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

Dominick Bartelme
University of Michigan

Arnaud Costinot
MIT

Dave Donaldson
MIT

Andres Rodriguez-Clare
UC Berkeley

July 2018 – PRELIMINARY AND INCOMPLETE

Abstract

We develop a new empirical strategy to estimate external economies of scale using trade data. Across 2-digit manufacturing sectors, our baseline estimates of scale elasticities range from 0 to 0.23 and average 0.06. We then use our estimates of external economies of scale across sectors to explore the structure and implications of optimal industrial policy. We find that gains from optimal industrial policy are only around 0.1% on average across the countries in our sample – this is small relative to the gains from optimal trade policy, which are around 0.6% on average.

*Contact: dbartelm@umich.edu, costinot@mit.edu, ddonald@mit.edu, andres@econ.berkeley.edu. We are grateful to Susantu Basu, Jonathan Eaton and David Weinstein for comments, and to Daniel Haanwinckel and Walker Ray for outstanding research assistance.

1 Introduction

When sector size goes up, does productivity go up as well? If it does, are productivity gains larger in some sectors than others? And if they are, what are the gains from implementing industrial policies that subsidize these sectors at the expense of others?

In this paper, we develop a new empirical strategy to estimate sector-level economies of scale and address these questions. Our main findings are as follows. First, we find clear evidence of sector-level economies of scale in manufacturing. Our baseline estimate implies that a 10% increase in sector size leads to a (quality adjusted) productivity increase of 0.06%. Second, across 2-digit manufacturing sectors, there is significant heterogeneity in the elasticity of (quality adjusted) productivity with respect to size. It ranges from 0 in the Food, Beverages and Tobacco, Texties, and Wood Products sectors, to 0.23 in the Motor Vehicles sector. Third, through the lens of a standard model of international trade, the gains from industrial policy are quite small. In our preferred calibration, a small open economy would gain on average around 0.1% of GDP by adopting the optimal industrial policy. In comparison, optimal trade policy would lead to a gain of 0.6% on average.

Section 2 describes the economic environment within which we propose to study sector-level economies of scale. We focus on a Ricardian economy with multiple sectors. Within each sector, external economies of scale may affect both the productivity of firms and the quality of the goods they produce. In this environment, we show that external economies of scale are non-parametrically identified from commonly available trade data and standard orthogonality conditions.

The starting point of our approach is the observation that in each destination and within each sector, trade flows from different origins reflect the optimal demand for inputs from these countries. Provided that demand system is invertible, changes in trade flows therefore reveal changes in the effective prices of these inputs, that is, prices after controlling for productivity and quality differences. Once input prices have been revealed, we can estimate external economies of scale by measuring the extent to which an exogenous increase in sector size lowers such prices.

Section 3 implements the previous strategy in the context of a multi-sector gravity model of trade, as in Chor (2010), Costinot et al. (2012), Caliendo and Parro (2015), and Levchenko and Zhang (2016). In this case, the demand for inputs from different countries has Constant Elasticity of Substitution (CES). We proceed in three steps.

The first step is to estimate revenue scale elasticities, defined as the elasticity of sector-level bilateral exports with respect to sector size, controlling for origin and destination-sector fixed effects. In the model, the revenue scale elasticity is the product of the scale elasticity and the trade elasticity. Our strategy to estimate such elasticities is to assume that there are no scale economies in agriculture. We can then estimate revenue scale elasticities as the effect of sector size on sector-level exports relative to agriculture in an otherwise standard gravity regression. Since such relative exports will also be affected by exogeneous comparative advantage, identification requires an instrument that is positively correlated with sector size yet uncorrelated with idiosyncratic productivity shocks. We use population as our instrument. Not surprisingly, this leads to a very strong first stage, since larger countries have larger manufacturing sectors. Our IV approach yields consistent estimates under the assumption that country size does not affect Ricardian comparative advantage – that is the exclusion restriction that we make. Reassuringly, in every sector the IV estimate is lower than the OLS estimate, as would be expected if sector size responds positively to productivity.

The second step is to estimate sector-level trade elasticities. As in Caliendo and Parro (2015), we estimate trade elasticities using the standard gravity equation using tariff data. We exploit variation in tariffs across country pairs coming from the inclusion of domestic sales and from trade agreements and tariff preferences conferred by developed countries to poor countries. Our estimates range from 1.5 to 8, with an average of 5.1 across manufacturing sectors.

The third and final step is to combine the previous estimates to infer scale elasticities. Our findings point to positive and significant scale elasticities in manufacturing sectors. As mentioned above, these range from 0 to 0.23, with an average of 0.06.

Section 4 uses our empirical estimates to characterize optimal industrial policy as well as the gains from implementing such policy. We do so in the context of a small open economy that can affect employment in each of its industries, but not in the rest of the world. The rest of the economic environment is the same as in Kucheryavyy et al. (2017). Differences in external economies of scale across sectors lead to different optimal subsidies for the standard Pigouvian reason. Although external economies of scale are large, gains from industrial policy are only 0.1% for the average country. This is very similar to the gains from optimal trade policy, which are on average 0.4%.

Our analysis is related to a large empirical literature on the estimation of production functions in industrial organization and macroeconomics, see Ackerberg et al. (2007) and Basu (2008). Compared to the former, we make no attempt at estimating internal economies of scale at the firm-level. Rather, we focus on external economies at the sector-level, which sector-level trade flows reveal. Our focus on economies of scale at the sector-level is closer in spirit to Caballero and Lyons (1992) and Basu and Fernald (1997). A key difference between our approach and theirs is that we do not rely on measures of real output, or price indices, collected by statistical agencies. Instead, we use estimates of the demand for foreign inputs, as in Adao et al. (2017), to infer the effective prices for inputs. This provides a theoretically-grounded way to adjust for quality differences across origins within the same sector. We come back to these issues in Section 2.3.

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

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

Finally, while a number of theoretical and empirical papers have discussed the rationale and potential consequences of industrial policy, as Harrison and Rodríguez-Clare (2010) discuss in their review of the literature, we are not aware of any paper trying to connect theory and data in order to estimate the gains from optimal industrial policy. We hope to fill this gap.

2 Theory

2.1 Economic Environment

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

Technology Technology is Ricardian. In any origin country i, the same composite input is used to produce all goods in all sectors.[1] Output of good ω in sector k is given by

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

where $l_{i,k}(\omega)$ denotes the amount of the composite input used by producers of good ω.[2]At the sector-level, however, production may be subject to economies of scale,

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

where $L_{i,k}$ is the total amount of the composite input used in country i and sector k. For expositional purposes, we shall simply refer to $L_{i,k}$ as sector size.

Preferences Preferences are weakly separable over goods from different sectors. In any destination country j, the subutility associated with goods from sector k is given by

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

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

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

Trade Frictions Trade is subject to iceberg trade costs. In order to sell one unit of a good from country i to country j in sector k, a firm must ship $\tau_{ij,k}$ units.

Competitive Equilibrium We focus on a perfectly competitive equilibrium with external economies of scale where the size of each sector, $L_{i,k}$, is taken as given by profit-maximizing firms and utility-maximizing consumers. For any origin country i, any des-

[1]This rules out cross-sectoral differences in either factor intensity or input-output linkages. We will explore those those in future draft.

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

tination country j, any sector k, and any good ω, profit maximization determines supply,

$$q_{ij,k}(\omega) \in \text{argmax}_{\tilde{q}_{ij,k}(\omega)}[p_{ij,k}(\omega) - (w_i \tau_{ij,k})/(\alpha_{i,k}(\omega) A_{i,k} E_k^A(L_{i,k}))]\tilde{q}_{ij,k}(\omega). \tag{1}$$

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

$$\{q_{ij,k}(\omega)\} \in \text{argmax}_{\{\tilde{q}_{ij,k}(\omega)\}}\{U_{j,k}(\{\beta_{ij,k}(\omega) B_{i,k} E_k^B(L_{i,k})\tilde{q}_{ij,k}(\omega)\})| \sum_{i,\omega} p_{ij,k}(\omega)\tilde{q}_{ij,k}(\omega) = X_{j,k}\}, \tag{2}$$

where $X_{j,k}$ denotes total expenditure in country j on goods from sector k.

Equations (1) and (2) are the two equilibrium conditions that we will use to establish the non-parametric identification of external economies of scale in the next subsection, taking as given input prices, $\{w_i\}$, sector sizes, $\{L_{i,k}\}$, and sector-level expenditures, $\{X_{j,k}\}$. In a competitive equilibrium, these variables are, in turn, determined by factor market clearing and the upper-level of the consumers' utility maximization problem. We will return to these conditions in Sections 3 and 4.

2.2 Identification of External Economies of Scale

Let $x_{ij,k} \equiv \int_\omega p_{ij,k}(\omega) q_{ij,k}(\omega)/X_{j,k}$ denote the share of expenditure in destination j on goods from sector k produced in country i. For expositional purposes, we shall simply refer to $\{x_{ij,k}\}$ as trade shares. As shown in Appendix A.2, trade shares in a perfectly competitive equilibrium satisfy

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

with

$$c_{ij,k} \equiv \frac{\eta_{ij,k} w_i}{E_k(L_{i,k})},$$

where $\chi_{j,k} \equiv (\chi_{1j,k}, ..., \chi_{Ij,k})$ is homogeneous of degree zero, invertible, and a function of, and only of, $U_{j,k}$, $\{\alpha_{i,k}(\omega)\}$ and $\{\beta_{ij,k}(\omega)\}$; $\eta_{ij,k} \equiv \tau_{ij,k}/ (A_{i,k} B_{i,k})$ captures trade costs adjusted by exogenous productivity and quality in the origin country; and $E_k(L_{i,k}) \equiv E_k^A(L_{i,k}) E_k^B(L_{i,k})$ captures the joint effect of external economies of scale on the supply and demand sides.

$\chi_{ij,k}$ can be interpreted as the demand for inputs from country i in country j, within a given sector k, and $c_{ij,k}$ as the effective price of such input. This is the sector-level counterpart of factor demand in Adao et al. (2017). Under the assumption that $U_{j,k}$ satisfies the connected substitutes property, $\chi_{ij,k}$ is invertible and non-parametrically identified under standard orthogonality conditions, as also discussed in Adao et al. (2017).

Our goal here is to provide conditions under which, given knowledge of $\chi_{ij,k}$, F_k is non-parametrically identified as well. The basic idea is to start by inverting demand in order to go from the trade shares, that are observed, to the effective input prices, that are not. Once the prices having been inferred, we can then estimate external economies of scale by measuring the extent to which an exogenous increase in sector size lowers such prices.

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

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

Taking a second difference relative to another sector k_2, we therefore have

$$\begin{aligned} &\ln \frac{\chi_{i_1 j,k_1}^{-1}(x_{1j,k_1},...,x_{Ij,k_1})}{\chi_{i_2 j,k_1}^{-1}(x_{1j,k_1},...,x_{Ij,k_1})} - \ln \frac{\chi_{i_1 j,k_2}^{-1}(x_{1j,k_2},...,x_{Ij,k_2})}{\chi_{i_2 j,k_2}^{-1}(x_{1j,k_2},...,x_{Ij,k_2})} \\ &\qquad = \ln \frac{E_{k_1}(L_{i_2,k_1})}{E_{k_1}(L_{i_1,k_1})} - \ln \frac{E_{k_2}(L_{i_2,k_2})}{E_{k_2}(L_{i_1,k_2})} + \ln \frac{\eta_{i_1 j,k_1}}{\eta_{i_2 j,k_1}} - \ln \frac{\eta_{i_1 j,k_2}}{\eta_{i_2 j,k_2}}. \end{aligned} \tag{4}$$

Given two origin countries, i_1 and i_2, two sectors, k_1 and k_2, and a destination country j, equation (4) is a nonparametric regression model with endogenous regressors and a linear error term,

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

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

$$\begin{aligned} y &\equiv \ln \frac{\chi_{i_1 j,k_1}^{-1}(x_{1j,k_1},...,x_{Ij,k_1})}{\chi_{i_2 j,k_1}^{-1}(x_{1j,k_1},...,x_{Ij,k_1})} - \ln \frac{\chi_{i_1 j,k_2}^{-1}(x_{1j,k_2},...,x_{Ij,k_2})}{\chi_{i_2 j,k_2}^{-1}(x_{1j,k_2},...,x_{Ij,k_2})}, \\ l &\equiv (L_{i_1,k_1}, L_{i_2,k_1}, L_{i_1,k_2}, L_{i_2,k_2}), \\ h(l) &\equiv \ln \frac{E_{k_1}(L_{i_2,k_1})}{E_{k_1}(L_{i_1,k_1})} - \ln \frac{E_{k_2}(L_{i_2,k_2})}{E_{k_2}(L_{i_1,k_2})}, \\ \epsilon &\equiv \ln \frac{\eta_{i_1 j,k_1}}{\eta_{i_2 j,k_1}} - \ln \frac{\eta_{i_1 j,k_2}}{\eta_{i_2 j,k_2}}. \end{aligned}$$

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

Newey and Powell (2003) provide general conditions for nonparametric identification in such environments. Specifically, if there exists a vector of instruments z that satisfies the exclusion restriction, $E[\epsilon|z] = 0$, as well as the completeness condition, $E[g(l)|z] = 0$ implies $g = 0$ for any g with finite expectation, then $h(\cdot)$ is nonparametrically identified. As shown in Appendix A.3, once $h(\cdot)$ is identified, both E_{k_1} and E_{k_2} are also identified, up to a normalization. In the next section, we will propose such a vector of instruments and use it to estimate sector-level external economies of scale.

2.3 Discussion

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

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

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

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

Perfect versus Imperfect Competition The estimation of production functions, either using micro or macro data, does not require any assumption on good market structure. With output data and exogenous variation in input use, one can directly estimate the elasticity of output with respect to input, and hence economies of scale, regardless of whether good markets are perfectly competitive or not. In contrast, the nonparametric identification of external economies of scale in Section 2.2 is conducted under the assumption of perfect competition. Under this assumption, prices are equal to unit costs. This allows us to infer how variation in sector sizes affects costs, and hence economies of scale, by estimating how the variation in sector sizes affects prices, as revealed by trade shares.

The previous discussion might suggest that perfect competition is critical for our empirical strategy. In an environment where the pass-through from costs into prices is incomplete, as documented in a large international macro literature, one might expect our approach to systematically misinterpret changes in markups as changes in costs. Fortunately, this is not the case.

This is best seen through an extreme example. Consider an economy where production is as described in Section 2.1, but there is now an imperfectly competitive retail sector that buys goods at marginal costs and sell them at a profit. We assume that retailers take sector-level expenditure as given. In this more general environment, retailers will impose different markups on different goods, $p_{ij,k}(\omega) = \mu_{ij,k}(\omega) w_i / \left(\alpha_{i,k}(\omega) A_{i,k} E_k^A(L_{i,k})\right)$. However, as we formally demonstrate in the Appendix A.4, markups in sector k and country j will remain a function of $(c_{1j,k}, ..., c_{Ij,k})$, and hence we can still express trade shares as a function of input prices, $\chi_{ij,k}(c_{1j,k}, ..., c_{Ij,k})$. Thus, given knowledge of χ, external economies are nonparametrically identified under the same condition as under perfect competition. The reason why the lack of market power by firms is not critical for our empirical strategy can be understood as follows. If we have access to an observable exogenous shifter of $c_{ij,k}$, such as freight costs or tariffs, which is what the knowledge of χ requires, then one can compare the elasticity of trade shares with respect to this observable cost shifter to the elasticity of trade shares with respect to sector size. The ratio of the latter to the former then identifies by how much sector sizes has affected costs, i.e. the extent of economies of scale. Whether or not good prices are equal to their marginal costs, the exact same inference remains valid.[3]

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

By using micro data, one could estimate these two functions sector by sector. Specifically, one could first use data on firm's physical output and input use to estimate firm-level production functions, $F(l, \phi)$. Given such estimates, one could then infer $E_k^A(L_{i,k})$ by

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

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

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

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

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

Absent free entry, good-level production functions may no longer be constant returns

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

and economies of scale estimated at the sector-level may therefore reflect a mixture of both internal and external economies of scale. To control for internal economies of scale, without assuming that they necessarily vanish at the good level, one would need micro data. This is the same issue that one faces when estimating sector-level production functions.

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

3 Estimation

3.1 Parametric Restrictions

The results presented in Section 2.2 are asymptotic in nature. They answer the question of whether, in theory, one could point-identify external economies of scale, E_k, in all sectors with a dataset that includes an infinite sequence of economies. In this context, we have established that given an exogenous shifter of sector sizes, one can identify external economies of scale by tracing out the impact of changes in sector sizes on prices, as revealed by changes in equilibrium trade shares.

In practice, our datasets only include a small number of observations. As we discuss below, our datasets only includes 4 time periods and 61 countries. So, our estimation must proceed parametrically. In the rest of our analysis, we impose the following functional-form assumptions:

$$\chi_{ij,k}(c_{1j,k}, ..., c_{Ij,k}) = \frac{c_{ij,k}^{-\theta_k}}{\sum_{i'} c_{i'j,k}^{-\theta_k}}, \tag{5}$$

$$E_k(L_{i,k}) = (L_{i,k})^{\gamma_k}. \tag{6}$$

Equation (5) states that bilateral trade shares between an origin country i and a destination j in any sector k satisfy a gravity equation with trade elasticity. θ_k Costinot et al. (2012) provide a multi-sector extension of Eaton and Kortum (2002) that provide micro-theoretical foundations for such functional form. The same micro-theoretical foundations can be invoked in the presence of external economies of scale, as in Kucheryavyy et al. (2017). Equation (6) allows external economies of scale to vary across sectors, but restricts the elasticity γ_k to be constant in each sector.

Let $x^t_{ij,k}$ now denote the trade share of exporter i for importer j in sector k in period t. Given equations (5) and (6), equation (4) simplifies into

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

where $L^t_{i,k}$ denotes the size of sector k in exporter i for period t and the residual $\eta^t_{ij,k} \equiv \tau^t_{ij,k}/\left(A^t_{i,k}B^t_{i,k}\right)$ captures variation in trade costs and productivity. Using 3 and the definition $c_{ij,k} \equiv \frac{\eta_{ij,k}w_i}{E_k(L_{i,k})}$, we can write the same relationship more directly as

$$\ln x^t_{ij,k} = \delta^t_{j,k} - \theta_k \ln w^t_i + \alpha_k \ln L^t_{i,k} + \theta_k \ln \eta^t_{ij,k}, \tag{7}$$

where $\delta^t_{j,k}$ is an importer-sector-time fixed effect that absorbs the rthe sector-level price index and expenditure of country j on sector k, and where $\alpha_k \equiv \theta_k\gamma_k$ is the revenue scale elasticity, namely the scale elasticity adjusted by the trade elasticity so as to capture the effect of scale on revenue rather than output.

3.2 Empirical Strategy

To estimate the scale elasticities γ_k we proceed in three steps: first, we estimate the revenue scale elasticities α_k building on Equation (7); second, we estimate trade elasticities θ_k from a standard gravity equation using tariff data; and third, we recover γ_k as $\hat{\alpha}_k/\hat{\theta}_k$.

3.2.1 Estimation of revenue scale elasticities

To use Equation (7) to estimate α_k we must first deal with the fact that we don't observe the composite input in efficiency units, and hence we don't have a proper empirical counterpart for the model objects $L^t_{i,k}$ nor w^t_i. Instead we observe sector-level value added, which in the model is given by $Y^t_{i,k} \equiv w^t_i L^t_{i,k}$. We can rewrite Equation (7) as

$$\ln x^t_{ij,k} = \delta^t_{j,k} - \theta_k(\gamma_k + 1)\ln w^t_i + \alpha_k \ln Y^t_{i,k} + \theta_k \ln \eta^t_{ij,k}.$$

We proceed by adding GDP per capita (in logs) interacted with a sector dummy to flexibly capture the term $-\theta_k(\gamma_k+1)\ln w^t_i$, and by adding a vector of bilateral determinants of trade frictions $V_{ij,k}$ to absorb part of the sector-level country-pair variation in $\theta_k \ln \eta^t_{ij,k}$,

$$\ln x^t_{ij,k} = \delta^t_{j,k} + V'_{ij,k}\Phi^t_k + \delta^t_k \ln(Y^t_i/N^t_i) + \alpha_k \ln Y^t_{i,k} + v^t_{ij,k},$$

where Φ^t_k is a vector of coefficients that will be estimated.

Next, we assume that there are no economies of scale in agriculture (we discuss this assumption below). Labeling agriculture as sector 1, the previous equation then implies

$$\ln\left(x^t_{ij,k}/x^t_{ij,1}\right) = \tilde{\delta}^t_{j,k} + V'_{ij}\tilde{\Phi}^t_k + \tilde{\delta}^t_k \ln(Y^t_i/N^t_i) + \alpha_k \ln Y^t_{i,k} + \tilde{v}^t_{ij,k}, \tag{8}$$

where $\tilde{\delta}^t_{j,k} \equiv \delta^t_{j,k} - \delta^t_{j,1}$, $\tilde{\Phi}^t_k = \Phi^t_k - \Phi^t_1$, and $\tilde{v}^t_{ij,k} \equiv v^t_{ij,k} - v^t_{ij,1}$. This equation says that a larger sector k in country i should lead to higher exports from i to any country j in sector k relative to those in sector 1. Notice that all coefficients (including fixed effects) are sector specific, so this equation holds sector by sector.

As discussed in Section 2.2, ordinary least squares (OLS) estimates of Equation (8) would be biased because of the fact that, even conditional on the controls included, an exporter's size in any sector $Y^t_{i,k}$ would respond endogenously to the idiosyncratic productivity shocks that are part of $\tilde{v}^t_{ij,k}$. We therefore pursue an instrumental variables (IV) strategy that exploits the simple fact that larger countries have larger sectors. In particular, we use the log of population, $\ln N^t_i$, as an instrument for the dependent variable $\ln Y^t_{i,k}$ in Equation (8). We expect that as country size increases then $Y^t_{i,k}$ will also increase and through economies of scale this would lead to higher productivity in sector k relative to sector 1 if $\gamma_k > \gamma_1 = 0$. In turn, this should be reflected in higher relative exports to any destination since $\alpha_k > \alpha_1 = 0$. This procedure will recover consistent estimates of α_k as long as there is zero covariance between Ricardian productivity and country size.

One concern with this IV strategy is that one could expect more populous countries to have a comparative advantage in manufacturing relative to agriculture. Formally, there could be a positive correlation between L^t_i and $\tilde{v}^t_{ij,k}$ through the negative impact of $\ln N^t_i$ on agricultural productivity in $v^t_{ij,1}$. The most likely channel would be through decreasing returns to more intensive use of a relatively fixed quantity of agricultural land as the population grows. We deal with this by adding a vector of controls for population density, both in terms of total land area and arable land area, which we denote by W^t_i. This leads to our baseline equation

$$\ln\left(x^t_{ij,k}/x^t_{ij,1}\right) = \zeta^t_{j,k} + V'_{ij}\Psi^t_k + (W^t_i)'\Xi^t_k + \zeta^t_k \ln(Y^t_i/N^t_i) + \alpha_k \ln Y^t_{i,k} + \varepsilon^t_{ij,k}, \tag{9}$$

where Ψ_k^t and Ξ_k^t are vectors of coefficients that will be estimated. [5]

While we use data from multiple time periods, this is not necessary for identification (that is, we could proceed with data from just one time period). Our main estimates will pool the data from all available time periods, but we will also describe how we obtain similar results when using the purely cross-sectional data from any particular time period alone. Accordingly, the identification of α_k will exploit both variation in sector size across origin countries, within a given time period, and variation across time periods, within a given origin.

To recap, our procedure for estimating α_k for sectors $k = 2, ..., K$ relies on Equation (9). We run this as a simple IV regression sector-by-sector using log population $\ln N_i^t$ as an instrument for sector size $\ln Y_{i,k}^t$. The critical assumptions for identification are that (1) there are no economies of scale in agriculture, (2) country size measured by population affects sector size, and (3) population is uncorrelated with Ricardian productivity differences conditional on the controls W_i. Assumption (2) seems incontrovertible, while assumption (3) is our basic non-testable exclusion restriction. Regarding assumption (1), note that this is a common assumption in the literature on industrial policy – see for example Matsuyama (1992). Now, if instead there were economies of scale in agriculture then we would have $-\alpha_1 \ln Y_{i,1}^t$ in the error term. Since the instrument $\ln N_i^t$ should be positively correlated with $\ln Y_{i,1}^t$, this would generate a downward bias in our estimates $\hat{\alpha}_k$. Thus, if we find that $\hat{\alpha}_k > 0$, the qualitative finding of positive economies of scale in manufacturing would not be compromised. For the specific estimation results, we will show how the estimates $\hat{\alpha}_k$ for $k = 2, ..., K$ vary with the assumed value of α_1, and we will show that the ranking of those estimates across sectors does not vary with α_1. Finally, for the quantitative analysis, we will study how the gains from industrial policy vary with different levels of α_1.

It is instructive to think about the reduced form regression associated with our IV approach. Abusing notation so that we can use the same labels for coefficients in the reduced form and in the second stage equation (9) above, the reduced form equation can be written as

$$\ln\left(x_{ij,k}^t / x_{ij,1}^t\right) = \zeta_{j,k}^t + V_{ij}' \Psi_k^t + W_i' \Xi_k^t + \zeta_k^t \ln(Y_i^t / N_i^t) + \beta_k \ln N_k^t + \varepsilon_{ij,k}^t.$$

[5] The coefficients Ξ_k^t are supposed to capture the effect of variables in W_i on the exports of sector k relative to sector 1. We allow such effects to vary by sector mostly to simplify the empirical analysis, since this will allow us to run the regressions sector by sector, but note that we can justify this from a structural perspective by noting that these coefficients will be affected by sector-specific trade elasticities, and because the variables in W_i may also affect the productivity in sector k directly.

According to the model, the coefficients β_k depends on upper-tier preferences, scale and trade elasticities, and the extent to which the country is open to international trade. Under mild conditions, larger countries would have higher relative productivity in sectors with high scale elasticities, and this would lead to higher exports in those sectors relative to agriculture as long as their trade elasticity is not too low. In other words, we expect the reduced-form coefficients $\hat{\beta}_k$ to line up with structural coefficients $\hat{\alpha}_k$. Similar reasoning implies that we should also see a positive correlation between the first and second stage coefficients as long as the upper-tier elasticity of substitution is not too low or countries are not too closed to trade. We will explore this below when we present the empirical results.

3.2.2 Estimation of Trade Elasticities

We now describe our procedure for estimating the trade elasticity parameter θ_k, separately for each sector k, following the functional form in equation (5). We do this by using variation in tariffs as in, for example, Caliendo and Parro (2015). To introduce this variation, we decompose the bilateral variable $\tau^t_{ij,k}$ into the portion due to *ad valorem* tariffs $T^t_{ij,k}$ and the portion due to transport costs, other trade barriers and unobserved demand shocks which we denote by $\kappa^t_{ij,k}$. This implies that $\tau^t_{ij,k} = (1 + T^t_{ij,k})\kappa^t_{ij,k}$.[6] We further assume that $\kappa^t_{ij,k}$ can be proxied for by a set of geographic variables G_{ij}—bilateral distance being the leading example—with impacts that may vary by sector and time. In particular, we write: $\ln \kappa^t_{ij,k} = \delta^t_k G_{ij} + \phi^t_{ij,k}$. Using this notation, and regardless of the particular shape that external economies take in $E_k(L^t_{i,k})$, the sector-specific gravity equation can be written as

$$\ln x^t_{ij,k} = \delta^t_{j,k} + \delta^t_{i,k} - \theta_k \ln(1 + T^t_{ij,k}) - \beta^t_k G_{ij} + \tilde{\phi}^t_{ij,k}, \tag{10}$$

where $\beta^t_k \equiv \theta_k \delta^t_k$, $\tilde{\phi}^t_{ij,k} \equiv -\theta_k \phi^t_{ij,k}$ and $\delta^t_{i,k}$ and $\delta^t_{j,k}$ are exporter-sector-year and importer-sector-year fixed effects, respectively.

Our focus is on the coefficient $-\theta_k$ on the tariff variable $\ln(1 + T^t_{ij,k})$, or (minus) the trade elasticity. We note that, conditional on the importer-sector-year fixed effects included here, only truly bilateral variation in tariffs (within importer-sector-year cells) will contribute to identification of θ_k. Such bilateral variation comes both from the presence of domestic trade observations (i.e., those for which $i = j$) or departures from the WTO's Most-Favored Nation (MFN) principle of tariffs due to unilateral preferences for poor countires granted by developed countries under the Generalized System of Preferences

[6] This multiplicative form is common in the literature, see for example Caliendo and Parro (2015); Hummels and Hillberry (2012); Head and Mayer (2013).

(GSP) and to the Preferential Trade Agreements (PTAs) that have proliferated in recent years (Limão, 2016).

Under the assumption that the unobserved variation in trade cost and demand shocks $\tilde{\phi}^t_{ij,k}$ is uncorrelated with the *ad valorem* tariffs $T^t_{ij,k}$, estimating equation 10 sector by sector using OLS will result in consistent estimates of θ_k. This is a standard assumption in the literature (e.g. (Caliendo and Parro, 2015; Hummels and Hillberry, 2012; Head and Mayer, 2013)), but it is not entirely unproblematic. As noted by Trefler (1993), this method tends to underestimate the true trade elasticities when sectors are composed of subsectors with varying trade elasticities. For political economy reasons, governments tend to set lower tariffs on subsectors with lower trade elasticities because those are precisely the sectors for which import competition is the least threatening to domestic industry. Given the high level of aggregation of our data relative to the tariff line sectors, this mechanism represents a nontrivial concern. Our use of variation in GSP and PTA relative to MFN tariff levels is also vulnerable to the potential endogeneity of the decision to grant GSP preferences or enter into a PTA to the level of trade between the participants. Further sources of downward bias include classical measurement error due to aggregation and other measurement issues, as well as our use of relatively short run variation in tariffs.[7]

To our knowledge, the literature on trade elasticity estimation has acknowledged these issues but not yet developed satisfactory methods of dealing with them. We have no significant methodological insights to add to this literature, and thus our estimation strategy shares its strengths and weaknessees. Ultimately, the trade elasticities play an important but restricted role in our analysis. They are not need to identify the α_k, which qualitatively establish the existence of scale economies independently of particular values of the trade elasticities. But they are necessary in order to separately identify the scale elasticities γ_k, which are in turn required to identify the optimal policies and to quantify their effects. We view our estimates as providing a reasonable benchmark, and explore the sensitivity of our results to alternative values.

3.3 Data

Our main estimation procedure seeks to estimate the external economies of scale elasticity γ_k within each sector k. This requires data on bilateral trade flows $X^t_{ij,k}$ and employment $L^t_{i,k}$ as in equation (7), as well as data on bilateral trade barriers, population, GDP per capita, and controls for agricultural comparative advantage that may be correlated with

[7] We utilize both cross-sectional and time series variation. However, given that tariffs have changed quite a bit over our sample period of 1995-2010 as well as in the period immediately preceding it, even the cross-sectional variation we exploit represents a mixture of long and short run variation.

population. We discuss each of these in turn.

We obtain data on bilateral trade flows $X^t_{ij,k}$ from the OECD's Inter-Country Input-Output (ICIO) tables. This source documents bilateral trade among 61 major exporters i and importers j, within each of 34 sectors k (27 of which are traded, with 15 in manufacturing) defined at a similar level to the 2-digit SIC, and for each year $t = 1995, 2000, 2005$, and 2010. The 15 manufacturing sectors s are those for which we aim to estimate γ_k.[8] Data on value added by country, sector and year are also available in the ICIO dataset. In our baseline model, value added corresponds to the wage bill $w^t_i \cdot L^t_{i,k}$, so with country fixed effects variation in value added corresponds to variation in sectoral employment levels both within and across countries.

We utilize data on two different sources of bilateral trade frictions: geographic variables and tariffs. Data on bilateral distance between the exporter i and importer j, denoted d_{ij}, come from the CEPII Gravity Database (Head et al., 2010). We use the population-weighted great circle distance between the large cities in the country pairs as our measure of bilateral distance. We use CEPII's measure of the own country distance d_{ii}, which is computed based on a measure of the average internal distance between a country's major population centers. We also use data on country contiguity in terms of land borders from the same source.

Data on tariffs (for the years 1995, 2000, 2005 and 2010) come from the United Nations Statistical Division, Trade Analysis and Information System (UNCTAD-TRAINS). When tariff data are missing for a given country-year we first use the nearest available preceding year for that country. Where data on an individual sector is missing, within a country-year, we interpolate using the country-wide average. And when no observations are available for a given country-year we interpolate using the worldwide sectoral average for that year. Our tariff measure is the applied tariff rate, based on the simple tariff line average as computed in TRAINS. We use UNCTAD's own product/industry concordance to map from tariff lines to ICIO sectors.

We take our baseline measure of population N^t_i from the POP^t_i 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. Our measure of GDP per capita also comes from the Penn World Tables 9.0; we use real GDP measured on the output side (rgdpo) and divide by population. Finally, we use two different variables to capture any correlation between population and agricultural productivity. The first is population density, constructed by dividing our population variable by country land area in square kilometers from CEPII. We also use a measure of arable land per person constructed by the World Bank in their

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

World Development Indicators dataset, which in turn relies on data from the FAO.

3.4 Baseline Results

3.4.1 Results for revenue scale elasticities

We first present our estimates of the parameter α_k for each sector. As described above, these estimates involve the logic of the nonparametric identification argument in Section 2.2, and the instrumental variable approach in Section 3.2.1. Our estimating equation is equation (9), with log relative exports as the dependent variable and log sectoral output as the independent variable of interest. The RHS of equation (9) features three additional sets of control variables: GDP per capita, bilateral trade barriers and controls for population density. We use a border dummy, a contiguity dummy, log bilateral distance and log bilateral tariffs as our measures of trade barriers, and we use both log population per square kilometer of land area and log population per square kilometer of arable land as our density measures. Each control variable enters into the regression with a sector-specific elasticity; border, contiguity and log bilateral distance have sector-time-specific elasticities . This specification allows us to flexibily capture differential and time-varying impacts of trade barriers while conveniently allowing estimation to proceed sector by sector.

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

First-stage estimates

We begin by reporting the first-stage regression from an IV specification that pools our estimates of α_k across sectors (and hence arrives at a single estimate of manufacturing-wide α). This amounts to simply using $\ln N_i^t$ as an IV for $\ln Y_{i,k}^t$ in a version of equation (9) that sets $\alpha_k = \alpha$ for all k.

The results from this pooled first-stage regression are presented in Table 1, column 3. It is clear that, on average across all sectors, our instrument is a powerful predictor

Table 1: Pooled (All Sectors) Estimates of α

	log (bilateral sales)		log (value added)	log (bilateral sales)
	OLS	IV	First Stage	Reduced Form
	(1)	(2)	(3)	(4)
log (value added)	0.39	0.25		
	(0.02)	(0.03)		
log (population)			1.14	0.28
			(0.02)	(0.03)
Within R^2	0.17	0.16	0.81	0.12
Observations	197,300	197,300	197,300	197,300
First-state F-statistic		3261		

Notes: Column (1) reports the OLS estimate, and column (2) the IV estimate, of equation (5) subject to the constraint that all sectors have the same economies of scale elasticity (i.e. $\gamma_k = \gamma$, for all sectors k). Column (3) reports the corresponding pooled first-stage specification while column (4) reports the reduced form. The instrument is the natural log of country population. All regressions control for importer-sector-year fixed-effects as well as other controls described in the text. Standard errors in parentheses are clustered at the exporter-sector level.

of the size (as measured by $\ln Y^t_{i,k}$) of any given exporter-sector-year; higher population predicts higher sectoral output in every sector. Indeed, the t-statistic on the IV (equal to 57.11) implies a first-stage F-statistic of 3261, and hence little concern about finite sample bias due to a weak instrument. This strong performance is unsurprising given the nature of the relationship that we expoit in the first stage, which is simply the fact that larger countries produce more in every sector.

Columns 3 and 5 of Table 2 (which reports a wider set of results that we discuss shortly) contains the corresponding first-stage coefficients and F-statistics for the case when α_k varies across sectors. Recall that our system has 15 endogenous variables ($\ln Y^t_{i,k}$ interacted with 15 sector indicators), so there are 15 first-stage equations. Since our coefficients are sector-specific there is no interaction across equations, so we run them (and compute F-statistics) sector by sector. In each case $\ln N^t_i$ is a strong instrument, with F-statistics in the triple digits. Again, this is to be expected based on the nature of the first stage.

Second-stage estimates

As in the previous discussion, we begin with a pooled estimate of α—one that is restricted to be common to all sectors. These pooled estimates are reported in Table 1. Column 1 starts with the OLS estimate of equation (9) for the pooled model. The precisely estimated

Table 2: Sector-specific Estimates of α_k

Sector	OLS (1)	IV (2)	First Stage (3)	Reduced Form (4)	First Stage F (5)
Food, Beverages and Tobacco	0.01	-0.03	0.99	-0.03	768.0
	(0.06)	(0.07)	(0.04)	(0.07)	
Textiles	0.17	0.04	0.96	0.04	460.0
	(0.10)	(0.11)	(0.05)	(0.11)	
Wood Products	0.31	-0.04	0.88	-0.04	145.7
	(0.11)	(0.17)	(0.07)	(0.15)	
Paper Products	0.36	0.29	1.05	0.30	917.0
	(0.09)	(0.10)	(0.03)	(0.10)	
Coke/Petroleum Products	0.42	0.40	1.50	0.60	178.4
	(0.05)	(0.07)	(0.11)	(0.11)	
Chemicals	0.29	0.24	1.22	0.29	888.9
	(0.07)	(0.09)	(0.04)	(0.11)	
Rubber and Plastics	0.28	0.21	1.08	0.22	455.9
	(0.08)	(0.10)	(0.05)	(0.12)	
Mineral Products	0.30	0.24	1.01	0.24	409.0
	(0.10)	(0.12)	(0.05)	(0.12)	
Basic Metals	0.49	0.34	1.29	0.44	119.6
	(0.07)	(0.11)	(0.12)	(0.17)	
Fabricated Metals	0.27	0.22	1.03	0.23	464.3
	(0.10)	(0.12)	(0.05)	(0.12)	
Machinery and Equipment	0.37	0.25	1.21	0.30	425.2
	(0.07)	(0.09)	(0.06)	(0.11)	
Computers and Electronics	0.54	0.20	1.10	0.22	129.4
	(0.08)	(0.11)	(0.10)	(0.14)	
Electrical Machinery, NEC	0.44	0.24	1.12	0.27	307.7
	(0.08)	(0.09)	(0.06)	(0.12)	
Motor Vehicles	0.62	0.51	1.56	0.80	291.7
	(0.07)	(0.08)	(0.09)	(0.15)	
Other Transport Equipment	0.49	0.38	1.20	0.45	192.1
	(0.06)	(0.07)	(0.09)	(0.10)	

Notes: Column (1) reports the OLS estimate, and column (2) the IV estimate, of equation (5) when the scale elasticity is allowed to differ across sectors. Column (3) reports the corresponding first-stage specifications while column (4) reports the reduced form coefficients. The instrument is the natural log of country population in each equation. The instrument is the natural log of country population. All regressions control for importer-sector-year fixed-effects as well as other controls described in the text. All coefficients are sector specific, so estimation is sector by sector. Standard errors in parentheses are clustered at the exporter-sector level. The correlation between the coefficients in columns (2) and (3) is 0.85, and the correlation between the coefficients in columns (2) and (4) is 0.97.

coefficient estimate of 0.39 implies that there is a strong positive correlation between the number of workers in an exporter-industry and that exporter-industry's aggregate productivity (as revealed by its relative success in foreign export markets). But as discussed above, the error term in equation (9) includes unobserved productivity variation that we would expect to be positively correlated with sector size to the extent that relatively productive sectors expand (hire more workers) precisely because of their productivity advantage. This logic implies that the OLS estimate in column 1 would be biased upwards.

Column 3 of Table 2 reports the corresponding IV estimate, with $\ln N_i^t$ serving as an instrument for $\ln Y_{i,k}^t$. As expected, the IV estimate of 0.25 is smaller that the OLS estimate, but positive and statistically significantly different from both zero and the OLS estimate at the 1% level.

Table 3 presents the equivalent estimates of equation (7) for the case with unrestricted coefficients, delivering independent estimates of α_k for each sector k. Column 1 reports the OLS estimate for each sector, and column 2 reports the corresponding IV estimate. We see considerable heterogeneity in both estimates, with the OLS ranging from roughly 0-0.6 and the IV from -0.5-0.5. However, note that in every single industry the IV estimate is below the OLS, sometimes by a large amount (e.g., Wood Products, Computers and Electronics). Standard errors are certainly larger than in the pooled case, but many of the IV estimates are statistically different from zero at the 5% or even 1% level. We can also easily reject the hypothesis of coefficient equality at the 1% level. This heterogeneity is important for the scope for industrial policy, as we discuss in Section 4 below.

Turning to the specific sectoral elasticity estimates, we see some correlation between the estimates and our priors regarding which sectors enjoy significant external economies of scale. Sectors producing simple products in which innovation is likely to be less important, such as Food, Beverages and Tobacco, Textiles and Wood Products, tend to have values of α_k that are statistically indistinguishable from zero. Sectors that produce more complex products, such as Motor Vehicles and Other Transportation Equipment tend to have higher estimates of α_k. Of course, estimates of α_k contain information about both the scale and trade elasticities, so by themselves these estimates do not imply that high α sectors have high scale elasticities. In order to distinguish between these elasticities we will need the trade elasticity estimates, which we present below.

One interesting feature of our strategy is that our theoretical framework gives a prediction for the magnitude of the reduced form and first stage coefficients as well as the signs. Recall from the discussion in Section 3.2.1 that we expect the reduced form coefficients to be positively correlated with the estimated αs. Moreover, under the additional assumption that the upper-tier elasticity of substitution is not too low or that countries are not too closed, the first stage coefficients will also be positively correlated with the second stage coefficients. Thus the theory provides a quasi-structural link between the three regressions that can be used as a sort of overidentifying test of the model mechanisms.

Column 4 of Table 1 reports the pooled reduced form cofficient of 0.28, while column 3 of the same table reports the pooled first stage estimate of 1.14. The reduced form coefficient being higher than 0 implies that, as population increases, exports of manufacturing sectors tend to increase faster than agricultural exports. The first stage coefficient being higher than 1 implies that, as population increases, the total size of the manufacturing sectors increases faster than the agricultural sectors. This is exactly what a model in which α in the manufacturing sector is larger than that in agriculture would predict (subject to the caveats mentioned above). If we rule out mechanical correlation between country size and exogenous comparative advantage in manufacturing, which (conditional on our controls) is our identifying assumption, then these regressions already provide powerful evidence of scale economies in manufacturing. An even sharper set of results emerges from comparing the sector specific estimates of α_k with the corresponding reduced form and first stage estimates in Table 2. Here the correlation between the reduced form and second stage estimates is 0.97 and that between the first and second stage estimates is 0.85. These high correlations provide a reassuringly tight fit between the data moments and our theoretical framework.

As discussed in Section 2.2, our estimates are predicted on the assumption that there are no scale economies in agriculture ($\alpha_1 = 0$). If instead it were the case that $\alpha_1 > 0$, we would expect the IV estimates of $\alpha_k$in Table 2 to be biased downwards, due to the negative correlation between the independent variable $\ln Y_{i,k}^t$ and the omitted term $-\alpha_1 \ln Y_{i,1}^t$, with the degree of bias in each sector depending on the strength of the correlation between these two terms. One tantalizing possibility is that this correlation is not too different across sectors. If that is the case then our IV estimates in Table 2 will correctly identify the *relative* α_k across manufacturing subsectors while being biased downward in absolute magnitude. This is important because the optimal industrial policy is scale invariant; as we show in Section 4, simply raising each scale elasticity by a constant amount will not change the size of the optimal subsidy in each sector.

We explore the robustness of our results to alternative values of α_1 by assuming that it takes on different known positive values, then adding the term $-\alpha_1 \ln Y_{i,1}^t$ to the LHS of our estimating equation (9) and re-estimating the α_k. Under the maintained assumptions that the rest of our identifying restrictions continue to hold, this procedure will consistently recover the true α_k in manufacturing so long as α_1 is correctly specified.

Table 3 reports the results of this exercise for $\alpha_1 = 0.1, ..., 0.5$. We see that each increase

Table 3: Estimates of α_k For Alternative Assumptions on Scale Economies in Agriculture

Sector	IV, $\alpha_1 = .1$ (1)	IV, $\alpha_1 = .2$ (2)	IV, $\alpha_1 = .3$ (3)	IV, $\alpha_1 = .4$ (4)	IV, $\alpha_1 = .5$ (5)
Food, Beverages and Tobacco	0.06	0.16	0.25	0.35	0.44
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Textiles	0.14	0.24	0.34	0.44	0.53
	(0.11)	(0.11)	(0.10)	(0.10)	(0.10)
Wood Products	0.07	0.17	0.28	0.39	0.50
	(0.17)	(0.16)	(0.15)	(0.15)	(0.14)
Paper Products	0.38	0.47	0.56	0.65	0.74
	(0.10)	(0.10)	(0.09)	(0.09)	(0.09)
Coke/Petroleum Products	0.46	0.52	0.59	0.65	0.71
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Chemicals	0.32	0.39	0.47	0.55	0.63
	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Rubber and Plastics	0.29	0.38	0.47	0.56	0.64
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
Mineral Products	0.34	0.43	0.53	0.62	0.71
	(0.12)	(0.11)	(0.11)	(0.11)	(0.11)
Basic Metals	0.42	0.49	0.57	0.64	0.71
	(0.10)	(0.10)	(0.09)	(0.09)	(0.09)
Fabricated Metals	0.31	0.41	0.50	0.59	0.68
	(0.12)	(0.11)	(0.11)	(0.11)	(0.11)
Machinery and Equipment	0.33	0.41	0.48	0.56	0.64
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Computers and Electronics	0.29	0.37	0.46	0.54	0.63
	(0.11)	(0.10)	(0.10)	(0.10)	(0.10)
Electrical Machinery, NEC	0.32	0.41	0.49	0.58	0.66
	(0.09)	(0.09)	(0.09)	(0.08)	(0.08)
Motor Vehicles	0.57	0.63	0.69	0.76	0.82
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Other Transport Equipment	0.46	0.54	0.61	0.69	0.77
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Avg. Diff. from Baseline	0.08	0.17	0.25	0.34	0.42
St. Dev. Diff. from Baseline	0.01	0.02	0.04	0.05	0.06
Rank Corr. with Baseline	0.99	0.97	0.95	0.94	0.90

Notes: Each column reports the IV estimates of Equation (7) under the alternative assumptions for the value of α_1, the α in agriculture. All other details are the same as in column 2 of Table 2. Deviations from baseline calculated as the difference between the estimate in each column and the corresponding estimate in Table 2, column 2.

of α_1 by 0.1 tends to increase the estimated α_k by 0.08 on average. There is some heterogeneity in the response across different manufacturing subsectors, and this heterogeneity tends to increase as the value of α_1 increases. For example, the standard deviation of the difference between the baseline estimates and the alternative estimates increases from 0.01 when $\alpha_1 = 0.1$ to 0.06 when $\alpha_1 = 0.5$. However, overall the relative rankings and magnitudes of the α_k are remarkably insensitive to the value of α_1. Large differences between the baseline estimates tend to remain large for alternative values of α_1, and small differences tend to remain small. The rank correlation between the baseline estimates and the alternatives is very close to 1 for low values of α_1, and remains very high (0.90) even when $\alpha_1 = 0.5$.

These results demonstrate that our conclusions regarding which sectors a central planner would like to subsidize as well as the magnitudes of the subsidies are fairly robust to our assumption of zero scale economies in agriculture. It is true that the welfare effects of such policies are not scale invariant; however, our analysis in 4 will show that the gains from industrial policy are only modestly affected by our assumptions regarding α_1.

3.4.2 Results for trade elasticities

As described above, we use the gravity equation (10) to estimate trade elasticities θ_k for each sector k. These estimates are shown in column 1 of Table 3. All estimates are in the expected range (with $\theta_k > 0$), and statistically significantly at standard levels. The magnitudes, which range from 1.45 in the "Food, Beverages and Tobacco" sector to 8.04 in the "Wood Products" sector, are all in the broad range implied by prior estimates (such as those from Caliendo and Parro (2015), which range from 0.37 to 12.79 if the estimate for the "Petroleum" sector, equal to 51.08, is excluded).

3.4.3 Results for scale elasticities

Finally, in column (2) of Table 3 we report the implied value of γ_k, the scale elasticity, for each of our manufacturing sectors. This is simply the ratio of the revenue scale elasticity (equal to $\alpha_k \equiv \theta_k \gamma_k$) reported in column (2) of Table 2 to the estimate of the trade elasticity θ_k in column 1 of Table 3. These estimates show considerable heterogeneity, ranging from -0.02 in the "Food, Beverages and Tobacco" sector to 0.23 in the "Motor Vehicles" sector.[9] The implied scale elasticities increase as we allow for positive values of α_1, in accordance of the results in Table 3, but the broad picture remains the same. Overall, there is scope

[9] Confidence intervals that correspond to these γ_k estimates, based on a bootstrapping procedure, are in progress.

for industrial policy that promotes the expansion of certain manufacturing sectors at the expense of others. Our calculations in the next section ask how large the benefits from such actions could be.

4 Gains from Trade and Industrial Policy

In the previous section, we have estimated external economies of scale, E_k, using knowledge of input demand, χ_k, in each sector k. In this section, we show how one can use the previous estimates to characterize the structure of optimal trade and industrial policy across sectors as well as the associated welfare gains. For expositional purposes, we focus in the main text on the case of a small open economy that can only affect the price of its own good relative to goods from other countries: relative prices in the rest of the world, employment, and expenditure are taken as exogenously given. The general case, which we will study in some of our quantitative exercises, can be found in the Appendix – **this is work in progress**.

4.1 Optimal Trade and Industrial Policy

We first describe the problem of a fictitious planner that directly controls consumption and production in a given country j to maximize utility in that country. We then show how that allocation can be decentralized through sector-level production and trade taxes.

The Planner's Problem. Like in Section 2, it is convenient to focus on inputs rather than goods. Following Adao, Costinot and Donaldson (2017), let $L_{ij,k}$ denote the demand, in efficiency units, for inputs from country i in country j within a given sector k, and let $V_j(\{L_{ij,k}\}_{i,k})$ denote the utility of the representative agent in country j associated with a given vector of input demand,

$$V_j(\{L_{ij,k}\}_{i,k}) \equiv \max_{\{q_{ij,k}(\omega), l_{ij}^{k}(\omega)\}} U(\{U_{j,k}(\{B_{ij,k}(\omega) q_{ij,k}(\omega)\}_{i,\omega})\}_k)$$

$$q_{ij,k}(\omega) \leq A_{i,k}(\omega) l_{ij,k}(\omega) \text{ for all } \omega,\ i, \text{ and } k,$$

$$\int l_{ij,k}(\omega) d\omega \leq L_{ij,k} \text{ for all } i \text{ and } k.$$

Table 4: Estimates of Trade Elasticities (θ_k) and External Economies of Scale (γ_k)

Sector	Trade elasticity (θ_k) (1)	Scale elasticity (γ_k) (2)
Food, Beverages and Tobacco	1.45 (0.29)	-0.02
Textiles	3.81 (0.98)	0.01
Wood Products	8.04 (1.19)	-0.01
Paper Products	6.50 (1.30)	0.04
Coke/Petroleum Products	2.46 (1.11)	0.16
Chemicals	5.79 (1.13)	0.04
Rubber and Plastics	5.66 (1.04)	0.04
Mineral Products	5.71 (0.92)	0.04
Basic Metals	6.58 (1.25)	0.05
Fabricated Metals	5.92 (1.08)	0.04
Machinery and Equipment	6.20 (1.40)	0.04
Computers and Electronics	2.60 (0.93)	0.08
Electrical Machinery, NEC	6.57 (1.20)	0.04
Motor Vehicles	2.25 (0.90)	0.23
Other Transport Equipment	6.45 (0.88)	0.06
Observations	207,553	
Within R^2	0.509	
Average Values	5.07	.06

Notes: Column (1) reports the OLS estimate of the trade elasticity θ_k for each manufacturing sector k in our sample. This specification, corresponding to equation (10), controls for importer-sector-year fixed-effects, exporter-sector-year fixed-effects, and contiguity dummies and log distance each interacted with sector and year indicators. Standard errors in parentheses are clustered at the exporter-sector level. The estimates in column (2) report the corresponding implied estimate of the scale elasticity γ_k for each sector, which is computed as the ratio of the revenue scale elasticity in column (2) of Table 2 to the trade elasticity in column (1) of the present table.

Expressed in terms of input choices, the planner's problem in country j is

$$\max_{\{\tilde{L}_{ij,k}\}_{i,k},\{\tilde{L}_{ji,k}\}_{i\neq j,k},\{L_{j,k}\}_k} V_j(\{\tilde{L}_{ij,k}\}_{i,k}) \tag{11a}$$

$$\sum_{i\neq j,k} c_{ij,k}\tilde{L}_{ij,k} \leq \sum_{i\neq j,k} c_{ji,k}(\tilde{L}_{ji,k})\tilde{L}_{ji,k}, \tag{11b}$$

$$\sum_i \eta_{ji,k}\tilde{L}_{ji,k} \leq E_k(\tilde{L}_{j,k})\tilde{L}_{j,k}, \text{ for all } k, \tag{11c}$$

$$\sum_k \tilde{L}_{j,k} \leq L_j. \tag{11d}$$

Equation (11b) is a trade balance condition. It states that the value of inputs imported by country j is no greater than the value of its exports. In general, the prices of both imports and exports, $c_{ij,k}$ and $c_{ji,k}$, should be functions of the entire vector of country j's input choices. In the small open economy case that we focus on, country j takes import prices, $c_{ij,k} \equiv \eta_{ij,k}w_i/E_k(L_{i,k})$, as given, but treats export prices, $c_{ji,k}(\tilde{L}_{ji,k})$, as a function of exports, $\tilde{L}_{ji,k}$, implicitly given by the solution to

$$\chi_{ji,k}(c_{1i,k}, ..., c_{ji,k}, ..., c_{Ii,k}) = c_{ji,k}\tilde{L}_{ji,k}/X_{i,k}, \tag{12}$$

where the costs of other exporters, $\{c_{li,k}\}_{l\neq j}$, and expenditure, $X_{i,k}$, are again taken as given by country j.

Equations (11c) and (11d) characterize the production possibility frontier in country j. Equation (11c) captures technological constraints; it states total demand for inputs across destinations i, adjusted by bilateral exogenous efficiency term $\eta_{ji,k} \equiv \tau_{ji,k}/\left(A_{j,k}B_{j,k}\right)$, can be no greater than the total supply, in efficient units, in country j and sector k. The term $A_{j,k}B_{j,k}E_k(\tilde{L}_{j,k})$ reflects the fact that because of economies of scale, an increase $\tilde{L}_{j,k}$ leads either to larger quantities or higher quality goods being produced with a given amount of inputs, and hence an increase the number of inputs supplied in efficiency units. Equation (11d) is a resource constraint; it states that the sum of inputs allocated across sectors k can be no greater than the total supply of inputs in country j.

The Decentralized Equilibrium with Taxes. Now consider a decentralized equilibrium with taxes. Production in a given sector k may be subject to an ad-valorem production subsidy, $s_{j,k}$, which creates a wedge between the prices faced by firms and consumers in country j. Imports and exports in a given sector k may also be subject to an import tariff, $t_{ij,k}$, and an export tax, $t_{ji,k}$, which creates a wedge between input prices in country j and the rest of the world. Net revenues from taxes and subsidies are rebated through a lump-sum transfer, T_j, to the representative agent in country j.

In the decentralized equilibrium, consumers maximize utility and firms maximize profits, taking after-tax prices and transfers as given, and markets clear. To prepare the characterization of optimal taxes, it is convenient to describe the allocation in the decentralized equilibrium as the solution to an alternative planning problem,

$$\max_{\{\tilde{L}_{ij,k}\}_{i,k},\{\tilde{L}_{ji,k}\}_{i\neq j,k},\{\tilde{L}_{j,k}\}_{k}} V_j(\{\tilde{L}_{ij,k}\}_{i,k}) \tag{13a}$$

$$\sum_{i\neq j} c_{ij,k}(1+t_{ij,k})\tilde{L}_{ij,k} \leq \sum_{i\neq j} c_{ji,k}(L_{ji,k})(1-t_{ji,k})\tilde{L}_{ji,k} + T_j, \tag{13b}$$

$$\sum_i \eta_{ji,k}\tilde{L}_{ji,k} \leq (1+s_{j,k})E_k(L_{j,k})\tilde{L}_{j,k}, \text{ for all } k, \tag{13c}$$

$$\sum_k \tilde{L}_{j,k} \leq L_j. \tag{13d}$$

There are two key differences between problems (11) and (13). First, whereas the planner internalizes sector-level economies of scale, $E_k(\tilde{L}_{j,k})$, firms and consumers do not, which explains why $E_k(L_{j,k})$ depends on the equilibrium sector size, $L_{j,k}$, rather than the choice variable, $\tilde{L}_{j,k}$, in equation (13c). This creates a rationale for Pigouvian taxation, that is production subsidies, $\{s_{j,k}\}$, that may be non-zero at the optimum. Second, whereas the planner recognizes its market power on foreign markets, $c_{ji,k}(\tilde{L}_{ji,k})$, firms and consumers do not, which explains why $c_{ji,k}(L_{ji,k})$ in equation (13b) depends on the quantity exported in equilibrium, $L_{ji,k}$, rather than the choice variable, $\tilde{L}_{ji,k}$. This creates a rationale for export taxes, $\{t_{ji,k}\}$, that manipulate country j's terms-of-trade.

The Structure of Optimal Policy. To characterize the structure of optimal policy, we compare the solutions to (11) and (13) and derive necessary conditions on production subsidies and trade taxes such that the two solutions coincide.

Consider first the solution to (11). The first-order conditions with respect to $\{\tilde{L}_{j,k}\}_k$, $\{\tilde{L}_{ji,k}\}_{i\neq j,k}$, and $\{\tilde{L}_{ij,k}\}_{i,k}$ imply

$$[E'_k(L_{j,k})L_{j,k} + E_k(L_{j,k})]\rho_{j,k} = \rho_j, \tag{14}$$

$$\lambda_j[c'_{ji,k}(L_{ji,k})L_{ji,k} + c_{ji,k}(L_{ji,k})] = \eta_{ji,k}\rho_{j,k}, \tag{15}$$

$$dV_j(\{L_{ij,k}\}_{i,k})/dL_{ij,k} = \lambda_j c_{ij,k}, \text{ if } i \neq j, \tag{16}$$

$$dV_j(\{L_{ij,k}\}_{i,k})/dL_{ij,k} = \eta_{jj,k}\rho_{j,k}, \text{ if } i = j. \tag{17}$$

where λ_j, $\{\rho_{j,k}\}$ and ρ_j denote the values of the Lagrange multipliers associated with constraints (11b)-(11d) at the optimal allocation.

Now suppose that the same allocation arises at the solution to (13). The first-order conditions associated with this problem imply

$$(1 + s_{j,k}) E_k(L_{j,k}) \rho^e_{j,k} = \rho^e_j, \tag{18}$$

$$\lambda^e_j (1 - t_{ji,k}) c_{ji,k}(L_{ji,k}) = \eta_{ji,k} \rho^e_{j,k}, \tag{19}$$

$$dV_j(\{L_{ij,k}\}_{i,k}) / dL_{ij,k} = \lambda^e_j (1 + t_{ij,k}) c_{ij,k}, \text{ if } i \neq j, \tag{20}$$

$$dV_j(\{L_{ij,k}\}_{i,k}) / dL_{jj,k} = \eta_{jj,k} \rho^e_{j,k}, \text{ if } i = j, \tag{21}$$

where λ^e_j, $\{\rho^e_{j,k}\}$ and ρ^e_j denote the values of the Lagrange multipliers associated with constraints (13b)-(13b). A simple comparison of equations (14)-(17) and equations (18)-(21) leads to the following proposition.

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

$$1 + s_{j,k} = (1 + s_j)(1 + \frac{d \ln E_k}{d \ln L_{j,k}}),$$

$$1 - t_{ji,k} = (1 + t_j)(1 + \frac{d \ln c_{ji,k}}{d \ln L_{ji,k}}),$$

$$1 + t_{ij,k} = 1 + t_j,$$

for all i, j, and k.

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

It is worth noting that while we have focused on the case of a small open economy, this restriction is only relevant for the structure of optimal trade policy, which would depend, in general, on the entire vector of imports and exports by country j. The optimal Pigouvian tax, in contrast, is always given by $\frac{d \ln E_k}{d \ln L_{j,k}}$. Formally, this can be seen easily from the fact that the technological constraints (11c) and (13c) would be unchanged in the case of a large open economy.

Table 5: **Gains from Optimal Policies, Selected Countries**

Country	Optimal Policy (1)	Classic Trade Pol. (2)	Add Industrial Policy (3)	Constrained Industrial Pol. (4)	Global Efficient Pol. (5)
United States	0.31%	0.23%	0.07%	0.08%	0.21%
China	0.40%	0.31%	0.09%	0.19%	0.07%
Germany	0.67%	0.49%	0.18%	0.12%	-0.30%
Ireland	1.29%	1.20%	0.10%	0.70%	-0.22%
Vietnam	1.14%	1.02%	0.11%	0.69%	0.93%
Avg, Unweighted	**.71%**	**0.60%**	**0.11%**	**0.24%**	**0.13%**
Avg, GDP Weighted	**0.46%**	**0.36%**	**0.10%**	**0.14%**	**0.06%**

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.

4.2 Welfare Gains from Optimal Policy

According to Proposition 1, optimal industrial and trade policy only require knowledge of two types of elasticities, $\frac{d\ln E_k}{d\ln L_{j,k}}$ and $\frac{d\ln c_{ji,k}}{d\ln L_{ji,k}}$. Under the parametric restrictions imposed in Section 3—equations (5) and (6)—we have $\frac{d\ln E_k}{d\ln L_{j,k}} = \gamma_k$ and $\frac{d\ln c_{ji,k}}{d\ln L_{ji,k}} = -\frac{1}{1+\theta_k}$, where the second expression uses the fact that $\frac{d\ln c_{ji,k}(L_{ji,k})}{d\ln L_{ji,k}} = -\frac{1}{1-\frac{d\ln \chi_{ji,k}}{d\ln c_{ij,k}}}$, by equation (12). Under our normalization, $s_j = t_j = 0$, this leads to

$$
\begin{aligned}
s_{j,k} &= \gamma_k, \text{ for all } k, \\
t_{ji,k} &= \frac{1}{1+\theta_k}, \text{ for all } k \text{ and } i \neq j, \\
t_{ij,k} &= 0, \text{ for all } k \text{ and } i.
\end{aligned}
$$

To quantify the welfare effect of implementing the optimal policy, we further assume that preferences are Cobb-Douglas with expenditure shares $\beta_{j,k}$ for $k = 1, ..., K$. As in Dekle et al. (2007) we allow for transfers across countries so that country i has a deficit D_i, with $\sum_i D_i = 0$ and compute counterfactuals using exact hat algebra under the assumption that the data comes from an equilibrium without taxes or subsidies. For the manufacturing subsectors we use our estimated values for the scale elasticities, and we set $\gamma_k = 0$ for all sectors outside of manufacturing. For all manufacturing sectors, we use our estimated values for trade elasticities, and for all non-manufacturing sectors we set the trade elasticity equal to the simple average across manufacturing sectors, which is 5.1.

In column 1 of Table 5 we report the gains from optimal policy for some selected countries assuming that each of them is a small economy. As we discuss below, this turns out to be an innocuous assumption even for the large countries in our sample. In column 2 we report the gains from imposing the export taxes that are part of the optimal policy, i.e., $1/\left(1+\theta_k\right)$, but now assuming that there are no production subsidies. One can think of these gains as those that would arise if the government rightly thought that it was a small open economy with trade elasticities θ_k but wrongly through that there were no external economies of scale. In column 3 we show the simple difference between columns 1 and 2 – this gives the gains that such a government would realize if it suddenly discovered that there are external economies of scale given by γ_k, and imposed the optimal subsidies to internalize those externalities.[10]

The results of the first three columns in Table 5 reveal that the gains from optimal policy (column 1) are not very large, and that the gains from industrial policy (column 3) are particularly small: the GDP-weighted average of these gains is 0.1%. This seems small for a government that is assumed to be benevolent and omniscient, and given our assumption that there are no scale economies outside manufacturing. For comparison, the average gains from trade policy alone are also 0.6%, while the average gains from trade are larger by an order of magnitude.

Why are the gains from industrial policy small? A necessary condition for industrial policy to generate large welfare gains is that there be significant differences in scale elasticities across sectors. Our estimated scale elasticities exhibit significant differences across manufacturing subsectors, and we have assumed that there are no scale economies in non-manufacturing, so one could have expected large gains from reallocating resources towards sectors with the largest scale elasticities. However, the economic forces at play are more subtle than that.

In a closed economy, the benefits of reallocating resources from low-γ to high-γ sectors are limited by the ability of domestic demand to absorb the additional output. With limited substitutability in consumption across sectors, as in our case with Cobb-Douglas upper-tier preferences, the gains from industrial policy in autarky are very small. The percentage gains from industrial policy under autarky for a country with expenditure shares $\beta_{j,k}$ can be computed directly as $\prod_k \left(\frac{1+\gamma_k}{1+\bar{\gamma}_j}\right)^{(1+\gamma_k)\beta_{j,k}} - 1$, where $\sum_k \beta_{j,k}\gamma_k$. Using this for-

[10]We have also computed the welfare effect of imposing production subsidies equal to γ_k but with no export taxes. This would be the policy of a government that doesn't understand that it has market power abroad and so ignores the effect of production subsidies on the country's terms of trade. We find that some countries actually lose from implementing this policy – a case of immiserizing growth. We choose to report results as in Table 3 because it seems more plausible that governments don't know about the gains from industrial policy than that they don't know about the gains from trade policy.

mula for the countries in our sample yields a GDP-weighted average gains of 0.04%.[11] When the high elasticity sectors are open to trade, a small economy can take advantage of the more elastic international demand for its products and the gains from industrial policy will be higher. Still, to benefit from exporting a country must also be able to replace the lost production in the low elasticity sectors with imports. If the low elasticity sectors are closed, the gains from reallocation are still limited by the inelastic domestic demand for the non-traded goods. For most countries non-manufacturing sectors both account for the lion's share of expenditure and are largely non-traded. This helps account for the small gains from industrial policy, even when non-manufacturing sectors are assumed to have zero scale elasticity. Thus the key to increasing the effectiveness of industrial policy may be trade liberalization in low scale elasticity sectors. Of course, from the point of view of the world as a whole, this cannot have a big impact, for the same reason that the gains from industrial policy are small in autarky.

In column 4 of Table 5 we present the results of a different exercise: we compute the gains from production subsidies chosen to maximize a country's welfare assuming that it cannot use trade taxes. This is motivated by the fact that such taxes are beggar thy neighbor policies, and so there may be an international agreement such as GATT/WTO preventing countries from using them. We solve for this numerically by finding the production subsidies that maximize utility conditional on zero trade taxes.[12] Intuitively, these constrained-optimum production subsidies involve a compromise between internalizing the production externalities via Pigouvian subsidies and improving the country's terms of trade by taxing the sectors with the lowest trade elasticities.[13] The results in column 4 differ somewhat from those in column 3. For example, the gains from industrial policy in column 4 are significantly higher for Ireland and Vietnam compared to those in column 3. The basic idea is that, without recourse to trade taxes, countries can use production subsidies to improve their terms of trade and capture some of the gains displayed in column 2. If those gains are large, as in the case of very open economies like Ireland and Vietnam,

[11] Moving away from Cobb-Douglas preferences would obviously affect this result. If the upper-tier elasticity of substitution was high then exploiting difference in scale elasticities through industrial policy would yield larger gains. The problem is that, if anything, a reasonable alternative to Cobb-Douglas preferences is to have the upper-tier elasticity of substitution be less than one – see Herrendorf et al. (2014) and Cravino and Sotelo (2017).

[12] Lashkaripour and Lugovskyy (2018) derive an implicit formula for these constrained-optimal production subsidies in a two country world.

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

then the gains realized in column 4 can be higher as well.[14]

Finally, in column 5 of Table 5 we show the gains that each country derives if all countries follow the policy that maximizes world welfare $\sum_i \pi_i U_i$ for any set of Pareto weights π_i – this entails $s_{i,k} = \gamma_k$ for all i, k. We see that there are large distributional implications associated with the imposition of globally efficient production subsidies. For example, Vietnam experiences a welfare gain of 0.93% while Ireland suffers a welfare loss of 0.22%. This is because of terms of trade changes: countries that specialize in sectors with high scale elasticities experience a deterioration of their terms of trade since those sectors expand everywhere thanks to positive production subsidies.[15]The GDP-weighted average gains from this policy are 0.06% – as one should expect, this is very close to the GDP-weighted average gains from industrial policy assuming that countries are in autarky (the 0.04% average gain reported above).

We now explore the robustness of our results in Table 5 regarding two assumptions we have made: that each country is a small open economy, and that there are no scale economies in non-manufacturing sectors. We will also explore the sensitivity of the results in Table 5 with respect to the trade elasticities.

The benefit of assuming that each country is a small open economy is that we can just take our explicit formulas for the optimal production subsidies and export taxes. When countries are large then those formulas are no longer formally correct. To explore the extent to which our small economy assumption affects the results presented in Table 5, we quantify the gains associated with the policies that are optimal for a small economy but now allowing for the full general equilibrium implications. The gains are virtually identical to those that we compute in the small economy case. In ongoing work we are computing the optimal policies when countries are not small – given that general equilibrium considerations do not seem to matter for the gains using the small-country optimal policies, we conjecture that this will not lead to any significant differences either. We tentatively conclude that the gains shown in Table 5 give a remarkably good approximation of the gains that would arise when taking the full general equilibrium implications of trade and industrial policy.

[14]How can we understand the slightly lower gains in column 4 than column 3 for Germany? Imagine again the case in which all production is exported, and assume that there is a subset of sectors for which $(1+\gamma_k)\frac{\theta_k}{1+\theta_k} = (1+\gamma_{k'})\frac{\theta_{k'}}{1+\theta_{k'}}$ for any k, k' in this set. A country that only has employment in sectors in this set would have exactly zero gains from constrained industrial policy as in column 4. In contrast, as long as γ_k and θ_k vary across sectors with positive employment, then a country can gain from both trade and industrial policy in columns 2 and 3.

[15]To confirm this intuition, we computed the correlation between the gains from industrial policy in column 5 and the country-level correlation between sectoral net exports and scale elasticities. The correlation is -0.68.

Table 6: Gains from Industrial Policy with Alternative α_1, Selected Countries

Country	$\alpha_1 = 0.1$ (1)	$\alpha_1 = 0.2$ (2)	$\alpha_1 = 0.3$ (3)	$\alpha_1 = 0.4$ (4)	$\alpha_1 = 0.5$ (5)
United States	0.07%	0.08%	0.09%	0.09%	0.10%
China	0.08%	0.09%	0.10%	0.11%	0.13%
Germany	0.18%	0.19%	0.20%	0.22%	0.25%
Ireland	.09%	0.10%	0.11%	0.14%	0.18%
Vietnam	0.11%	0.11%	0.13%	0.16%	0.20%
Avg, Unweighted	**0.11%**	**0.12%**	**0.13%**	**0.15%**	**0.18%**
Avg, GDP Weighted	**0.10%**	**0.10%**	**0.11%**	**0.12%**	**0.14%**

Notes: Each column reports the gains, expressed as a share of initial real national income, that could be achieved industrial policy under different assumptions regarding α in non-manufacturing sectors. The exercise in each case corresponds to column 3 in Table 5. See the text for detailed descriptions of the exercises.

We next study the implications of our assumption that there are no scale economies in non-manufacturing. We assume that the scale elasticity is positive but common across agriculture and services, and for given values of that elasticity (α_1) we redo the estimation of α_k and the computation of gains from industrial policy as in column 3 of Table 5. The results show that the gains from industrial policy increase a bit relative to those shown in Table 5, going from an average of 0.1% for $\alpha_1 = 0$ to 0.18% for $\alpha_1 = 0.5$. There are two effects operating here. On the hand hand, the heterogeneity in scale elasticities falls slightly as we increase α_1, and this would tend to reduce the gains from industrial policy. On the other hand, the level of those elasticities increases as we increase α_1, and this has the opposite effect. Table 6 reveals that the second effect dominates, but the overall picture remains the same – on average, the gains from industrial policy remain small.

Finally, we explore the sensitivity of our results in Table 5 to the values of trade elasticities. We do so by considering the case in which the trade elasticity is constrained to be the same across sectors. We use the average trade elasticity across the sector-level elasticities estimated in Section 3.4.2, which is equal to 5.1. In Table 4 we show the analogous results to those in Table 5 but now setting $\hat{\theta} = 5.1$ and $\hat{\gamma}_k = \hat{\alpha}_k/\hat{\theta}$, with $\hat{\alpha}_k$ again being those from Section 3.4.1. The lower variation in trade and scale elasticities lead to slightly lower gains from optimal policy, from trade policy, and from industrial policy in columns 1 to 3 compared to those in Table 3, but the overall picture remains the same. Thus, although trade elasticities matter importantly for the implied scale elasticities and for the production subsidies that are part of the optimal policy, they matter little for the implied gains from industrial policy as defined by column 3.

Table 7: Gains from Optimal Policies, Selected Countries, Common TE

Country	Optimal Policy (1)	Classic Trade Pol. (2)	Add Industrial Policy (3)	Constrained Industrial Pol. (4)	Global Efficient Pol. (5)
United States	0.30%	0.26%	0.04%	0.04%	0.07%
China	0.35%	0.31%	0.04%	0.07%	0.00%
Germany	0.66%	0.57%	0.08%	0.04%	-0.22%
Ireland	1.14%	1.08%	0.07%	0.28%	-0.25%
Vietnam	0.95%	0.88%	0.07%	0.27%	0.59%
Avg, Unweighted	**0.69%**	**0.63%**	**0.06%**	**0.12%**	**0.04%**
Avg, GDP Weighted	**0.43%**	**0.39%**	**0.05%**	**0.07%**	**0.00%**

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.

5 Concluding Remarks

Perennial arguments for industrial policy rest on three beliefs. First, that production processes display external economies of scale—such that a nation's productivity in a given sector is increasing in its scale in that sector. Second, that such scale economies differ across sectors—such that any productivity-enhancing expansion of scale in one sector does not just lead to an equal and opposite contraction of productivity in some other sector. And third, that countries produce highly substitutable and tradable goods—such that a country can simultaneously expand scale in one sector without driving down the price of its own output, and find useful foreign alternative versions of the goods in the sector that it chooses to shrink.

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

Our results are sobering. External economies of scale do indeed exist, and do indeed differ substantially across sectors (ranging from an elasticity of 0 to 0.23), but the gains from unilateral industrial policy for all countries in our sample are never particularly large (and equal to just 0.1% of GDP on average across all countries) because countries cannot much expand in attractive sectors without both depressing the price of their goods and forcing consumers to import, often at high trade costs, imperfect substitutes for these goods. By contrast, the gains from unilateral trade policy in our estimated framework are often larger than this (0.6% of GDP on average), reflecting the fact that those gains hinge on the international substitutability of products alone. Interestingly, these two sources of gains to optimal policy design appear largely orthogonal to one another, so most countries could pursue an optimal combination of industrial and trade policy that would be roughly equal to the sum of these two effects (e.g. 0.71% of GDP on average across all countries).

References

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

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

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

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

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

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

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

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

Chor, Davin, "Unpacking Sources of Comparative Advantage: A Quantitative Approach," *forth. Journal of International Economics*, 2010.

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

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

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

Cravino, Javier and Sebastian Sotelo, "Trade induced structural change and the skill premium," Technical Report, National Bureau of Economic Research 2017.

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

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

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

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

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

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

_ , _ , and John Ries, "The erosion of colonial trade linkages after independence," *Journal of International Economics*, 2010, *81* (1), 1–14.

Herrendorf, Berthold, Richard Rogerson, and Akos Valentinyi, "Growth and structural transformation," 2014, *2*, 855–941.

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

Hummels, David and Russell Hillberry, *Trade Elasticity Parameters for a CGE Model* Handbook of Computable General Equilibrium Modeling, Elsevier, 2012.

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

Lashkaripour, Ahmad and Vova Lugovskyy, "National Differentiation and Industry-Wide Scale Effects," *mimeo Indiana University*, 2017.

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

Limão, Nuno, "Preferential trade agreements," in "Handbook of commercial policy," Vol. 1, Elsevier, 2016, pp. 279–367.

Matsuyama, Kiminori, "Agricultural Productivity, Comparative Advantage, and Economic Growth," *Journal of Economic Theory*, 1992, *58*, 317–334.

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

Trefler, Daniel, "Trade liberalization and the theory of endogenous protection: an econometric study of US import policy," *Journal of political Economy*, 1993, *101* (1), 138–160.

A Proofs

A.1 Firm-Level Economies of Scale

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

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

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

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

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

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

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

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

The two previous observations imply that producer prices must satisfy

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

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

A.2 Factor Demand

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

By condition (1), equilibrium quantities and prices must satisfy

$$p_{ij,k}(\omega) = \frac{\tau_{ij,k} w_i}{\alpha_{i,k}(\omega) A_{i,k} E_k^A(L_{i,k})} \text{ if } q_{ij,k}(\omega) > 0. \tag{22}$$

By condition (2), since $U_{j,k}$ is homothetic and taste shocks, $\beta_{ij,k}(\omega) B_{i,k} E_k^A(L_{i,k})$, enter utility multiplicatively, optimal quantities consumed must also satisfy

$$q_{ij,k}(\omega)\beta_{ij,k}(\omega) B_{i,k} E_k^B(L_{i,k}) = \delta_{ij,k}(\{p_{i'j,k}(\omega')/(\beta_{i'j,k}(\omega') B_{i,k} E_k^B(L_{i',k}))\}_{i',\omega'}|\omega) X_{j,k} \tag{23}$$

where $\delta_{ij,k}(\cdot|\omega)$ only depends on $U_{j,k}$ and $\{p_{i'j,k}(\omega')/(\beta_{i'j,k} B_{i,k} E_k^B(L_{i',k})\}_{i',\omega'}$ represents the vector of quality-adjusted prices faced by the representative consumer in destination j and sector k. Combining equation (22) and (23), we can express the share of expenditure, $x_{ij,k} \equiv \sum_\omega p_{ij,k}(\omega) q_{ij,k}(\omega)/X_{j,k}$, in destination j on goods from sector k produced in country i as

$$x_{ij,k} = \sum_\omega \frac{\eta_{ij,k} w_i}{\alpha_{i,k}(\omega)\beta_{ij,k}(\omega) E_k^A(L_{i,k}) E_k^B(L_{i,k})} \delta_{ij,k}(\{p_{i'j,k}(\omega')/(\beta_{i'j,k}(\omega') B_{i,k} E_k^B(L_{i',k}))\}_{i',\omega'}|\omega).$$

Thus, as argued above, we can write

$$x_{ij,k} = \chi_{ij,k}(\eta_{1j,k} w_1/E_k(L_{1,k}), ..., \eta_{Ij,k} w_I/E_k(L_{I,k})),$$

with

$$\chi_{ij,k}(c_{1j,k}, ..., c_{Ij,k}) = \sum_\omega \frac{c_{ij,k}}{\alpha_{i,k}(\omega)\beta_{ij,k}(\omega)} \delta_{ij,k}(\{\frac{c_{i'j,s}}{\alpha_{i',k}(\omega')\beta_{i'j,k}(\omega')}\}_{i',\omega'}|\omega),$$
$$E_k(L_{i,k}) = E_k^A(L_{i,k}) E_k^B(L_{i,k}).$$

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

A.3 Nonparametric Identification

In Section 2.2, we have argued that if there exists a vector of instruments z that satisfies the exclusion restriction, $E[\epsilon|z] = 0$, as well as the completeness condition, $E[g(l)|z] = 0$ implies $g = 0$ for any g with finite expectation, then for any k, E_k is identified, up to a normalization. We now establish this result formally.

Fix i_1, i_2, k_1, k_2, and j. Starting from equation (4), the exclusion restriction implies

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

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

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

By the completeness condition, we therefore have

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

which can be rearranged as

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

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

A.4 Imperfect Competition

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

From our analysis in Appendix A.2, we know that the price at which the retailer from sector k in destination j can buy goods is given by

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

Let $\bar{p}_{ij,k}(\omega)$ denote the price at which the same retailer sells to consumers. From our analysis in Appendix A.2, we also know that the demand of the consumer in destination j for goods from sector k can be expressed as

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

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

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

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

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

The solution to the previous problem must take the form

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

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

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

this implies that

$$x_{ij,k} = \chi_{ij,k}(\eta_{1j,k} w_1/E_k(L_{1,k}), ..., \eta_{Ij,k} w_I/E_k(L_{I,k})),$$

with

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

as argued in the main text.

B Additional Empirical Results