad (A.48)$$

$$z^d_{ij,k} = z^s_{ij,k}, \quad (A.49)$$

$$\sum_{j,k} \ell_{ij,k} = L_i, \quad (A.50)$$

$$\sum_{j,k} k_{ij,k} = K_i. \quad (A.51)$$

The government's budget constraint generalizes to

$$T_i = -\sum_k s_{i,k}w_{i,k}Z_{i,k} + \sum_{j\neq i,k} t^m_{ji,k}\bar{p}_{ji,k}(c_{ji,k} + \sum_{d,l} dm_{jid,kl}) + \sum_{j\neq i} t^x_{ij,k}\bar{p}_{ij,k}y_{ij,k}. \quad (A.52)$$

In the environment with physical capital and input-output linkages, a competitive equilibrium with input subsidies, $\{s_{j,k}\}$, import tariffs, $\{t^m_{ij,k}\}$, and export taxes, $\{t^x_{ij,k}\}$ corresponds to an allocation, $\{c_{ij,k}, \ell_{ij,k}, k_{ij,k}, m_{ij,k}, z^s_{ij,k}, z^d_{ij,k}, y_{ij,k}\}$, with sector sizes, $\{Z_{i,k}\}$, good prices, $\{p_{ij,k}, q_{ij,k}, \bar{p}_{ij,k}\}$, and input prices, $\{w_i, r_i, w_{i,k}, v_{i,k}\}$, and lump-sum transfers, $\{T_j\}$, such that equations (2), (3), (6), (36), (37), and (A.46)-(A.52) hold.

A.4.2 Proof of Proposition 3

Let us first show that in the environment with physical capital and input-output linkages, equation (11) generalizes to

$$dU_j/\lambda_j = -\sum_k s_{j,k}w_{j,k}dZ_{j,k} + \sum_{i\neq j,k} t^m_{ij,k}\bar{p}_{ij,k}(dc_{ij,k} + \sum_{d,l} dm_{ijd,kl}) + \sum_{i\neq j,k} t^x_{ji,k}\bar{p}_{ji,k}dy_{ji,k}$$
$$+ \sum_k \epsilon^E_{j,k}(1 - s_{j,k})w_{j,k}dZ_{j,k} + \sum_{i\neq j,k} \epsilon^p_{ji,k}\bar{p}_{ji,k}dy_{ji,k}, \quad (A.53)$$

where the extra term, $\sum_{d,l} dm_{ijd,kl}$, captures the changes in the imports of intermediate goods from country i in sector k. We follow the same steps as in Appendix A.1 and start from the identity,

$$U_j(\tau) = V_j(p_j(\tau), I_j(\tau)), \tag{A.54}$$

The total income of country j's representative agent is now

$$\begin{aligned} I_j(\tau) = & w_j L_j + r_j K_j + \sum_{i,k} \pi^{\text{input}}_{ji,k} + \sum_{i,k} \pi^{\text{output}}_{ji,k} \\ & - \sum_k s_{j,k} w_{j,k} Z_{j,k} + \sum_{j \neq i,k} t^m_{ij,k} \bar{p}_{ij,k} (c_{ij,k} + \sum_{d,l} m_{ijd,kl}) + \sum_{j \neq i} t^x_{ji,k} \bar{p}_{ji,k} y_{ji,k}. \end{aligned}$$

Totally differentiating (A.54) therefore implies

$$\begin{aligned} dU_j = & \sum_{i,k} \frac{\partial V_j}{\partial p_{ij,k}} \times dp_{ij,k} \\ & + \frac{\partial V_j}{\partial I_j} \times \Big\{ L_j dw_j + K_j dr_j + \sum_{i,k} \left[\frac{\partial \pi^{\text{input}}_{ji,k}}{\partial w_{j,k}} dw_{j,k} + \frac{\partial \pi^{\text{input}}_{ji,k}}{\partial w_j} dw_j + \frac{\partial \pi^{\text{input}}_{ji,k}}{\partial r_j} dr_j + \sum_{o,s} \frac{\partial \pi^{\text{input}}_{ji,k}}{\partial p_{oj,s}} dp_{oj,s} \right] \\ & + \sum_{i,k} \left[\frac{\partial \pi^{\text{output}}_{ji,k}}{\partial q_{ji,k}} dq_{ji,k} + \frac{\partial \pi^{\text{output}}_{ji,k}}{\partial v_{j,k}} dv_{j,k} + \frac{\partial \pi^{\text{output}}_{ji,k}}{\partial Z_{j,k}} dZ_{j,k} \right] \\ & + \sum_k (dv_{j,k} - dw_{j,k}) Z_{j,k} + \sum_k (v_{j,k} - w_{j,k}) dZ_{j,k} + \sum_{i,k} (dp_{ij,k} - d\bar{p}_{ij,k})(c_{ij,k} + \sum_{d,l} m_{ijd,kl}) \\ & + \sum_{i,k} (p_{ij,k} - \bar{p}_{ij,k})(dc_{ij,k} + \sum_{d,l} dm_{ijd,kl}) - \sum_{i,k} (dq_{ji,k} - d\bar{p}_{ji,k}) y_{ji,k} - \sum_{i,k} (q_{ji,k} - \bar{p}_{ji,k})\, dy_{ji,k} \Big\} . \end{aligned} \tag{A.55}$$

Applying the Envelope Theorem to the utility maximization problem of country j's representative agent and the profit maximization problem of the firms now implies

$$\frac{\partial V_j}{\partial p_{ij,k}} = -\frac{\partial V_j}{\partial I_j} c_{ij,k}, \text{ for all } i \text{ and } k, \tag{A.56}$$
$$\frac{\partial \pi^{\text{input}}_{ji,k}}{\partial w_{j,k}} = z^s_{ji,k}, \text{ for all } i \text{ and } k, \tag{A.57}$$
$$\frac{\partial \pi^{\text{input}}_{ji,k}}{\partial w_j} = -\ell_{ji,k}, \text{ for all } i \text{ and } k, \tag{A.58}$$
$$\frac{\partial \pi^{\text{input}}_{ji,k}}{\partial r_j} = -k_{ji,k}, \text{ for all } i \text{ and } k, \tag{A.59}$$
$$\frac{\partial \pi^{\text{input}}_{ji,k}}{\partial p_{oj,s}} = -m_{oji,sk}, \text{ for all } i, k, o, \text{ and } s, \tag{A.60}$$
$$\frac{\partial \pi^{\text{output}}_{ji,k}}{\partial q_{ji,k}} = y_{ji,k}, \text{ for all } i \text{ and } k, \tag{A.61}$$
$$\frac{\partial \pi^{\text{output}}_{ji,k}}{\partial v_{j,k}} = -z^d_{ji,k}, \text{ for all } i \text{ and } k. \tag{A.62}$$

Substituting (A.56)-(A.62) into (A.55) then gives an equation analogous to equation (A.6),

$$dU_j = \frac{\partial V_j}{\partial I_j} \times \Bigg\{ \sum_k (v_{j,k} - w_{j,k}) dZ_{j,k} + \sum_{i,k} (p_{ij,k} - \bar{p}_{ij,k})(dc_{ij,k} + \sum_{d,l} dm_{ijd,kl}) - \sum_{i,k} (q_{ji,k} - \bar{p}_{ji,k}) dy_{ji,k}$$
$$+ \sum_{i,k} \frac{\partial \pi^{\text{output}}_{ji,k}}{\partial Z_{j,k}} dZ_{j,k} - \sum_{i \neq j,k} (c_{ij,k} + \sum_{d,l} m_{ijd,kl}) d\bar{p}_{ij,k} + \sum_{i \neq j,k} y_{ji,k} d\bar{p}_{ji,k} \Bigg\},$$

where we have also used the good market clearing condition (A.48) for local goods, $c_{jj,k} = y_{jj,k} + \sum_{d,l} m_{jjd,kl}$, as well as country j's market clearing conditions (A.49)-(A.50) for composite inputs, labor, and capital. The final part of the argument is the same as in Appendix A.1 and omitted.

Proof. Since changes in utility associated with any tax variation should be zero at the optimal policy mix, equation (A.53) implies

$$\sum_k s^*_{j,k} w^*_{j,k} dZ_{j,k} - \sum_{i \neq j,k} t^{m*}_{ij,k} \bar{p}^*_{ij,k} dc_{ij,k} - \sum_{i \neq j,k} t^{x*}_{ji,k} \bar{p}^*_{ji,k} dy_{ji,k}$$
$$= \sum_k \epsilon^{E*}_{j,k} (1 - s^*_{j,k}) w^*_{j,k} dZ_{j,k} + \sum_{i \neq j,k} \epsilon^{p*}_{ji,k} \bar{p}^*_{ji,k} dy_{ji,k}, \tag{A.63}$$

where now $\epsilon^E_{j,k} \equiv d \ln E_k(Z_{j,k}) / d \ln Z_{j,k}$. Now consider a variation that lowers final consumption of the domestic good in sector k and raises the use of that good as an intermediate by firms in that sector so that $dZ_{j,k} > 0$, holding all other equilibrium variables fixed. Specializing (A.63) to this variation implies

$$s^*_{j,k} = \epsilon^{E*}_{j,k}/(1+\epsilon^{E*}_{j,k}).$$

The variations leading to the characterization of optimal import and export taxes are the same as in the proof of Proposition 1 and omitted. □

A.4.3 Proof of Proposition 4

Proof. The only difference between the environment with and without physical capital and input-output linkages comes from the expression for the first-order welfare change. Without physical capital and input-output linkages, the proof of Proposition 2 uses the fact that

$$\begin{aligned}
&\frac{1}{\lambda^*_j}\frac{\partial \hat{V}_j(\hat{\tau}_j(\epsilon^*_j),\epsilon^*_j,0)}{\partial \tau_{j,n}} \\
&= -\sum_k s^*_{j,k}\hat{w}_j\frac{\partial \hat{L}_{j,k}}{\partial \tau_{j,n}} + \sum_{i\neq j,k} t^{x*}_{ji,k}\hat{\bar{p}}_{ji,k}\frac{\partial \hat{y}_{ji,k}}{\partial \tau_{j,n}} + \sum_k \epsilon^{E*}_{j,k}(1-s^*_{j,k})\hat{w}_j\frac{\partial \hat{L}_{j,k}}{\partial \tau_{j,n}} + \sum_{i\neq j,k}\epsilon^{p*}_{ji,k}\hat{\bar{p}}_{ji,k}\frac{\partial \hat{y}_{ji,k}}{\partial \tau_{j,n}}.
\end{aligned}$$

With physical capital and input-output linkages, equation (A.53) instead implies

$$\begin{aligned}
&\frac{1}{\lambda^*_j}\frac{\partial \hat{V}_j(\hat{\tau}_j(\epsilon^*_j),\epsilon^*_j,0)}{\partial \tau_{j,n}} \\
&= -\sum_k s^*_{j,k}\hat{w}_{j,k}\frac{\partial \hat{Z}_{j,k}}{\partial \tau_{j,n}} + \sum_{i\neq j,k} t^{x*}_{ji,k}\hat{\bar{p}}_{ji,k}\frac{\partial \hat{y}_{ji,k}}{\partial \tau_{j,n}} + \sum_k \epsilon^{E*}_{j,k}(1-s^*_{j,k})\hat{w}_{j,k}\frac{\partial \hat{Z}_{j,k}}{\partial \tau_{j,n}} + \sum_{i\neq j,k}\epsilon^{p*}_{ji,k}\hat{\bar{p}}_{ji,k}\frac{\partial \hat{y}_{ji,k}}{\partial \tau_{j,n}}.
\end{aligned}$$

Going through the same steps as in the proof of Proposition 2 delivers equations (41)-(43). □

B Online Appendix: Empirics

B.1 Baseline Specification

Given our parametric assumptions in (22), (23) and (24), equation (21) implies

$$\frac{1}{J}\sum_j[\frac{1}{\theta_k}(\ln X_{ij,k}-\ln X_{i_0 j,k})-\frac{1}{\theta_{k_0}}(\ln X_{ij,k_0}-\ln X_{i_0 j,k_0})]$$
$$=\gamma_k(\ln L_{i,k}-\ln L_{i_0,k})-\gamma_1(\ln L_{i,k_0}-\ln L_{i_0,k_0})+\frac{1}{J}\sum_j[\ln\alpha_{ij,k}-\ln\alpha_{i_0 j,k}-(\ln\alpha_{ij,k_0}-\ln\alpha_{i_0,k_0})].$$

This can be written equivalently as

$$Y_{i,k}=\delta_i+\delta_k+\gamma_k\ln L_{i,k}+\varepsilon_{i,k},$$

with the following definitions

$$\begin{aligned}
Y_{i,k}\equiv&(\frac{1}{J}\sum_j\ln X_{ij,k})/\theta_k,\\
\varepsilon_{i,k}\equiv&\frac{1}{J}\sum_{j,t}\ln\alpha_{ij,k}-\mathbb{E}[\frac{1}{J}\sum_j\ln\alpha_{ij,k}|i]-\mathbb{E}[\frac{1}{J}\sum_j\ln\alpha_{ij,k}|k]+\mathbb{E}[\frac{1}{J}\sum_j\ln\alpha_{ij,k}],\\
\delta_i\equiv&-\gamma_1\ln L_{i,k_0}+Y_{i,k_0}-\varepsilon_{i,k_0},\\
\delta_k\equiv&-\gamma_k\ln L_{i_0,k}+Y_{i_0,k}-\varepsilon_{i_0,k}+\gamma_{k_0}\ln L_{i_0,k_0}-Y_{i_0,k_0}+\varepsilon_{i_0,k_0}.
\end{aligned}$$

By construction of the productivity shocks $\varepsilon_{i,k}$, we have

$$\mathbb{E}[\varepsilon_{i,k}|i]=0\text{ for all }i,$$
$$\mathbb{E}[\varepsilon_{i,k}|k]=0\text{ for all }k.$$

In dummy variable notation, the second and first stages corresponding to our IV estimator can be expressed as,

$$Y_{i,k}=\sum_{n\in\mathcal{I}}\delta_n\times\mathbb{1}_{n=i}+\sum_{n\in\mathcal{K}}\delta_n\times\mathbb{1}_{n=k}+\sum_{n\in\mathcal{K}}\gamma_n\times(\mathbb{1}_{n=k}\times\ln L_{i,n})+\varepsilon_{i,k},\tag{B.1}$$

$$(\mathbb{1}_{n=k}\times\ln L_{i,n})=\sum_{m\in\mathcal{I}}\tilde{\delta}_{m,n}\times\mathbb{1}_{m=i}+\sum_{m\in\mathcal{K}}\tilde{\delta}_{m,n}\times\mathbb{1}_{m=k}+\sum_{m\in\mathcal{K}}\tilde{\gamma}_{m,n}\times(\mathbb{1}_{m=k}\times\ln\hat{L}_{i,m})+\tilde{\varepsilon}^n_{i,k},\text{ for all }n\in\mathcal{K},\tag{B.2}$$

where $\mathcal{I}$ and $\mathcal{K}$ denote the set of countries and sectors, respectively; $\{\mathbb{1}_{n=i}\}_{n\in\mathcal{I}}$ is a vector of country-specific dummy variables; $\{\mathbb{1}_{n=k}\}_{n\in\mathcal{K}}$ is a vector of sector-specific dummy variables; and $\tilde{\varepsilon}^n_{i,k}$ is the first-stage residual. In equation (B.2) the coefficient $\tilde{\gamma}_{m,n}$ represents the (conditional on other regressors) projection of the endogenous variable $(\mathbb{1}_{n=k}\times\ln L_{i,n})$ on the instrument $(\mathbb{1}_{m=k}\times$

$\ln \hat{L}_{i,m}$). Table B.4 reports the "diagonal" element (i.e. $\tilde{\gamma}_{n,n}$) of the matrix of these first-stage coefficients, as well as the maximum (in absolute value) of the "off-diagonal" coefficients (i.e. $\max_{m\neq n} |\tilde{\gamma}_{m,n}|$), for each sector $n \in \mathcal{K}$.

B.2 Threats to the Exclusion Restriction

Condition (28). By definition, we have

$$\varepsilon_{i,k} \equiv \frac{1}{J}\sum_j \ln \alpha_{ij,k} - \mathbb{E}[\frac{1}{J}\sum_j \ln \alpha_{ij,k}|i] - \mathbb{E}[\frac{1}{J}\sum_j \ln \alpha_{ij,k}|k] + \mathbb{E}[\frac{1}{J}\sum_j \ln \alpha_{ij,k}],$$
$$\alpha_{ij,k} \equiv A_{ij,k}(1-t^x_{ij,k})/[(1-s_{i,k})(1+t^m_{ij,k})].$$

Suppose that $A_{ij,k} \propto (\hat{L}_i)^{\Gamma_k}$ and that other determinants of productivity as well as taxes and subsidies are orthogonal to $\hat{L}_i$. Then,

$$\begin{aligned}
&\mathbb{E}[\ln \hat{L}_i \times \varepsilon_{i,k}|k] \\
&= \mathbb{E}[\ln \hat{L}_i \times \{\frac{1}{J}\sum_j \ln A_{ij,k} - \mathbb{E}[\frac{1}{J}\sum_j \ln A_{ij,k}|i] - \mathbb{E}[\frac{1}{J}\sum_j \ln A_{ij,k}|k] + \mathbb{E}[\frac{1}{J}\sum_j \ln A_{ij,k}]\}|k] \\
&\quad = \Gamma_k \mathbb{E}[(\ln \hat{L}_i)^2|k] - \mathbb{E}(\Gamma_k|i)\mathbb{E}[(\ln \hat{L}_i)^2|k] - \Gamma_k(\mathbb{E}[\ln \hat{L}_i|k])^2 + \mathbb{E}[\Gamma_k \ln \hat{L}_i]\mathbb{E}[\ln \hat{L}_i|k] \\
&\qquad\qquad = \Gamma_k \mathbb{E}[(\ln \hat{L}_i)^2] - \mathbb{E}(\Gamma_k)\mathbb{E}[(\ln \hat{L}_i)^2] - \Gamma_k(\mathbb{E}[\ln \hat{L}_i])^2 + \mathbb{E}[\Gamma_k \ln \hat{L}_i]\mathbb{E}[\ln \hat{L}_i].
\end{aligned}$$

By the Law of Iterated Expectations, $\mathbb{E}[\Gamma_k \ln \hat{L}_i] = \mathbb{E}[\mathbb{E}(\Gamma_k \ln \hat{L}_i|k)] = \mathbb{E}[\Gamma_k \mathbb{E}(\ln \hat{L}_i|k)] = \mathbb{E}[\Gamma_k]\mathbb{E}[\ln \hat{L}_i]$. Thus, the previous expression simplifies into

$$\mathbb{E}[\ln \hat{L}_i \times \varepsilon_{i,k}|k] = [\Gamma_k - \mathbb{E}(\Gamma_k)] \times \text{Var}(\ln \hat{L}_i).$$

Condition (29). Similarly, suppose that $\hat{\beta}_{i,k} \propto (\bar{\alpha}_{i,k})^{\phi}$, with $\ln \bar{\alpha}_{i,k} \equiv \frac{1}{J}\sum_j \ln \alpha_{ij,k} - \mathbb{E}[\frac{1}{J}\sum_j \ln \alpha_{ij,k}|i]$ and other determinants of $\hat{\beta}_{i,k}$ orthogonal to $\varepsilon_{i,k}$. Then

$$\begin{aligned}
&\mathbb{E}[\ln \hat{\beta}_{ik} \times \varepsilon_{i,k}|k] \\
&\qquad = \phi \mathbb{E}[(\varepsilon_{i,k} + \mathbb{E}[\frac{1}{J}\sum_j \ln \alpha_{ij,k}|k] - \mathbb{E}[\frac{1}{J}\sum_j \ln \alpha_{ij,k}]) \times \varepsilon_{i,k}|k] \\
&\qquad\quad = \phi \mathbb{E}[(\varepsilon^2_{i,k}|k)] + \phi \mathbb{E}[\varepsilon_{i,k}|k] \times (\mathbb{E}[\frac{1}{J}\sum_j \ln \alpha_{ij,k}|k] - \mathbb{E}[\frac{1}{J}\sum \ln_j \alpha_{ij,k}]).
\end{aligned}$$

Since $\mathbb{E}[\varepsilon_{i,k}|k] = 0$, this simplifies into

$$\mathbb{E}[\ln \hat{\beta}_{ik} \times \varepsilon_{i,k}|k] = \phi \times \text{Var}(\varepsilon_{i,k}|k).$$

B.3 Estimation of Trade Elasticities

B.3.1 Specifications from Prior Studies

We first demonstrate that the empirical specifications used to estimate sector-level trade elasticities by Caliendo and Parro (2015), Shapiro (2016), and Giri, Yi and Yilmazkuday (2021) remain valid in the context of our model.

Caliendo and Parro's (2015) Specification. In our model, (2), (3), (4), and (19) imply

$$p_{ij,k} = \frac{(1+t^m_{ij,k})}{(1-t^x_{ij,k})} \frac{(1-s_{i,k})w_i}{A_{ij,k}E_k(L_{i,k})}. \tag{B.3}$$

Using this expression to substitute for the consumer price into (30) implies

$$\ln(X_{ij,k}/X_{j,k}) = \psi_{j,k} + \psi_{i,k} - \theta_k \ln(1+t^m_{ij,k}) + \psi_{ij,k}, \tag{B.4}$$

where $\psi_{j,k}$, $\psi_{i,k}$, and $\psi_{ij,k}$ are such that

$$\begin{aligned}
\psi_{j,k} &\equiv \theta_k \times \ln P_{j,k}, \\
\psi_{i,k} &\equiv \theta_k \times (-\ln(1-s_{i,k}) - \ln w_i + \ln E_k(L_{i,k})), \\
\psi_{ij,k} &\equiv \theta_k \times (\ln A_{ij,k} + \ln(1-t^x_{ij,k})).
\end{aligned}$$

Starting from (B.4) and comparing bilateral trade flows between three countries, i, j, and l, within any given sector, k, we obtain (using our notation) Caliendo and Parro's (2015) specification,

$$\ln\left(\frac{X_{ij,k}X_{li,k}X_{jl,k}}{X_{lj,k}X_{ji,k}X_{il,k}}\right) = -\theta_k \ln\left[\frac{(1+t^m_{ij,k})(1+t^m_{lj,k})(1+t^m_{jl,k})}{(1+t^m_{lj,k})(1+t^m_{ji,k})(1+t^m_{il,k})}\right] + \psi_{ijl,k},$$

with $\psi_{ijl,k} \equiv \psi_{ij,k} - \psi_{lj,k} + \psi_{lj,k} - \psi_{ji,k} + \psi_{jl,k} - \psi_{il,k}$.

Shapiro's (2016) Specification. Starting again from (B.4) and letting $\xi_{ij,k} \equiv A_{ij,k}(1+\text{shipping}_{ij,k})$, with $\text{shipping}_{ij,k}$ denoting the observed freight costs between country i and country j in sector k, we obtain (using our notation) Shapiro's (2016) specification,

$$\ln(X_{ij,k}/X_{j,k}) = \psi_{j,k} + \psi_{i,k} - \theta_k \ln(1+\text{shipping}_{ij,k}) + \tilde{\psi}_{ij,k},$$

with $\tilde{\psi}_{ij,k} \equiv \theta_k \times (\ln \xi_{ij,k} + \ln(1-t^x_{ij,k}) - \ln(1+t^m_{ij,k}))$.

Giri et al.'s (2021) Specification. Starting directly from equation (30) implies

$$\theta_k = -\frac{\sum_{i,j}[\ln(X_{ij,k}/X_{j,k}) - \ln(X_{ii,k}/X_{i,k})]}{\sum_{i,j}[\ln(p_{ij,k}/P_{j,k}) - \ln(p_{ii,k}/P_{i,k})]}.$$

This is the starting point of Giri et al.'s (2021) SMM estimation using micro-level price data, as in Simonovska and Waugh (2012).[45]

B.3.2 Estimates from Prior Studies

Table B.1: Trade Elasticity (θ_k) Estimates from Prior Studies

Sector	CP (1)	Shapiro (2)	GYY (3)	Median (4)
Food, Beverages and Tobacco	2.6	5.3	3.6	3.6
Textiles	8.1	18.6	4.4	8.1
Wood Products	11.5	5.9	4.2	5.9
Paper Products	16.5	5.8	3.0	5.8
Coke/Petroleum Products	64.9	9.0	3.8	9.0
Chemicals	3.1	1.6	3.8	3.1
Rubber and Plastics	1.7	1.6	4.1	1.7
Mineral Products	2.4	12.9	5.1	5.1
Basic Metals	3.3	12.9	8.9	8.9
Fabricated Metals	7.0	12.9	5.1	7.0
Machinery and Equipment	1.5	10.8	3.3	3.3
Computers and Electronics	13.0	10.8	3.3	10.8
Electrical Machinery, NEC	12.9	10.8	3.3	10.8
Motor Vehicles	1.8	6.9	4.5	4.5
Other Transport Equipment	0.4	6.9	4.5	4.5

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 1, column (4) in Caliendo and Parro (2015); column (2) to Table 4, column (2) in Shapiro (2016); column (3) to Table 4 in Giri et al. (2021); and column (4) reports the median of columns (1)-(3).

[45] Compared to the other two procedures, this estimator takes into account the existence of differentiated varieties within each sector. Although we have abstracted from modeling such heterogeneity explicitly in Section 2, it can be accommodated without changing any of our results, as shown in our working paper Bartelme et al. (2019).

B.4 Estimation of Elasticity of Substitution Between Sectors

B.4.1 Baseline Specification

Equation (30) implies

$$\ln P_{j,k} = \frac{1}{J}\sum_i \frac{1}{\theta_k}\ln(X_{ij,k}/X_{j,k}) + \frac{1}{J}\sum_i \ln p_{ij,k}.$$

Equation (B.3) and the definition of $\alpha_{ij,k} \equiv A_{ij,k}(1-t^x_{ij,k})/[(1-s_{i,k})(1+t^m_{ij,k})]$ further imply

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

Letting $\ln \hat{P}_{j,k} \equiv \frac{1}{J}\sum_i \ln(X_{ij,k}/X_{j,k})/\theta_k$, $\ln\hat{\alpha}_{j,k} \equiv \frac{1}{J}\sum_i \ln\alpha_{ij,k}$, and $\ln\chi_k \equiv \frac{1}{J}\sum_i \ln[E_k(L_{i,k})/w_i]$, equation (31) can be rearranged as

$$\ln X_{j,k} = \phi_j + \phi_k - \rho\ln\hat{P}_{j,k} + \phi_{j,k},$$

where ϕ_j, ϕ_k, and $\phi_{j,k}$ are such that

$$\begin{aligned}
\phi_j &\equiv \ln(X_j P_j^{\rho}) + \mathbb{E}[\ln(\beta_{j,k}\hat{\alpha}^{\rho}_{j,k})|j],\\
\phi_k &\equiv \ln\chi_k + \mathbb{E}[\ln(\beta_{j,k}\hat{\alpha}^{\rho}_{j,k})|k] - \mathbb{E}[\ln(\beta_{j,k}\hat{\alpha}^{\rho}_{j,k})],\\
\phi_{j,k} &\equiv \ln(\beta_{j,k}\hat{\alpha}^{\rho}_{j,k}) - \mathbb{E}[\ln(\beta_{j,k}\hat{\alpha}^{\rho}_{j,k})|j] - \mathbb{E}[\ln(\beta_{j,k}\hat{\alpha}^{\rho}_{j,k})|k] + \mathbb{E}[\ln(\beta_{j,k}\hat{\alpha}^{\rho}_{j,k})].
\end{aligned}$$

By construction, the error term $\phi_{j,k}$ is demeaned so that

$$\begin{aligned}
\mathbb{E}[\phi_{j,k}|j] &= 0 \text{ for all } j,\\
\mathbb{E}[\phi_{j,k}|k] &= 0 \text{ for all } k.
\end{aligned}$$

The additional orthogonality conditions imposed by our IV estimator are

$$\mathbb{E}[\ln\hat{L}_j \times \phi_{j,k}|k] = 0 \text{ for all } k.$$

In dummy variable notation, the second and first stages corresponding to our IV estimator can be expressed as,

$$\ln X_{j,k} = \sum_{n\in\mathcal{I}}\phi_n \times \mathbb{1}_{n=i} + \sum_{n\in\mathcal{K}}\phi_n \times \mathbb{1}_{n=k} - \rho\ln\hat{P}_{j,k} + \phi_{j,k}, \tag{B.5}$$

$$\ln\hat{P}_{j,k} = \sum_{n\in\mathcal{I}}\tilde{\phi}_n \times \mathbb{1}_{n=i} + \sum_{n\in\mathcal{K}}\tilde{\phi}_n \times \mathbb{1}_{n=k} + \sum_{n\in\mathcal{K}}\tilde{\rho}_n \times (\mathbb{1}_{n=k}\times\hat{L}_j) + \tilde{\phi}_{j,k}. \tag{B.6}$$

B.4.2 Estimates

Table B.2: First-Stage Estimates from Upper-Level Preference Parameter (ρ)

Sector	Coeff.	Sector	Coeff.
Food, Beverages and Tobacco	-0.04	Basic Metals	0.03
	(0.02)		(0.01)
Textiles	0.01	Fabricated Metals	0.02
	(0.01)		(0.01)
Wood Products	0.00	Machinery and Equipment	0.00
	(0.01)		(0.01)
Paper Products	0.01	Computers and Electronics	0.03
	(0.01)		(0.01)
Coke/Petroleum Products	0.02	Electrical Machinery, NEC	0.03
	(0.01)		(0.01)
Chemicals	0.00	Motor Vehicles	-0.04
	(0.01)		(0.01)
Rubber and Plastics	-0.12	Other Transport Equipment	0.00
	(0.02)		(0.00)
Mineral Products	-0.01		
	(0.01)		
Within R^2		0.25	
Observations		915	

Notes: This table reports the first-stage coefficients corrresponding to $\{\tilde{\rho}_n\}$ in equation (B.6). Standard errors clustered at the country-sector level.

Table B.3: Estimate of Upper-Level Preference Parameter (ρ)

	log (sectoral expenditure share)	log (sectoral expenditure share)
	OLS	IV
	(1)	(2)
Log Prices	1.65	0.28
	(0.21)	(0.20)
Within R^2	0.18	
First-state F-statistic		12.9
Observations	915	915

Notes: This table reports OLS and IV estimates of ρ from equation (B.5). Table B.2 reports the first-stage coefficients from the IV specification. Standard errors in parentheses are clustered at the country level.

B.5 Scale Elasticities

Table B.4: Summary of First-Stage Regressions, Sector-Level Scale Elasticities (γ_k)

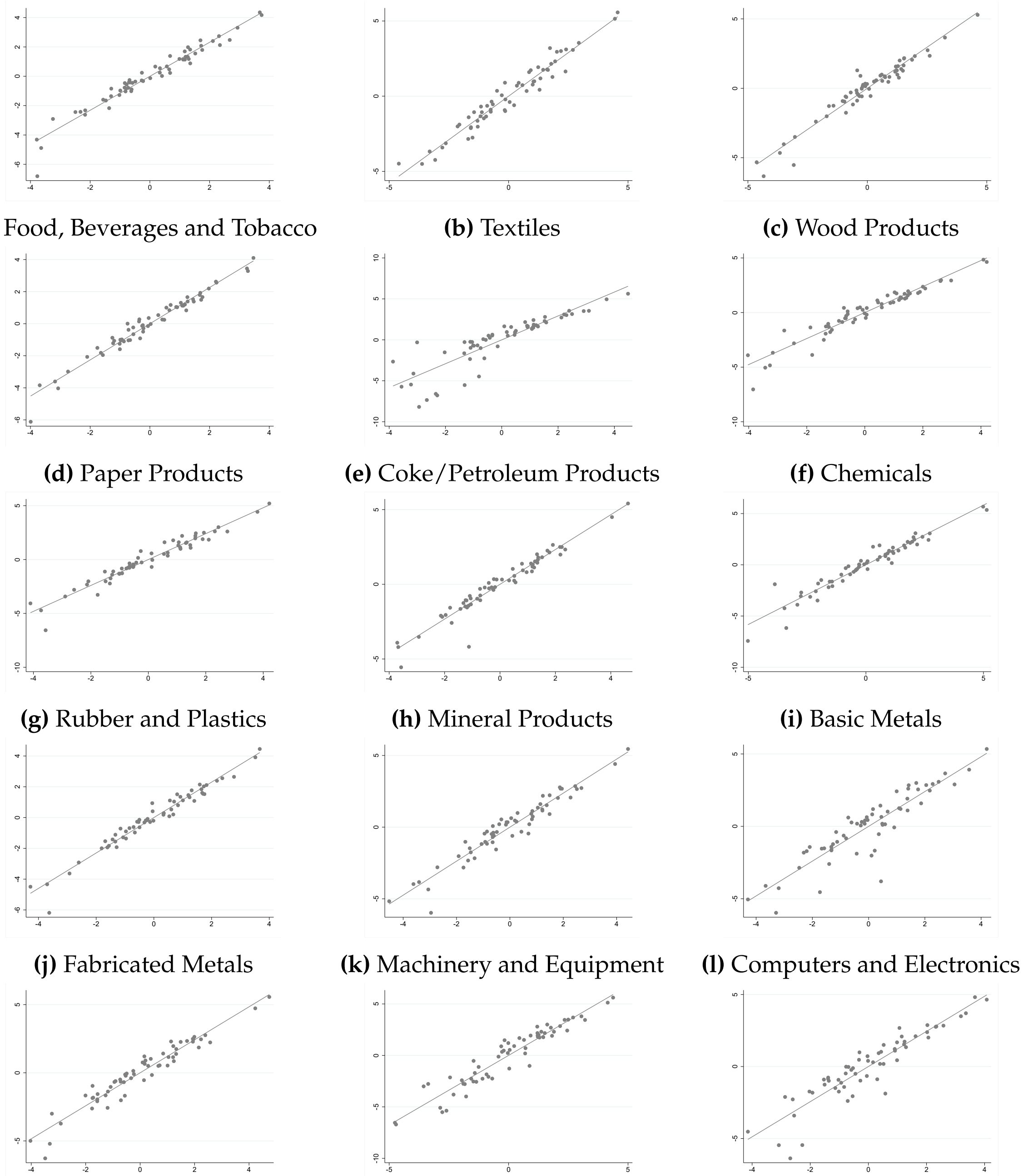
Sector	Diagonal Coeff. (1)	Max. (Abs) Off-Diagonal Coeff. (2)	F-Stat (3)	SW F-Stat (4)
Food, Beverages and Tobacco	1.16	0.01	172.80	1176.6
	(0.06)	(0.01)		
Textiles	1.18	0.02	105.60	1999.7
	(0.04)	(0.01)		
Wood Products	1.20	0.02	156.90	1126.9
	(0.06)	(0.02)		
Paper Products	1.12	-0.01	155.90	1671.5
	(0.05)	(0.01)		
Coke/Petroleum Products	1.40	-0.08	15.40	383.0
	(0.14)	(0.05)		
Chemicals	1.19	0.00	63.30	902.6
	(0.07)	(0.01)		
Rubber and Plastics	1.21	0.01	94.80	1426.4
	(0.06)	(0.01)		
Mineral Products	1.18	0.01	279.10	1091.7
	(0.04)	(0.01)		
Basic Metals	1.17	0.02	184.70	1234.1
	(0.06)	(0.01)		
Fabricated Metals	1.16	0.01	179.70	2020.7
	(0.05)	(0.01)		
Machinery and Equipment	1.19	0.01	122.20	1334.4
	(0.04)	(0.01)		
Computers and Electronics	1.24	0.04	77.60	608.3
	(0.07)	(0.05)		
Electrical Machinery, NEC	1.22	0.02	113.90	1193.0
	(0.05)	(0.01)		
Motor Vehicles	1.34	-0.01	86.60	1055.1
	(0.06)	(0.03)		
Other Transport Equipment	1.16	-0.07	26.40	745.1
	(0.07)	(0.02)		

Notes: This table summarizes the first-stage coefficients $\{\tilde{\gamma}_{m,n}\}$ in equation (B.2). Column (1) reports the diagonal coefficient ($\tilde{\gamma}_{n,n}$) and column (2) the maximum (in absolute value) among the 14 off-diagonal coefficients ($\{\tilde{\gamma}_{m,n}\}_{m \neq n}$) for each of the first-stage regressions corresponding to the endogenous regressor formed by interacting the log of sector size with an indicator for the sector named in each row. Standard errors in parentheses are clustered at the exporter level. Column (3) reports the corresponding conventional F-statistic, and column (4) the Sanderson-Windmeijer F-statistic, from each first-stage.

Table B.5: Estimates of Sector-Level Scale Elasticities (γ_k), Heterogeneity over Time

Sector	IV 1995 (1)	IV 2000 (2)	IV 2005 (3)	IV All Years (4)
Food, Beverages and Tobacco	0.17	0.18	0.20	0.19
	(0.02)	(0.02)	(0.02)	(0.02)
Textiles	0.10	0.10	0.11	0.11
	(0.02)	(0.02)	(0.01)	(0.01)
Wood Products	0.07	0.10	0.11	0.10
	(0.03)	(0.03)	(0.02)	(0.02)
Paper Products	0.11	0.09	0.14	0.12
	(0.02)	(0.02)	(0.02)	(0.01)
Coke/Petroleum Products	0.07	0.06	0.08	0.08
	(0.02)	(0.01)	(0.01)	(0.01)
Chemicals	0.21	0.20	0.22	0.22
	(0.02)	(0.02)	(0.02)	(0.02)
Rubber and Plastics	0.41	0.37	0.40	0.40
	(0.06)	(0.06)	(0.05)	(0.05)
Mineral Products	0.13	0.14	0.16	0.15
	(0.02)	(0.02)	(0.02)	(0.02)
Basic Metals	0.06	0.06	0.08	0.07
	(0.02)	(0.01)	(0.01)	(0.01)
Fabricated Metals	0.08	0.08	0.10	0.10
	(0.02)	(0.02)	(0.02)	(0.02)
Machinery and Equipment	0.20	0.19	0.21	0.21
	(0.03)	(0.03)	(0.02)	(0.02)
Computers and Electronics	0.05	0.05	0.07	0.06
	(0.02)	(0.01)	(0.01)	(0.01)
Electrical Machinery, NEC	0.06	0.05	0.07	0.06
	(0.02)	(0.02)	(0.02)	(0.01)
Motor Vehicles	0.16	0.16	0.18	0.17
	(0.02)	(0.02)	(0.01)	(0.01)
Other Transport Equipment	0.14	0.15	0.15	0.15
	(0.03)	(0.02)	(0.02)	(0.02)
Observations	915	914	913	3,657

Notes: Columns (1)-(3) report the IV estimates of equation (25) when the sample is a single cross-section of countries and sectors from the indicated year; column (4) does the same for a sample that pools across all years. The instruments are the log of (country population $\times$ sectoral demand shifter), interacted with sector indicators. The regressions in columns (1)-(3) include exporter and sector fixed effects, while the regression in column (4) includes exporter-year and sector-year fixed effects. Standard errors in parentheses are clustered at the exporter level.



(a) Food, Beverages and Tobacco **(b)** Textiles **(c)** Wood Products

(d) Paper Products **(e)** Coke/Petroleum Products **(f)** Chemicals

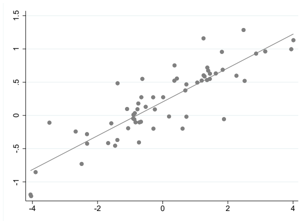
(g) Rubber and Plastics **(h)** Mineral Products **(i)** Basic Metals

(j) Fabricated Metals **(k)** Machinery and Equipment **(l)** Computers and Electronics

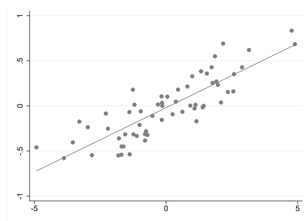
(m) Electrical Machinery, NEC **(n)** Motor Vehicles **(o)** Other Transport Equipment

Figure B.1: First-Stage Fit, Scale Elasticity Estimation

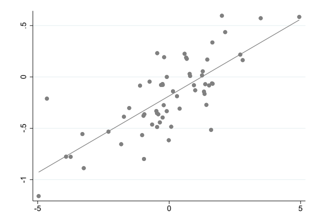
Notes: Each sub-figure corresponds to a sector. To construct the figure for any given sector, n, we proceed as follows. Step 1: project the endogenous variable corresponding to sector n (an interaction between $\ln L_{i,k}$ and an indicator for $k = n$) on all other regressors in equation (25), namely a country fixed effect, a sector fixed effect, and the 14 other interactions between sector indicators and (log) sector size. Step 2: project the instrument corresponding to sector n (an interaction between $\ln \hat{L}_{i,k}$ and an indicator for $k = n$) on the same regressors as in the previous step. The figure for sector n then plots the residuals from step 1 on the y-axis against those from step 2 on the x-axis, for the I observations with $k = n$.



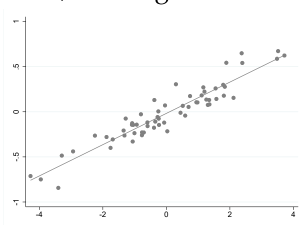
(a) Food, Beverages and Tobacco



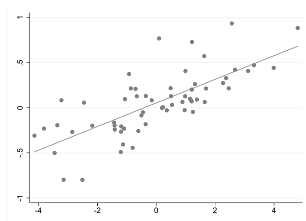
(b) Textiles



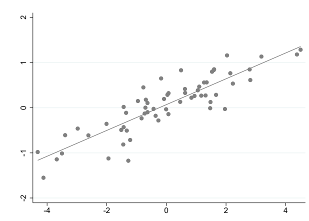
(c) Wood Products



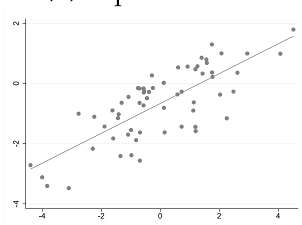
(d) Paper Products



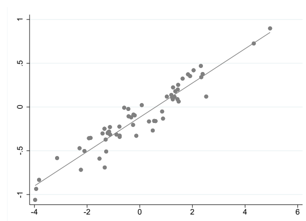
(e) Coke/Petroleum Products



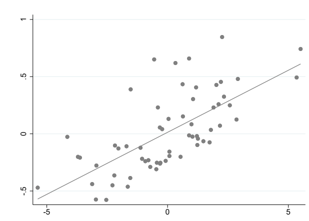
(f) Chemicals



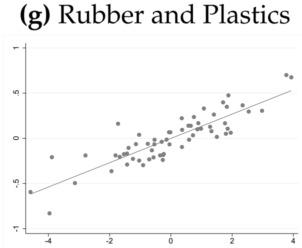
(g) Rubber and Plastics



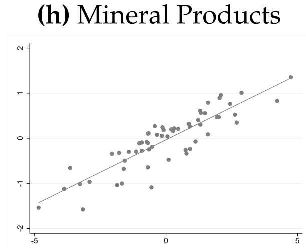
(h) Mineral Products



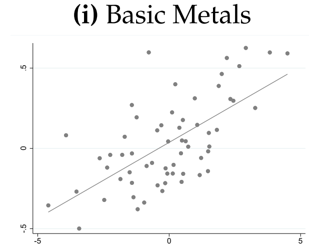
(i) Basic Metals



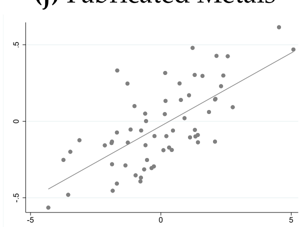
(j) Fabricated Metals



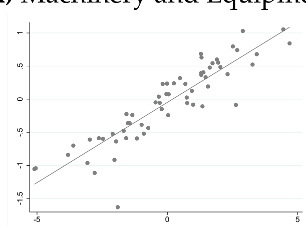
(k) Machinery and Equipment



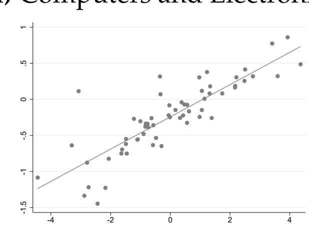
(l) Computers and Electronics



(m) Electrical Machinery, NEC



(n) Motor Vehicles



(o) Other Transport Equipment

Figure B.2: Reduced-Form Fit, Scale Elasticity Estimation

Notes: Each sub-figure corresponds to a sector. To construct the figure for any given sector, n, we proceed as follows. Step 1: project $Y_{i,k}$ on a country fixed effect, a sector fixed effect, and the 14 instruments (interactions between sector indicators and $\ln \hat{L}_{i,k}$) defined for all sectors apart from n. Step 2: project the sector-n instrument (an interaction between $\ln \hat{L}_{i,k}$ and an indicator for $k = n$) on the same regressors as in the previous step. The figure for sector n then plots the residuals from step 1 on the y-axis against those from step 2 on the x-axis, but only for the I observations with $k = n$.

Table B.6: Estimates of Sector-Level Scale Elasticities (γ_k), Alternative IVs

Sector	Baseline (1)	ρ=-.9 (2)	ρ=0 (3)	ρ=2 (4)	Top 50% Distances (5)
Food, Beverages and Tobacco	0.22	0.22	0.22	0.24	0.23
	(0.02)	(0.02)	(0.02)	(0.04)	(0.02)
Textiles	0.12	0.12	0.12	0.14	0.13
	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)
Wood Products	0.13	0.13	0.13	0.14	0.13
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
Paper Products	0.15	0.15	0.15	0.17	0.16
	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)
Coke/Petroleum Products	0.09	0.08	0.09	0.10	0.10
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Chemicals	0.24	0.24	0.24	0.26	0.25
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
Rubber and Plastics	0.42	0.41	0.41	0.45	0.44
	(0.04)	(0.04)	(0.04)	(0.07)	(0.04)
Mineral Products	0.17	0.16	0.17	0.19	0.18
	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)
Basic Metals	0.09	0.09	0.09	0.11	0.10
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
Fabricated Metals	0.12	0.11	0.12	0.14	0.13
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
Machinery and Equipment	0.24	0.24	0.24	0.26	0.26
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
Computers and Electronics	0.08	0.07	0.08	0.10	0.09
	(0.01)	(0.02)	(0.01)	(0.03)	(0.01)
Electrical Machinery, NEC	0.08	0.07	0.08	0.10	0.09
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
Motor Vehicles	0.18	0.18	0.18	0.19	0.19
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Other Transport Equipment	0.18	0.17	0.18	0.20	0.19
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)

Notes: IV estimates of equation (25) for alternative values of the demand residuals $\hat{\beta}_{i,k}$ entering demand-predicted sector size $\hat{L}_{i,k} \equiv \hat{\beta}_{i,k} \times \hat{L}_i$. Column (1) repeats column (2) from Table 1 for purposes of comparison. Columns (2)-(4) construct demand residuals $\hat{\beta}_{i,k}$ using alternative values of the upper-level preference parameter ρ. Column (5) constructs demand residuals $\hat{\beta}_{i,k}$ using alternative price proxies $\hat{P}_{i,k}$, as described in Section 4.4. All regressions include exporter and sector fixed effects. Standard errors in parentheses are clustered at the exporter level.

Table B.7: Estimates of Sector-Level Scale Elasticities (γ_k), Additional Controls

Sector	(1)	(2)	(3)
Food, Beverages and Tobacco	0.22	0.23	0.23
	(0.02)	(0.03)	(0.03)
Textiles	0.12	0.10	0.09
	(0.01)	(0.02)	(0.01)
Wood Products	0.13	0.13	0.13
	(0.02)	(0.02)	(0.02)
Paper Products	0.15	0.16	0.16
	(0.01)	(0.03)	(0.02)
Coke/Petroleum Products	0.09	0.07	0.07
	(0.01)	(0.02)	(0.02)
Chemicals	0.24	0.27	0.28
	(0.02)	(0.02)	(0.02)
Rubber and Plastics	0.42	0.53	0.55
	(0.04)	(0.02)	(0.04)
Mineral Products	0.17	0.19	0.20
	(0.01)	(0.03)	(0.01)
Basic Metals	0.09	0.07	0.06
	(0.01)	(0.02)	(0.01)
Fabricated Metals	0.12	0.12	0.12
	(0.02)	(0.02)	(0.01)
Machinery and Equipment	0.24	0.26	0.27
	(0.02)	(0.02)	(0.01)
Computers and Electronics	0.08	0.06	0.06
	(0.01)	(0.02)	(0.02)
Electrical Machinery, NEC	0.08	0.06	0.05
	(0.01)	(0.02)	(0.01)
Motor Vehicles	0.18	0.19	0.19
	(0.01)	(0.02)	(0.01)
Other Transport Equipment	0.18	0.19	0.19
	(0.02)	(0.02)	(0.02)
(Contract enforcement)×(sector fixed-effects)		✓	✓
(Financial development)×(sector fixed-effects)		✓	✓
(GDP per capita)×(sector fixed-effects)			✓

Notes: IV estimates of equation (25). All regressions include exporter and sector fixed effects. Column (1) repeats column (2) from Table 1 for purposes of comparison. Standard errors in parentheses are clustered at the exporter level. The number of observations for column (1) is 915 and for columns (2) and (3) it is 735.

B.5.1 Beyond Ricardian Economies

Like in Appendix C.5, we assume that production functions take the nested CES form as in equations (38) and (39). Maintaining our other parametric restrictions described in equations (22)-(24), this leads to the following generalization of equation (25),

$$Y^{IO}_{i,k} = \delta_i + \delta_k + \gamma_k \ln Z_{i,k} + \varepsilon_{i,k}, \tag{B.7}$$

where the left-hand side variable now also controls for differences in input costs across sectors,

$$Y^{IO}_{i,k} \equiv \frac{1}{J}\sum_j (\ln X_{ij,k}/\theta_k) + (1-a_k)b_k \ln(r_i/w_i) + \sum_s b_{sk} \ln(P_{i,s}/w_i). \tag{B.8}$$

To measure $Y^{IO}_{i,k}$ and $Z_{i,k}$, we first draw on trade and national accounts data in order to obtain $\{a_k, b_k, b_{sk}, r_i, w_i\}$. We use the OECD's ICIO tables to compute b_{sk} as the share of total input costs of firms from sector k spent on intermediate goods from sector s. This also gives us the value added share $b_k = 1 - \sum_s b_{sk}$. Unfortunately, this dataset does not report labor shares of value added a_k, so we obtain those from the World Input Output Database (after creating an industry concordance with our dataset) by computing the ratio of total labor spending by firms from sector k to total value added by those firms.[46] We next compute factor prices w_i and r_i by using the formulas $w_i = (\sum_k a_k b_k S_{i,k})/\hat{L}_i$ and $r_i = (\sum_k (1-a_k) b_k S_{i,k})/\hat{K}_i$, with $\hat{L}_i$ and $\hat{K}_i$ as in the Penn World Tables 9.0 (the variables "POP" and "CK", respectively), and with $S_{i,k} \equiv \sum_j X_{ij,k}$ representing total sales by country i in sector k.

In order to measure $Y^{IO}_{i,k}$ and $Z_{i,k}$, we also need proxies of price indices in both traded sectors, $k \in \mathcal{T}$, and non-traded sectors, $k \in \mathcal{N}$. For traded sectors, we use the same approach as in Section 4.2 and use as our proxy for $\ln P_{i,k}$,

$$\ln \hat{P}_{i,k} \equiv \frac{1}{J}\sum_j \ln(X_{ji,k}/X_{i,k})/\theta_k, \text{ for all } k \in \mathcal{T}. \tag{B.9}$$

For non-traded sectors $k \in \mathcal{N}$, we turn to the zero-profit conditions,

$$\ln P_{i,k} = a_k b_k \ln w_i + (1-a_k) b_k \ln r_i + \sum_s b_{sk} \ln P_{i,s} + \varphi_{i,k}, \text{ for all } k \in \mathcal{N},$$

with $\varphi_{i,k} \equiv -[\ln A_{ii,k} + a_k b_k \ln a_k + (1-a_k) b_k \ln(1-a_k) + b_k \ln b_k + \sum_s b_{sk} \ln b_{sk}]$, where we have invoked our baseline assumption of no scale economies in non-manufacturing sectors, which include all non-tradable sectors. The vector of non-tradable prices, $\ln P^N_i \equiv \{\ln P_{i,k}\}_{k \in \mathcal{N}}$, that solves

[46] We use data from 2005-2007, since there are minor discrepancies between total value added and the sum of labor and capital compensation in the data for later years. Results are not sensitive to this choice.

the previous system of linear equations is equal to

$$\ln P_i^N = (I - B^{NN})^{-1}[\ln w_i \times AB + \ln r_i \times (1-A)B + B^{NT} \ln P_i^T + \Phi_i]$$

where I is the identity matrix with dimension the number of non-traded sectors, $B^{NN} \equiv \{b_{sk}\}_{(s,k)\in\mathcal{N}\times\mathcal{N}}$, $B^{NT} \equiv \{b_{sk}\}_{(s,k)\in\mathcal{T}\times\mathcal{N}}$, $AB \equiv \{a_k b_k\}_{k\in\mathcal{N}}$, $(1-A)B \equiv \{(1-a_k)b_k\}_{k\in\mathcal{N}}$, $\ln P_i^T \equiv \{\ln P_{i,k}\}_{k\in\mathcal{T}}$, and $\Phi_j \equiv \{\varphi_{i,k}\}_{k\in\mathcal{N}}$. We then use as our proxy for non-tradable prices,

$$\ln \hat{P}_i^N = (I - B^{NN})^{-1}[\ln w_i \times AB + \ln r_i \times (1-A)B + B^{NT} \ln \hat{P}_i^T]. \tag{B.10}$$

We are now ready to derive measures of $Y_{i,k}^{\text{IO}}$ and $Z_{i,k}$. Substituting for $\{\ln P_{i,k}\}$ in equation (B.8) using (B.9) and (B.10), we can express $Y_{i,k}^{\text{IO}}$ as

$$Y_{i,k}^{\text{IO}} = \hat{Y}_{i,k}^{\text{IO}} - \zeta_{i,k} - \hat{E}_k, \tag{B.11}$$

where $\hat{Y}_{i,k}^{\text{IO}}$ is only a function of observables, $\zeta_{i,k}$ captures potential sources of discrepancies between tradable and non–tradable prices and our proxies that are related to primitive productivity differences, and $\hat{E}_k$ captures those related to endogenous wages and scale effects,

$$\begin{aligned}
\hat{Y}_{i,k}^{\text{IO}} &\equiv \frac{1}{J}\sum_j (\ln X_{ij,k}/\theta_k) + (1-a_k)b_k \ln(r_i/w_i) + \sum_s b_{sk} \ln(\hat{P}_{i,s}/w_i),\\
\zeta_{i,k} &\equiv \sum_{s\in\mathcal{K}^T} b_{sk} \ln \hat{\alpha}_{i,s} + \sum_{s\in\mathcal{K}^N} b_{sk}\xi_{i,s}\\
\hat{E}_k &\equiv \sum_{s\in\mathcal{K}^T} b_{sk}\bar{E}_s + \sum_{s\in\mathcal{K}^N} b_{sk}\tilde{E}_s,\\
\bar{E}_k &\equiv \frac{1}{J}\sum_i \ln\left(\frac{E_k(Z_{i,k})}{v_{i,k}}\right)\\
\{\xi_{i,k}\} &\equiv (I - B^{NN})^{-1}[-\Phi_i + B^{NT}\{\ln \hat{\alpha}_{i,k}\}]\\
\{\tilde{E}_k\} &\equiv (I - B^{NN})^{-1}[B^{NT}\{\bar{E}_k\}].
\end{aligned}$$

Similarly, in order to measure $Z_{i,k}$, we can use the fact that

$$\begin{aligned}
w_{i,k}Z_{i,k} &= S_{i,k},\\
w_{i,k} &= \left[(a_k b_k)^{-a_k b_k}\left((1-a_k)\,b_k\right)^{-(1-a_k)b_k}\prod_s b_{sk}^{-b_{sk}}\right](w_i^{a_k} r_i^{1-a_k})^{b_k}\prod_s P_{i,s}^{b_{sk}}.
\end{aligned}$$

The same algebra that we used to get (B.11), in turn, implies

$$\ln Z_{i,k} = \ln \hat{Z}_{i,k} + \zeta_{i,k} + \hat{E}_k,$$

with

$$\ln \hat{Z}_{i,k} \equiv \ln(S_{i,k}/w_i) - (1-a_k)b_k \ln(r_i/w_i) - \sum_s b_{sk} \ln(\hat{P}_{i,s}/w_i) - \ln\left[(a_k b_k)^{-a_k b_k}\left((1-a_k)\, b_k\right)^{-(1-a_k)b_k} \prod_s b_{sk}^{-b_{sk}}\right].$$

Combining (B.7) and (B.11), we obtain

$$\hat{Y}^{\mathrm{IO}}_{i,k} = \hat{\delta}_i + \hat{\delta}_k + \gamma_k \ln \hat{Z}_{i,k} + \hat{\varepsilon}_{i,k} \tag{B.12}$$

with

$$\begin{aligned}
\hat{\varepsilon}_{i,k} &\equiv \varepsilon_{i,k} + (\gamma_k+1)\zeta_{i,k} - \mathbb{E}[(\gamma_k+1)\zeta_{i,k}|i] - \mathbb{E}[(\gamma_k+1)\zeta_{i,k}|k] + \mathbb{E}[(\gamma_k+1)\zeta_{i,k}],\\
\hat{\delta}_i &\equiv \delta_i + \mathbb{E}[(\gamma_k+1)\zeta_{i,k}|i] - \mathbb{E}[(\gamma_k+1)\zeta_{i,k}],\\
\hat{\delta}_k &\equiv \delta_k + \mathbb{E}[(\gamma_k+1)\zeta_{i,k}|k] + (\gamma_k+1)\hat{E}_k.
\end{aligned}$$

Our estimation of γ_k then proceeds via the instrumental variable procedure described in Section 3.3. The only difference comes from the way we estimate the demand parameters $\hat{\beta}_{i,k}$ to construct demand-predicted sector size $\hat{L}_{i,k} \equiv \hat{\beta}_{i,k} \times \hat{L}_i$. Rather than using (26), which depends on the share $x_{i,k}$ of expenditure by country i on goods from sector k across all origins, we use

$$\beta_{i,k} = \frac{x^F_{i,k}/(P_{i,k})^{-\rho}}{\sum_l x^F_{i,l}/(P_{i,l})^{-\rho}},$$

which only depends on the share $x^F_{i,k}$ of final expenditure by country i on goods from sector k across all origins. The exclusion restriction then continues to be given by

$$\mathbb{E}[\ln \hat{L}_{i,k} \times \hat{\varepsilon}_{i,k}|k] = 0, \text{ for all } k,$$

with the error term $\hat{\varepsilon}_{i,k}$ in (B.12) now comprising the productivity of country i in both sector k and in other sectors (via the IO linkages due to non-tradable sectors that gives rise to $\zeta_{i,k}$).

OLS and IV estimates of the scale elasticity parameters γ_k from equation (B.7) are reported in Table B.8.

Table B.8: Estimates of Sector-Level Scale Elasticities (γ_k), Beyond Ricardian Economies

Sector	OLS (1)	IV (2)	First-stage F-stat (3)	SW F-stat (4)
Food, Beverages and Tobacco	0.23	0.17	34.6	214.3
	(0.01)	(0.03)		
Textiles	0.10	0.07	75.4	273.1
	(0.01)	(0.02)		
Wood Products	0.11	0.09	14.0	132.0
	(0.02)	(0.02)		
Paper Products	0.15	0.09	23.7	653.5
	(0.01)	(0.03)		
Coke/Petroleum Products	0.07	0.07	13.6	215.4
	(0.02)	(0.02)		
Chemicals	0.27	0.20	27.3	244.1
	(0.01)	(0.02)		
Rubber and Plastics	0.48	0.38	26.6	354.6
	(0.03)	(0.04)		
Mineral Products	0.16	0.12	15.4	314.4
	(0.02)	(0.02)		
Basic Metals	0.08	0.06	14.3	177.4
	(0.01)	(0.03)		
Fabricated Metals	0.11	0.07	35.6	299.1
	(0.02)	(0.02)		
Machinery and Equipment	0.28	0.19	23.8	286.9
	(0.02)	(0.02)		
Computers and Electronics	0.06	0.04	21.5	197.1
	(0.01)	(0.02)		
Electrical Machinery, NEC	0.06	0.04	38.3	521.2
	(0.01)	(0.02)		
Motor Vehicles	0.18	0.14	34.2	372.5
	(0.01)	(0.02)		
Other Transport Equipment	0.18	0.14	20.7	0.0
	(0.02)	(0.03)		

Notes: Column (1) reports the OLS estimate, and column (2) the IV estimate, of equation (B.7). The instruments are the log of (country population $\times$ sectoral demand shifter), interacted with sector indicators. Column (3) reports the conventional F-statistic, and column (4) the Sanderson-Windmeijer F-statistic, from the first-stage regression corresponding to the endogenous regressor formed by interacting the log of sector size with an indicator for the sector named in each row. All regressions include exporter and sector fixed effects. Standard errors in parentheses are clustered at the exporter level.

C Online Appendix: Counterfactuals

C.1 Construction of Counterfactual Equilibria

We first describe a competitive equilibrium as the solution to a system of non-linear equations. We then show how to use this system to conduct counterfactual and welfare analysis.

Competitive Equilibrium with Taxes and Subsidies. Starting from (1)-(9), we can describe a competitive equilibrium with employment subsidies, $\{s_{j,k}\}$, import tariffs, $\{t^m_{ij,k}\}$, export taxes, $\{t^x_{ij,k}\}$, as a set of sector sizes, $\{L_{i,k}\}$, within-sector expenditure shares, $\{x_{ij,k} \equiv X_{ij,k}/X_{j,k}\}$, between-sector expenditure shares, $\{x_{j,k} \equiv X_{j,k}/\sum_{l\in\mathcal{K}} X_{j,l}\}$, consumer prices, $\{p_{ij,k}\}$, price indices, $\{P_{j,k}\}$, wages, $\{w_j\}$, and lump-sum transfers, $\{T_j\}$, such that

$$
\begin{aligned}
w_i L_{i,k} &= \frac{1}{(1-s_{i,k})}\sum_j \frac{1-t^x_{ij,k}}{1+t^m_{ij,k}} x_{ij,k} x_{j,k}(w_j L_j + T_j + D_j),\\
x_{ij,k} &= \frac{p_{ij,k}^{-\theta_k}}{\sum_{i'} p_{i'j,k}^{-\theta_k}},\\
x_{j,k} &= \frac{\beta_{j,k} P_{j,k}^{-\rho}}{\sum_{k'} \beta_{j,k'} P_{j,k'}^{-\rho}},\\
p_{ij,k} &= \frac{(1+t^m_{ij,k})(1-s_{i,k})w_i}{(1-t^x_{ij,k})A_{ij,k}L_{i,k}^{\gamma_k}},\\
P_{j,k} &= (\sum_i p_{ij,k}^{-\theta_k})^{-1/\theta_k},\\
L_i &= \sum_k L_{i,k},\\
T_j &= \sum_{i,k} \frac{t^m_{ij,k}}{1+t^m_{ij,k}} x_{ij,k} x_{j,k}(w_j L_j + T_j + D_j) + \sum_{i,k}\left[t^x_{ji,k} - \frac{s_{j,k}(1-t^x_{ji,k})}{1-s_{j,k}}\right]\frac{x_{ji,k}}{1+t^m_{ji,k}} x_{i,k}(w_i L_i + T_i + D_i),
\end{aligned}
$$

where D_j denotes the trade deficit of country j, with $\sum_j D_j = 0$, and where we use the convention $t^m_{ii,k} = t^x_{ii,k} = 0$ for all i and k.

Counterfactual Changes. Suppose that the initial equilibrium has no taxes or subsidies. We are interested in a counterfactual equilibrium with taxes, subsidies, and transfers:

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

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 proportional change in this variable. We assume that D_js are fixed and do not change as we move to the counterfactual equilibrium. After simplifications, counterfactual changes are given by the solution to

$$\hat{w}_i \hat{L}_{i,k} S_{i,k} = \frac{1}{(1-s_{i,k})} \sum_j \frac{(1-t^x_{ij,k})\hat{x}_{ij,k}}{1+t^m_{ij,k}} \hat{x}_{j,k} \left(\frac{\hat{w}_j S_j + T_j + D_j}{S_j + D_j} \right) X_{ij,k}, \tag{C.1}$$

$$\hat{x}_{ij,k} = \frac{\hat{p}_{ij,k}^{-\theta_k}}{\sum_{i'} \hat{p}_{i'j,k}^{-\theta_k} x_{i'j,k}}, \tag{C.2}$$

$$\hat{x}_{j,k} = \frac{\hat{P}_{j,k}^{-\rho}}{\sum_{k'} \hat{P}_{j,k'}^{-\rho} x_{j,k'}}, \tag{C.3}$$

$$\hat{p}_{ij,k} = \frac{(1+t^m_{ij,k})(1-s_{i,k})\hat{w}_i}{(1-t^x_{ij,k})\hat{L}_{i,k}^{\gamma_k}}, \tag{C.4}$$

$$\hat{P}_{j,k} = (\sum_i \hat{p}_{ij,k}^{-\theta_k} x_{ij,k})^{-1/\theta_k}, \tag{C.5}$$

$$S_i = \sum_k \hat{L}_{i,k} S_{i,k}, \tag{C.6}$$

$$\begin{aligned} T_j = & \sum_{i,k} \frac{t^m_{ij,k} \hat{x}_{ij,k} \hat{x}_{j,k}}{1+t^m_{ij,k}} \left(\frac{\hat{w}_j S_j + T_j + D_j}{S_j + D_j} \right) X_{ij,k} \\ & + \sum_{i,k} \left[t^x_{ji,k} - \frac{s_{j,k}(1-t^x_{ji,k})}{1-s_{j,k}} \right] \frac{\hat{x}_{ji,k} \hat{x}_{i,k}}{1+t^m_{ji,k}} \left(\frac{\hat{w}_i S_i + T_i + D_i}{S_i + D_i} \right) X_{ji,k}, \end{aligned} \tag{C.7}$$

where bilateral trade flows $X_{ij,k}$, sectoral sales, $S_{i,k} \equiv \sum_j X_{ij,k}$, and total sales $S_i = \sum_{j,k} X_{ij,k}$ are all observed in the initial equilibrium. Once changes in the previous variables have been computed using (C.1)-(C.7), welfare changes are given by

$$\hat{U}_j = \frac{\hat{w}_j S_j + T_j + D_j}{S_j + D_j} \frac{1}{\hat{P}_j},$$

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

Under the assumption that country i_0 is small and that only country i_0 imposes trade taxes and employment subsidies, the system is as described above for country i_0; that is, equations (C.1)-(C.7) continue to hold if either i or j is equal to i_0. For all other countries, we set $\hat{w}_i = \hat{P}_{i,k} = \hat{L}_{i,k} = 1$ for all k and drop equations (C.1)-(C.7).[47]

[47]Since foreign wages are unaffected by changes in country i_0's policies, our assumption that D_js are fixed is equivalent to assuming that D_js are constant fractions of foreign GDP.

C.2 Gains from Optimal Policies

Table C.1: Gains from Optimal Policies

Country	Fully Optimal Policy (1)	Trade Policy (2)	Industrial Policy (3)
Argentina	1.14%	0.88%	0.93%
Australia	0.96%	0.54%	0.56%
Austria	1.23%	1.92%	1.07%
Belgium	1.33%	1.41%	0.99%
Brazil	0.89%	0.41%	0.72%
Brunei Darussalam	2.41%	0.91%	0.96%
Bulgaria	1.41%	1.77%	1.09%
Cambodia	1.87%	0.77%	0.88%
Canada	1.01%	1.31%	0.87%
Chile	1.32%	0.83%	0.87%
China	0.97%	0.43%	0.83%
Hong Kong	1.46%	1.12%	0.63%
Colombia	1.14%	0.44%	0.78%
Costa Rica	1.58%	0.88%	0.99%
Croatia	1.46%	0.79%	0.85%
Cyprus	1.95%	0.91%	0.93%
Czech Republic	1.26%	2.59%	1.27%
Denmark	1.33%	1.78%	0.99%
Estonia	1.57%	2.12%	1.11%
Finland	1.00%	1.68%	0.91%
France	0.97%	1.13%	0.83%
Germany	1.16%	1.84%	1.10%
Greece	1.12%	0.82%	0.72%
Hungary	1.37%	3.51%	1.37%
Iceland	1.24%	1.95%	0.99%
India	1.07%	0.56%	0.74%
Indonesia	1.28%	0.65%	0.89%
Ireland	1.86%	3.96%	1.77%
Israel	1.12%	1.66%	0.88%
Italy	0.94%	1.20%	0.85%
Japan	0.79%	0.56%	0.70%
Latvia	1.15%	1.21%	0.86%

Table C.1 (Continued): Gains from Optimal Policies

Country	Fully Optimal Policy (1)	Trade Policy (2)	Industrial Policy (3)
Lithuania	1.51%	1.60%	1.16%
Luxembourg	2.92%	3.21%	1.69%
Malaysia	1.64%	2.64%	1.21%
Malta	1.79%	1.29%	0.90%
Mexico	1.17%	1.60%	0.94%
Netherlands	1.10%	1.03%	0.84%
New Zealand	1.13%	0.95%	0.87%
Norway	1.64%	1.18%	0.89%
Philippines	1.20%	0.77%	0.76%
Poland	1.21%	1.67%	1.12%
Portugal	1.03%	1.31%	0.85%
Republic of Korea	1.19%	1.39%	0.96%
Romania	1.05%	1.19%	0.90%
Russian Federation	1.29%	0.76%	0.80%
Saudi Arabia	2.10%	2.02%	1.13%
Singapore	1.90%	2.32%	1.27%
Slovakia	1.18%	2.73%	1.25%
Slovenia	1.36%	2.56%	1.28%
South Africa	1.10%	0.95%	0.82%
Spain	1.11%	1.13%	0.89%
Sweden	1.19%	2.08%	1.03%
Switzerland	1.43%	3.05%	1.29%
Taiwan	1.37%	2.60%	1.10%
Thailand	1.44%	2.09%	1.25%
Tunisia	1.64%	1.75%	1.02%
Turkey	1.08%	0.67%	0.85%
United Kingdom	1.08%	1.31%	0.83%
United States	0.72%	0.50%	0.58%
Vietnam	1.75%	1.82%	1.13%
Avg., Unweighted	**1.34%**	**1.49%**	**0.98%**
Avg., GDP-weighted	**0.94%**	**0.87%**	**0.76%**

Notes: Column (1) reports the gains associated with fully optimal policies, as defined in equation (13); column (2) reports the gains associated with optimal trade policy, as defined in equation (15); and column (3) reports the gains associated with optimal industrial policy, as defined in equation (14). All gains are reported as a share of initial income.

C.3 Gains from Constrained and Globally Efficient Industrial Policies

Table C.2: Gains from Constrained and Globally Efficient Industrial Policies

Country	Baseline Industrial Policy (1)	Constrained Industrial Policy (2)	Globally Efficient Industrial Policy (3)
Argentina	0.93%	0.65%	0.24%
Australia	0.56%	0.59%	0.75%
Austria	1.07%	0.54%	0.00%
Belgium	0.99%	0.62%	0.00%
Brazil	0.72%	0.57%	0.59%
Brunei Darussalam	0.96%	1.98%	1.11%
Bulgaria	1.09%	0.65%	0.83%
Cambodia	0.88%	1.20%	1.79%
Canada	0.87%	0.48%	0.69%
Chile	0.87%	0.79%	0.61%
China	0.83%	0.65%	0.46%
Hong Kong	0.63%	1.01%	0.76%
Colombia	0.78%	0.78%	1.27%
Costa Rica	0.99%	0.89%	1.63%
Croatia	0.85%	0.95%	1.27%
Cyprus	0.93%	1.40%	2.64%
Czech Republic	1.27%	0.47%	-0.25%
Denmark	0.99%	0.70%	-0.18%
Estonia	1.11%	0.77%	0.76%
Finland	0.91%	0.42%	-0.27%
France	0.83%	0.47%	0.49%
Germany	1.10%	0.48%	-0.47%
Greece	0.72%	0.66%	1.31%
Hungary	1.37%	0.49%	-0.37%
Iceland	0.99%	0.61%	-0.22%
India	0.74%	0.66%	0.64%
Indonesia	0.89%	0.79%	0.77%
Ireland	1.77%	1.01%	-2.13%
Israel	0.88%	0.56%	0.15%
Italy	0.85%	0.42%	0.20%
Japan	0.70%	0.50%	0.14%
Latvia	0.86%	0.58%	1.20%

Table C.2 (Continued): Gains from Constrained and Globally Efficient Industrial Policies

Country	Baseline Industrial Policy (1)	Constrained Industrial Policy (2)	Globally Efficient Industrial Policy (3)
Lithuania	1.16%	0.77%	0.78%
Luxembourg	1.69%	1.72%	0.66%
Malaysia	1.21%	0.75%	-0.35%
Malta	0.90%	1.02%	1.15%
Mexico	0.94%	0.65%	1.16%
Netherlands	0.84%	0.54%	-0.28%
New Zealand	0.87%	0.61%	0.28%
Norway	0.89%	1.02%	0.68%
Philippines	0.76%	0.66%	0.63%
Poland	1.12%	0.54%	0.51%
Portugal	0.85%	0.48%	0.88%
Republic of Korea	0.96%	0.56%	-0.23%
Romania	0.90%	0.54%	0.95%
Russian Federation	0.80%	0.90%	0.74%
Saudi Arabia	1.13%	1.59%	0.86%
Singapore	1.27%	0.98%	-1.11%
Slovakia	1.25%	0.38%	0.13%
Slovenia	1.28%	0.56%	0.13%
South Africa	0.82%	0.63%	0.77%
Spain	0.89%	0.55%	0.51%
Sweden	1.03%	0.49%	-0.33%
Switzerland	1.29%	0.73%	-0.31%
Taiwan	1.10%	0.52%	-0.07%
Thailand	1.25%	0.60%	-0.50%
Tunisia	1.02%	0.91%	1.58%
Turkey	0.85%	0.65%	0.98%
United Kingdom	0.83%	0.56%	0.66%
United States	0.58%	0.41%	0.55%
Vietnam	1.13%	1.00%	1.87%
Avg., Unweighted	**0.98%**	**0.73%**	**0.49%**
Avg., GDP-Weighted	**0.76%**	**0.53%**	**0.35%**

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

C.4 Sensitivity to Calibrated Parameters

Table C.3: Gains from Industrial Policy, Alternative Trade Elasticities (θ_k)

Country	Baseline (1)	CP (2)	Shapiro (3)	GYY (4)
United States	0.58%	2.46%	0.76%	0.63%
China	0.83%	2.96%	1.09%	0.79%
Germany	1.10%	5.09%	0.38%	1.11%
Ireland	1.78%	3.54%	2.61%	2.25%
Vietnam	1.13%	5.35%	1.60%	1.22%
Avg., Unweighted	**0.98%**	**3.28%**	**1.13%**	**1.14%**
Avg., GDP-Weighted	**0.76%**	**2.92%**	**0.80%**	**0.84%**
Corr. with Baseline	**-**	**0.38**	**0.71**	**0.85**

Notes: Each column reports the gains from industrial policy, as defined in equation and (14), for the corresponding trade elasticity estimates reported in Table B.3. Column (1) replicates the baseline calculation, which uses the median trade elasticity in each sector across the studies in columns (2)-(4). Columns (2)-(4) reports a similar calculation using trade elasticity estimates from each of these studies (Caliendo and Parro, 2015; Shapiro, 2016; and Giri et al., 2021) individually. The last row reports the correlation, across countries, of the gains from industrial policy with those in column (1).

C.5 Beyond Ricardian Economies

C.5.1 Construction of Counterfactuals

Competitive Equilibrium with Taxes and Subsidies. Starting from equations (2), (3), (6), (36), (37), and (A.46)-(A.52), we can describe a competitive equilibrium with production subsidies, $\{s_{j,k}\}$, import tariffs, $\{t^m_{ij,k}\}$, and export taxes, $\{t^x_{ij,k}\}$, as a set of sector sizes, $\{L_{i,k}, K_{i,k}, Z_{i,k}\}$, within-sector expenditure shares, $\{x_{ij,k}\}$, between-sector final expenditure shares, $\{x^F_{j,k}\}$, between-sector expenditures, $\{X_{j,k}\}$, consumer prices, $\{p_{ij,k}\}$, sector price indices, $\{P_{j,k}\}$, wages, $\{w_j\}$, rental rates, $\{r_j\}$, and lump-sum transfers, $\{T_j\}$, such that

$$
\begin{aligned}
w_{i,k}Z_{i,k} &= \frac{1}{(1-s_{i,k})}\sum_j \frac{1-t^x_{ij,k}}{1+t^m_{ij,k}} x_{ij,k}X_{j,k},\\
w_i L_{i,k} &= a_k b_k w_{i,k} Z_{i,k}\\
r_i K_{i,k} &= (1-a_k) b_k w_{i,k} Z_{i,k}\\
x_{ij,k} &= \frac{p_{ij,k}^{-\theta_k}}{\sum_{i'} p_{i'j,k}^{-\theta_k}},\\
x^F_{j,k} &= \frac{\beta_{j,k} P_{j,k}^{-\rho}}{\sum_{k'} \beta_{j,k'} P_{j,k'}^{-\rho}},\\
X_{j,k} &= x^F_{j,k}\left(w_j L_j + r_j K_j + T_j + D_j\right) + \sum_{i,s} b_{ks} \frac{1-t^x_{ji,s}}{1+t^m_{ji,s}} x_{ji,s} X_{i,s},\\
p_{ij,k} &= \frac{(1+t^m_{ij,k})(1-s_{i,k})w_{i,k}}{(1-t^x_{ij,k}) A_{ij,k} Z_{i,k}^{\gamma_k}},\\
P_{j,k} &= (\sum_i p_{ij,k}^{-\theta_k})^{-1/\theta_k},\\
w_{i,k} &= \left[(a_k b_k)^{-a_k b_k}\left((1-a_k)\, b_k\right)^{-(1-a_k) b_k} \prod_s b_{sk}^{-b_{sk}}\right] (w_i^{a_k} r_i^{1-a_k})^{b_k} \prod_s P_{i,s}^{b_{sk}},\\
L_i &= \sum_k L_{i,k},\\
K_i &= \sum_k K_{i,k},\\
T_j &= \sum_{i,k} \frac{t^m_{ij,k}}{1+t^m_{ij,k}} x_{ij,k} X_{j,k} + \sum_{i,k}\left[t^x_{ji,k} - \frac{s_{j,k}(1-t^x_{ji,k})}{1-s_{j,k}}\right] x_{ji,k} X_{i,k}.
\end{aligned}
$$

Counterfactual Changes. The counterfactual changes associated with moving from laissez-faire to a counterfactual equilibrium with employment subsidies, export taxes, and import tariffs are given by the solution to

$$\hat{w}_{i,k}\hat{Z}_{i,k}S_{i,k} = \frac{1}{(1-s_{i,k})}\sum_j \frac{(1-t^x_{ij,k})\hat{x}_{ij,k}}{1+t^m_{ij,k}} x_{ij,k}X'_{j,k}, \tag{C.8}$$

$$\hat{w}_i\hat{L}_{i,k} = \hat{w}_{i,k}\hat{Z}_{i,k}, \tag{C.9}$$

$$\hat{r}_i\hat{K}_{i,k} = \hat{w}_{i,k}\hat{Z}_{i,k}, \tag{C.10}$$

$$\hat{x}_{ij,k} = \frac{\hat{p}_{ij,k}^{-\theta_k}}{\sum_l \hat{p}_{lj,k}^{-\theta_k} x_{lj,k}}, \tag{C.11}$$

$$\hat{x}^F_{j,k} = \frac{\hat{P}_{j,k}^{-\rho}}{\sum_s \hat{P}_{j,s}^{-\rho} x^F_{j,s}}, \tag{C.12}$$

$$X'_{j,k} = \hat{x}^F_{j,k}x^F_{j,k}\left(\hat{w}_j\sum_k a_k b_k S_{j,k} + \hat{r}_j\sum_k(1-a_k)b_k S_{j,k} + T_j + D_j\right) + \sum_{i,s} b_{ks}\frac{1-t^x_{ji,s}}{1+t^m_{ji,s}}\hat{x}_{ji,s}x_{ji,s}X'_{i,s}, \tag{C.13}$$

$$\hat{p}_{ij,k} = \frac{(1+t^m_{ij,k})(1-s_{i,k})\hat{w}_{i,k}}{(1-t^x_{ij,k})\hat{Z}_{i,k}^{\gamma_k}}, \tag{C.14}$$

$$\hat{P}_{j,k} = (\sum_i \hat{p}_{ij,k}^{-\theta_k} x_{ij,k})^{-1/\theta_k}, \tag{C.15}$$

$$\hat{w}_{i,k} = (\hat{w}_i^{a_k}\hat{r}_i^{1-a_k})^{b_k}\prod_s \hat{P}_{i,s}^{b_{sk}} \tag{C.16}$$

$$\sum_k a_k b_k S_{i,k} = \sum_k \hat{L}_{i,k}a_k b_k S_{i,k}, \tag{C.17}$$

$$\sum_k (1-a_k)b_k S_{i,k} = \sum_k \hat{K}_{i,k}(1-a_k)b_k S_{i,k}, \tag{C.18}$$

$$T_j = \sum_{i,k}\frac{t^m_{ij,k}\hat{x}_{ij,k}x_{ij,k}}{1+t^m_{ij,k}}X'_{j,k} + \sum_{i,k}\left[t^x_{ji,k} - \frac{s_{j,k}(1-t^x_{ji,k})}{1-s_{j,k}}\right]\hat{x}_{ji,k}x_{ji,k}X'_{i,k}. \tag{C.19}$$

We solve this system of equations using trade data as in the baseline analysis, augmented with data on factor and input shares from the OECD ICIO and WIOD datasets, as described in Appendix B.5.1, for 2010.[48] Once changes in the previous variables have been computed using (C.8)-(C.19), counterfactual welfare changes are given by

$$\hat{U}_j = \frac{\hat{w}_j\sum_k a_k b_k S_{j,k} + \hat{r}_j\sum_k(1-a_k)b_k S_{j,k} + T_j + D_j}{\sum_k b_k S_{j,k} + D_j}\frac{1}{\hat{P}_j},$$

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

[48]In practice, with factor and input shares set to their global levels, a small number of country-sectors have implied domestic consumption that is negative. In such cases we increase domestic consumption entries to zero and recalculate the resulting global factor and input shares.

Under the assumption that country i_0 is small and that only country i_0 imposes trade taxes and employment subsidies, we follow the same procedure as in Appendix C.1 and impose $\hat{w}_i = \hat{r}_i = \hat{P}_{i,k} = \hat{L}_{i,k} = \hat{K}_{i,k} = 1$ for all k and $i \neq i_0$.

C.5.2 Gains from Optimal Policies

Table C.4: Gains from Optimal Policies, Beyond Ricardian Economies

Country	Fully Optimal Policy (1)	Trade Policy (2)	Industrial Policy (3)
Argentina	3.33%	5.23%	3.39%
Australia	2.35%	1.22%	1.83%
Austria	2.97%	12.92%	3.53%
Belgium	3.44%	10.09%	3.59%
Brazil	2.65%	1.15%	2.50%
Brunei Darussalam	4.43%	1.40%	2.43%
Bulgaria	2.62%	9.54%	2.67%
Cambodia	3.96%	1.97%	2.81%
Canada	2.30%	8.62%	2.55%
Chile	3.11%	2.52%	2.89%
China	3.48%	1.44%	3.44%
Hong Kong	2.81%	1.99%	1.69%
Colombia	2.71%	1.63%	2.27%
Costa Rica	3.07%	2.35%	2.60%
Croatia	3.01%	2.71%	2.43%
Cyprus	2.88%	2.22%	1.67%
Czech Republic	3.28%	16.99%	4.61%
Denmark	3.04%	7.28%	3.01%
Estonia	2.85%	13.33%	2.56%
Finland	2.44%	7.10%	2.90%
France	2.29%	5.52%	2.67%
Germany	3.44%	10.73%	4.59%
Greece	1.92%	2.81%	1.59%
Hungary	3.20%	21.62%	4.31%
Iceland	2.85%	3.29%	2.66%
India	2.86%	1.98%	2.63%
Indonesia	3.30%	2.10%	2.83%
Ireland	4.40%	17.84%	5.28%
Israel	2.37%	7.85%	2.69%
Italy	2.30%	6.44%	2.78%
Japan	2.47%	2.58%	2.82%
Latvia	2.07%	4.02%	1.87%

Table C.4 (Continued): Gains from Optimal Policies, Beyond Ricardian Economies

Country	Fully Optimal Policy (1)	Trade Policy (2)	Industrial Policy (3)
Lithuania	3.38%	8.31%	3.46%
Luxembourg	4.81%	43.29%	3.36%
Malaysia	4.49%	17.47%	4.94%
Malta	3.95%	3.90%	2.64%
Mexico	2.75%	8.28%	2.80%
Netherlands	2.93%	5.24%	2.85%
New Zealand	2.51%	2.36%	2.37%
Norway	3.84%	4.04%	3.09%
Philippines	3.17%	5.65%	2.90%
Poland	2.92%	11.73%	3.55%
Portugal	2.03%	9.36%	2.19%
Republic of Korea	4.04%	8.93%	4.82%
Romania	2.30%	7.77%	2.63%
Russian Federation	3.24%	2.25%	2.66%
Saudi Arabia	3.77%	8.44%	3.09%
Singapore	5.48%	10.00%	5.26%
Slovakia	2.69%	17.00%	3.73%
Slovenia	2.99%	17.77%	3.86%
South Africa	2.64%	5.42%	2.62%
Spain	2.72%	8.45%	3.05%
Sweden	3.34%	11.71%	4.04%
Switzerland	3.12%	11.78%	3.55%
Taiwan	3.98%	11.35%	4.54%
Thailand	4.11%	14.91%	5.31%
Tunisia	3.67%	7.10%	3.34%
Turkey	3.02%	3.50%	3.06%
United Kingdom	2.39%	7.07%	2.48%
United States	1.82%	1.92%	1.86%
Vietnam	3.74%	5.29%	3.34%
Avg., Unweighted	**3.12%**	**7.85%**	**3.11%**
Avg., GDP-weighted	**2.57%**	**4.34%**	**2.71%**

Notes: Column (1) reports the gains associated with fully optimal policies, as defined in equation (13); column (2) reports the gains associated with optimal trade policy, as defined in equation (15); and column (3) reports the gains associated with optimal industrial policy, as defined in equation (14). All gains are reported as a share of initial income.

C.5.3 Comparison Between Exact Gains and Harberger Triangles

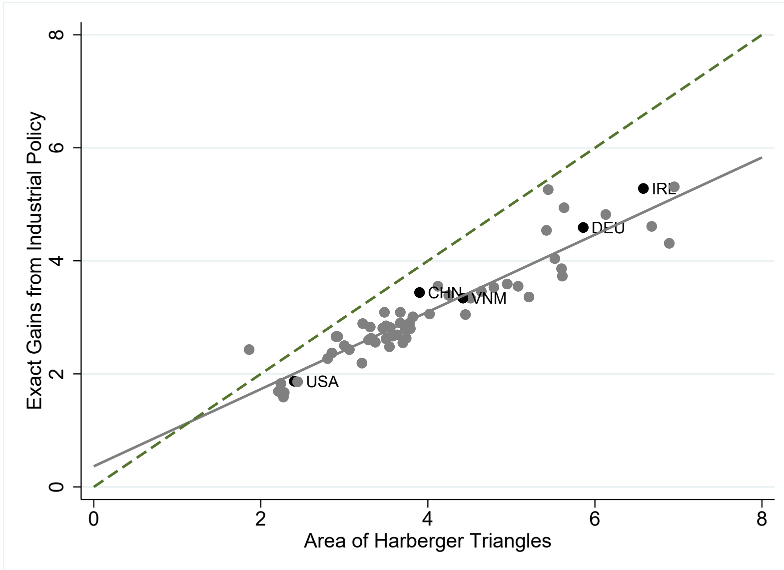
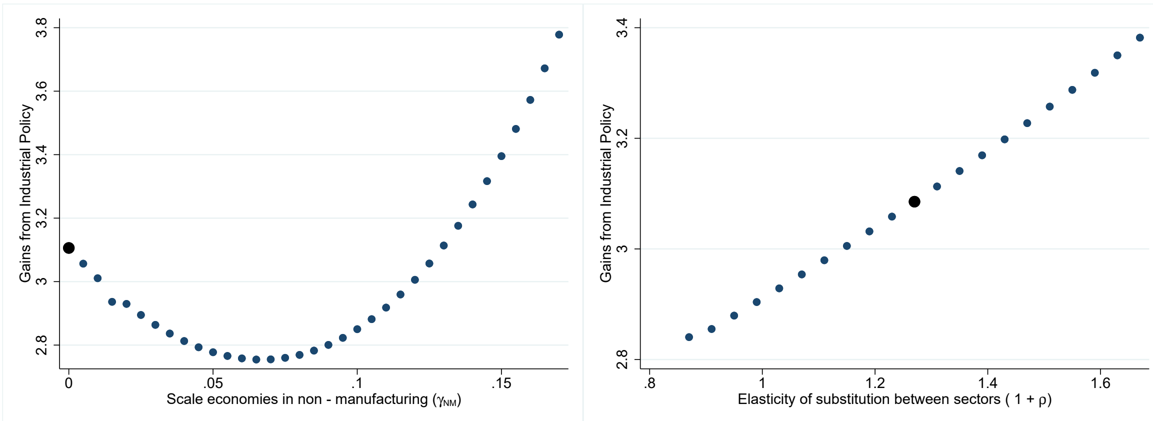


Figure C.1: Exact Gains from Industrial Policy versus Areas of Harberger Triangles, Beyond Ricardian Economies

Notes: Figure C.1 reports the exact gains from industrial policy on the y-axis, as defined in equation (14), against the second-order approximation from Proposition 4 on the x-axis, as described in equation (42). All gains are computed under the assumptions of Section 6.1 and reported as a share of initial income.

C.5.4 Sensitivity to Calibrated Parameters



(a) Scale Elasticity Outside Manufacturing **(b) Elasticity of Substitution Between Sectors**

Figure C.2: Gains from Industrial Policy, Alternative Parameter Values, Beyond Ricardian Economies

Notes: Figure C.2a reports the gains from industrial policy, as defined in equation (14), , computed under the assumptions of Section 6.1 for different values of the scale elasticity outside manufacturing, γ_{NM}. Figure C.2b does the same for different values of the elasticity of substitution between sectors, $1 + \rho$.

Table C.5: Gains from Industrial Policy, Alternative Trade Elasticities (θ_k), Beyond Ricardian Economies

Country	Baseline (1)	CP (2)	Shapiro (3)	GYY (4)
United States	1.86%	5.41%	3.15%	2.72%
China	3.44%	11.05%	5.76%	4.77%
Germany	4.59%	15.66%	6.79%	5.24%
Ireland	5.28%	6.14%	11.23%	9.53%
Vietnam	3.34%	5.96%	6.24%	4.35%
Avg., Unweighted	**3.11%**	**6.90%**	**5.35%**	**4.23%**
Avg., GDP-Weighted	**2.71%**	**7.75%**	**4.53%**	**3.63%**
Corr. with Baseline	**-**	**0.73**	**0.85**	**0.83**

Notes: Each column reports the gains from industrial policy, as defined in equation and (42), for the corresponding trade elasticity estimates reported in Table B.3. Column (1) replicates the baseline calculation, which uses the median trade elasticity in each sector across the studies in columns (2)-(4). Columns (2)-(4) reports a similar calculation using trade elasticity estimates from each of these studies (Caliendo and Parro, 2015; Shapiro, 2016; and Giri et al., 2021) individually. In the case of the elasticities from Caliendo and Parro, 2015, we found that the estimate in the "Other Transportation Equipment" was a clear outlier and lead to numerical instability in the algorithm computing counterfactuals. We replaced it with the next highest estimate ("Machinery and Equipment") when computing the policy counterfactuals reported above. The last row reports the correlation, across countries, of the gains from industrial policy with those in column (1).