Online Appendix

"Rise and Fall of Empires in the Industrial Era:
A Story of Shifting Comparative Advantages"
Bonfatti and Cosar

A Appendix: Theory

Condition for imperfect specialisation

Suppose $n \geq 2$, and all industrial locations are imperfectly specialised (so that $p_i = 1 \ \forall i$). Then, demand for any variety i cannot exceed $\beta\alpha + \beta(N-2)/(1+\tau^{1-\sigma})$. To see this, note first that, due to symmetric trade costs amongst the industrial locations, aggregate demand for any variety by the industrial locations is $n\beta\alpha/n = \beta\alpha$, irrespective of empire formation. As for aggregate demand by the agricultural locations, this is largest if 1) i's empire is the largest possible, that is it includes all agricultural workers; and 2) it is $n = 2$. This can be seen by noticing that aggregate demand for variety i by the agricultural locations can be written as

$$\beta\left(\frac{E_i}{1+(n-1)\,\tau^{1-\sigma}} + \frac{E_{-i}\tau^{1-\sigma}}{1+(n-1)\,\tau^{1-\sigma}} + \frac{N-n-E_i-E_{-i}}{n}\right) < \beta\frac{E_i + E_{-i} + N - n - E_i - E_{-i}\tau^{1-\sigma}}{1+(n-1)\,\tau^{1-\sigma}}$$
$$< \beta\frac{N-n}{1+(n-1)\,\tau^{1-\sigma}} \qquad (11)$$
$$\leq \beta\frac{N-2}{1+\tau^{1-\sigma}}. \qquad (12)$$

where (11) and (12) are demand when i's empire includes all agricultural workers, and (12) is the case $n = 2$. Then, $\alpha > \beta\alpha + \beta(N-2)/(1+\tau^{1-\sigma})$, or

$$\alpha > \frac{\beta}{1-\beta}\frac{N-2}{1+\tau^{1-\sigma}}$$

is a sufficient condition for imperfect specialisation, since it implies that even with maximum demand, industrial locations remains imperfectly specialised. □

Proof that $g(n,\sigma)$ is increasing in σ for $n > 1$

We show that $[(\tau^{\sigma-1}-1+n)/n]^{\frac{\beta}{\sigma-1}}$, or equivalently $\beta/(\sigma-1)\log[(\tau^{\sigma-1}-1+n)/n]$, is increasing in σ. A sufficient condition for this is that the first derivative of the latter with

respect to σ be greater than zero, or

$$-\frac{\beta}{(\sigma-1)^2}\left[\log\left(\tau^{\sigma-1}+n-1\right)-\log\left(n\right)\right]+\frac{\beta}{\sigma-1}\frac{\tau^{\sigma-1}\log\tau}{\tau^{\sigma-1}+n-1}>0$$
$$-\left[\log\left(\tau^{\sigma-1}+n-1\right)-\log\left(n\right)\right]+\frac{\tau^{\sigma-1}\log\tau^{\sigma-1}}{\tau^{\sigma-1}+n-1}>0 \tag{13}$$

But condition (13) holds for any $\sigma>1$. To see this, note that, for $\sigma=1$, it is $\tau^{\sigma-1}=1$, and the left-hand size is equal to zero. At the same time, for $\sigma>1$, $\tau^{\sigma-1}$ is increasing in σ, and the left-hand side is increasing in $\tau^{\sigma-1}$. A sufficient condition for the latter is, using the simplified notation $a\equiv\tau^{\sigma-1}$,

$$\begin{aligned}\frac{d\left[-\log\left(a+n-1\right)+\frac{a\log a}{a+n-1}\right]}{da}&=-\frac{1}{a+n-1}+\frac{\left(\log a+1\right)\left(a+n-1\right)-a\log a}{\left(a+n-1\right)^2}\\&=\frac{\log a\left(a+n-1\right)-a\log a}{\left(a+n-1\right)^2}\\&=\frac{\log a\left(n-1\right)}{\left(a+n-1\right)^2}>0\end{aligned}$$

which is true given $n>1$ and $a>1$. □

Proof to Proposition 1

Point 1 follows by definition, since empires must contain one industrial location, and a positive measure of workers from agricultural locations. Suppose then $0<n<N$. The sum of individual utilities can be written as a term which captures the aggregate utility of industrial locations (and which does not depend on empire formation), plus the aggregate utility of the agricultural locations, that is

$$\beta^{\beta}(1-\beta)^{1-\beta}\frac{1}{\left[n(\delta\tau)^{1-\sigma}\right]^{\frac{\beta}{1-\sigma}}}\left(N-n+\sum_{i=1}^{n}\int_0^{E_i}\left[g(n,\sigma)-c(x_i)\right]dx_i\right). \tag{14}$$

For any i, consider the planner's optimal choice of E_i. This must satisfy

$$\begin{aligned}&\max_{E_i}\int_0^{E_i}\left[g(n,\sigma)-c(x_i)\right]dx_i \qquad (15)\\&\text{s.t.}\quad E_i\geq 0.\end{aligned}$$

Given that $c(\cdot)$ is continuous and increasing, the maximand is strictly concave in E_i. Then, the necessary conditions for a maximum, $g(n,\sigma)-c(E_i)+\lambda_i=0$ and $E\lambda_i=0$ (where λ_i is the

Lagrange multiplier), are also sufficient. If $g(n,\sigma) - c(0) \leq 0$, then the necessary conditions require $\lambda_i \geq 0$ and $E_i = 0$. Otherwise, they require $\lambda_i = 0$ and $E_i \equiv \arg_y [g(n,\sigma) = c(y)]$. Since the planner solves the same independent problem for any i, empires must be equally sized at the planner's optimum. $\square$

Proof to Proposition 2

The surplus from industrial location i's empire can be written as

$$\beta^{\beta}(1-\beta)^{1-\beta}\frac{1}{[n(\delta\tau)^{1-\sigma}]^{\frac{\beta}{1-\sigma}}}\int_0^{E_i}[g(n,\sigma)-c(x_i)]dx_i.$$

Thus, the industrial location's problem is

$$\begin{aligned}&\max_{E_i}\int_0^{E_i}[g(n,\sigma)-c(x_i)]dx_i\\ \text{s.t.}\quad &E_i \geq 0\end{aligned}$$

which is independent of the other industrial locations' choice of empires, and coincides with the Social Planner problem. Thus, the Nash equilibrium of the game in which the industrial locations simultaneously choose the size of their empires must be the same as the Social Planner's optimum. $\square$

Proof to Proposition 3

If $n = 0$, then $N^*(n,\sigma) = N$ follows from equation (10). Suppose then $n > 0$. There exists $\overline{n} > 1$ such that $g(n,\sigma) \geq c(1)$ if $0 < n < \overline{n}$, and $g(n,\sigma) < c(1)$ if $n \geq \overline{n}$. This follows from the fact that, by Assumption 2, $g(1,\sigma) = \tau^{\beta} - 1 > c(1) > 0$, and by the fact that $g(n,\sigma)$ decreases to zero as n increases to infinity. That $\overline{n} \leq N$ is implied by assumption (9). To see this, note that, by construction, $g(\overline{n}-1,\sigma) > c(0)$. Then, by the said assumption, it is $N > (\overline{n}-1)[E^*(\overline{n}-1,\sigma)+1] > \overline{n}-1$, which proves the point. By Proposition 1 and 2, it follows that $\underline{E}^*(n,\sigma) \geq 1$ if $0 < n < \overline{n}$, and $\underline{E}^*(n,\sigma) = 0$ if $n \geq \overline{n}$. That $N^*(n,\sigma) < N$ in the first case and $N^*(n,\sigma) = N$ in the second case then follows from equation (10). Finally, the last sentence of the proposition follows from the fact that $g(n,\sigma)$—and hence, by Proposition 1, $E^*(n,\sigma)$—is increasing in σ, and $\underline{E}^*(n,\sigma)$ is either increasing or constant in $E^*(n,\sigma)$. $\square$

B Appendix: Data and Historical Evidence

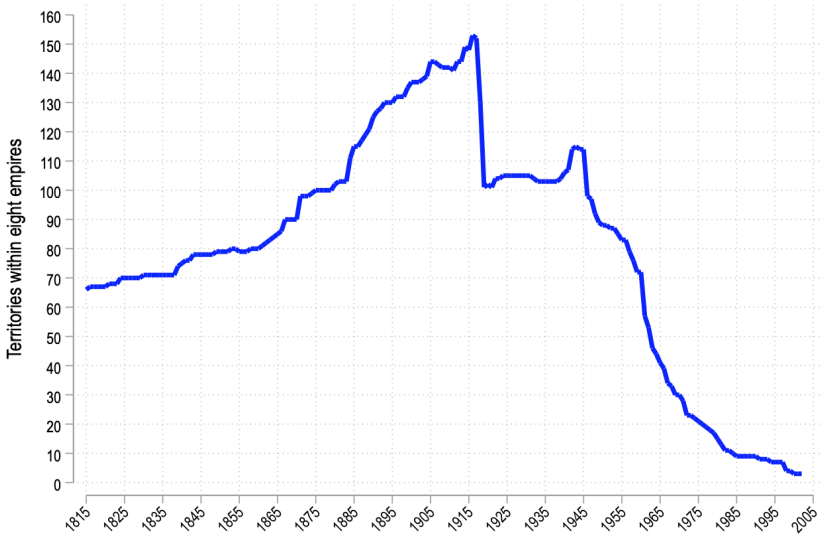
- Figure Ia (Number of sovereign entities and size of empires): The GeoPolHist database (Dedinger and Girard, 2021) provides a comprehensive list of geopolitical entities in the world and their political status (sovereign, colony, dependency etc) since 1815. It improves upon the Correlates of War (CoW) dataset in three aspects. First, CoW accounts for an entity only if it appears in a conflict. Second, CoW imposes a threshold of 500,000 people to count as a political entity. Finally, CoW records as sovereign only those entities in which both Britain and France established diplomatic missions after 1815. GeoPolHist database relaxes all there constraints and therefore provides the list and status of geopolitical entities, and their sovereignty status in each year.

  To generate the time series on the number of sovereign entities, we keep from the baseline GeoPolHist database those that are coded as *sovereign, sovereign (limited)*, or *sovereign (unrecognized)* for the period before 1945. We exclude sovereign entities that have eventually formed the unified Germany and Italy, and by doing so, we count both of these countries *as if* they existed throughout the data period. The qualitative pattern of the U-shape of number of sovereign entities remains unchanged if we separately include pre-unification German and Italian states. For the post WW2 era, we drop unrecognized entities to avoid counting separatist regions such as the Donetsk People's Republic as independent.

- Figure IIa (Per capita income dispersion): Using the Maddison dataset (Bolt and van Zanden, 2020), we construct a balanced sample of 57 countries/regions that are comparable over time and report per capita real income for at least one year in each decade.

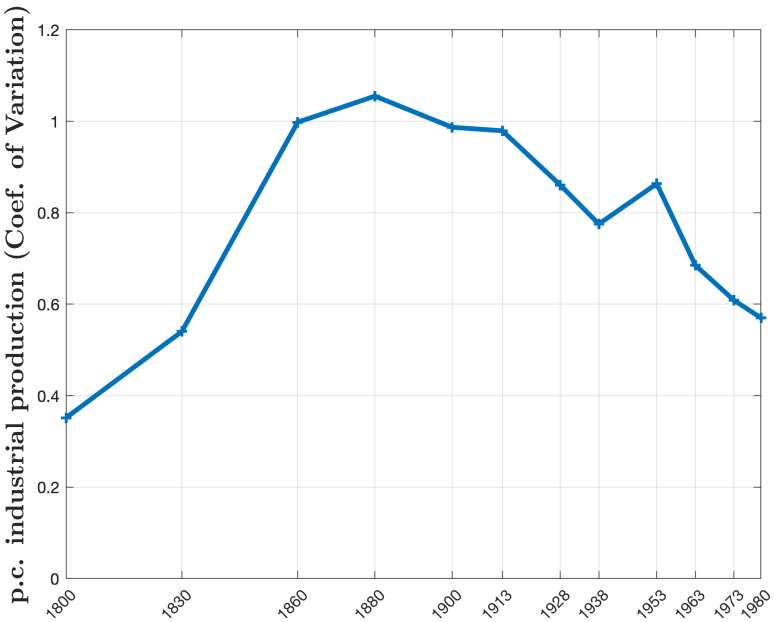
- Figure IIIa (IIT index and import share of primary products): Both series use the NBER-UN world trade data (Feenstra et al., 2005) after 1962. Aggregate manufacturing imports and exports until 1933 are digitized from Mitchell (1988) by the authors and the share of primary products in UK imports between 1850-1938 is based on the Federico and Tena Junguito (2019) dataset. Both the pre-1938 and post-1962 data define primary products as SITC sections 0-4. For manufacturing IIT index after 1962, we use aggregate UK imports and exports of SITC sections 5-8. Manufacturing trade data before 1938 is from table 19 (total finished manufactures) in Mitchell (1988).

Figure B.1: **Size of Empires - Alternative Dataset**



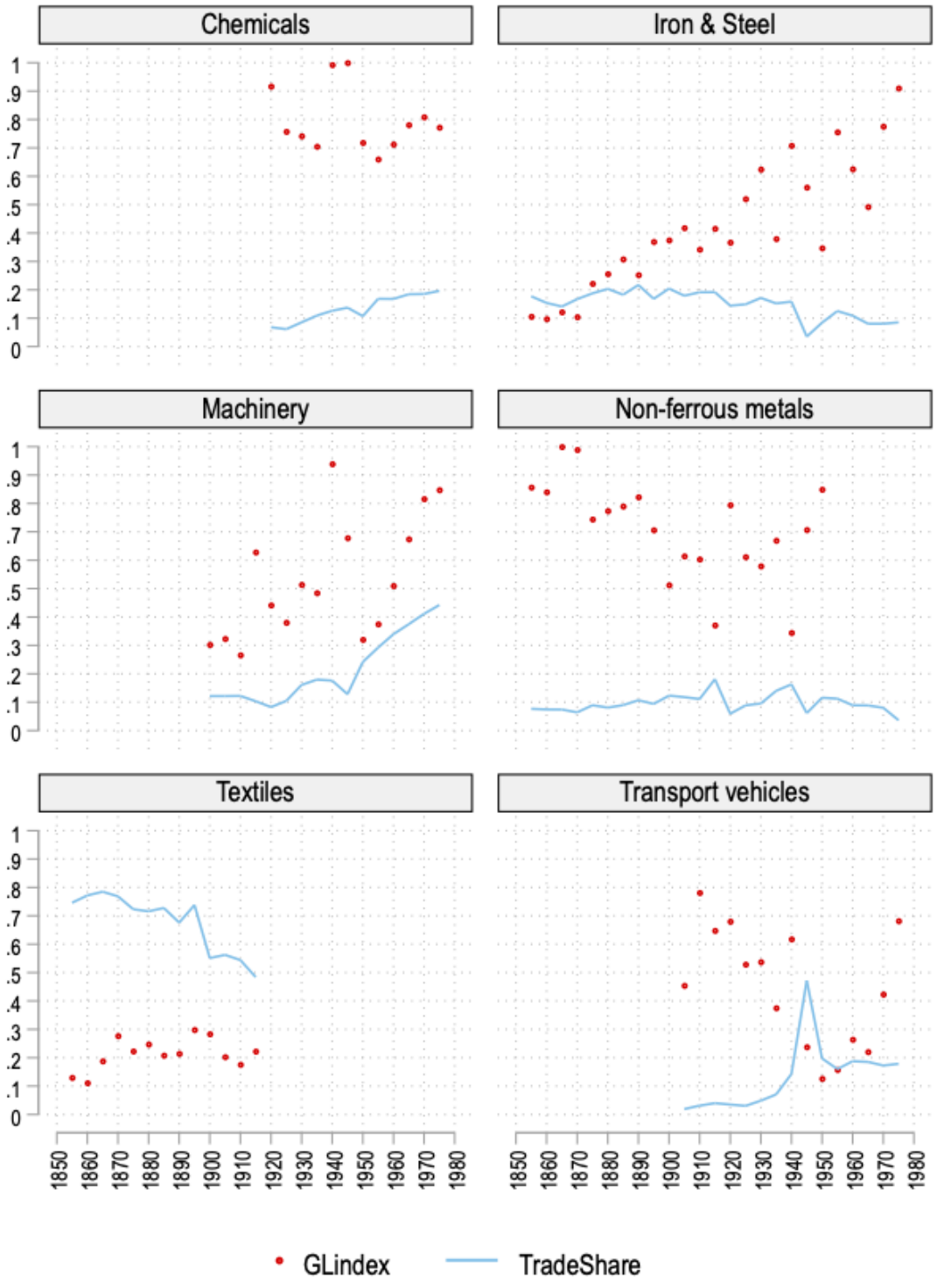
Notes: Source is data on empires and the territories they control by Gokmen, Vermeulen and Vézina (2020).

Figure B.2: **Dispersion of Industrial Production**



Notes: Coefficient of variation in per capita industrial output across a balanced sample of 29 countries/regions comparable over time, constructed by the authors from Bairoch (1982).

Figure B.3: **Sectoral Shares and GL Indices of Intra-industry Trade**



C Appendix: Numerical Comparative Statics

Calibration

We set the number of locations to $N = 200$ to match the scale of the empirical number of countries (Figure Ia). We let $\beta = 0.1$ corresponding to low shares of modern industry and urban populations in early to mid-19th century. In terms of historical realism, one could also specify β to be increasing over time due to demand- and supply-driven structural change. Since not doing so does not affect the qualitative patterns that we demonstrate, we abstract from this channel in order to keep the exposition focused. Using a gravity framework, Jacks, Meissner and Novy (2010) estimate trade costs between 1870-1913, and the reduction thereof within the British empire. In backing out the level of trade costs, they assume an elasticity of substitution equal to 11. More recent elasticity estimates are lower and lie around 5. Consistent with these results, we set $\tau = 1.6$ and $\delta = 1.3$, and the lower and upper bounds of the elasticity of substitution as $\underline{\sigma} = 5$ and $\overline{\sigma} = 11$, respectively.

We set the productivity term α to satisfy the imperfect specialisation condition. Given $(\beta, N, \tau, \overline{\sigma})$, the right hand side of equation (3) equals 21.8. We let $\alpha = 22$. A high sectoral TFP gap between industrialised countries and the agricultural periphery in early to mid-$19th$ century is realistic. In the Maddison data (Bolt and van Zanden, 2020), the ratio of the highest to lowest income in 1865 is about 18. This is likely to be a lower bound for the actual gap between the frontier and locations living on subsistence agriculture since countries with historical statistics that make it into the Maddison dataset are a selected group with relatively higher income levels. Even today, after decades of technology diffusion, calibrated TFP gaps to rationalize per capita income differences are in this order of magnitude (Klenow and Rodriguez-Clare, 1997). Moreover, condition (3) for imperfect specialisation is derived under the assumption that all locations have equal population. For $n = 2$ and $N = 200$, this implies a population share of 1%. In reality, world the population share of the first two industrial imperial powers, England and France, was consistently about 5% throughout the $19th$ century (Joerg et al., 2014). Accounting for this would have made them even more capable in satisfying the demand for industrial goods from the rest of world without perfectly specialising in manufacturing, therefore relaxing the minimum bound on α under the sufficient condition (3). To keep the analysis parsimonious, we do not model heterogeneous population growth across locations.

We chose the governance cost parameters (γ_1, γ_2) by exactly fitting two empirical moments. First, when $n = 1$, the first empire governs 25% of the world, which was Britain's share of world area at its peak. Second, when $t = 9$ such that $n = 9$ and $\sigma_t = \underline{\sigma}$, we have $E^* = 0$. That is, when there are 9 industrialised locations and the elasticity of substitution

is at its lower bound, empires dissolve. This choice implies the emergence of eight industrial empires, corresponding to the number of empires whose total size we plotted in Figure Ia.

Next, we show how to solve two equations recursively to back out these two parameters. The first condition is straightforward. For $n = 1$, we have $E^*(n = 1) = \frac{\tau^\beta - 1 - \gamma_1}{\gamma_2}$, so

$$\frac{\tau^\beta - 1 - \gamma_1 + \gamma_2}{\gamma_2 \cdot N} = 0.25,$$

and the second condition implies

$$\gamma_1 \geq \left(\frac{\tau^{\underline{\sigma}-1} + \tilde{n} - 1}{\tilde{n}} \right)^{\frac{\beta}{\underline{\sigma}-1}} - 1.$$

Given $(N = 200, \beta = 0.1, \tau = 1.6, \underline{\sigma} = 5, \tilde{n} = 8)$, we solve both equations with equality, which yields $\gamma_1 = 0.0133$ and $\gamma_2 = 0.00071$.

Figure C.1: **Size of Empire Comparative Statics**

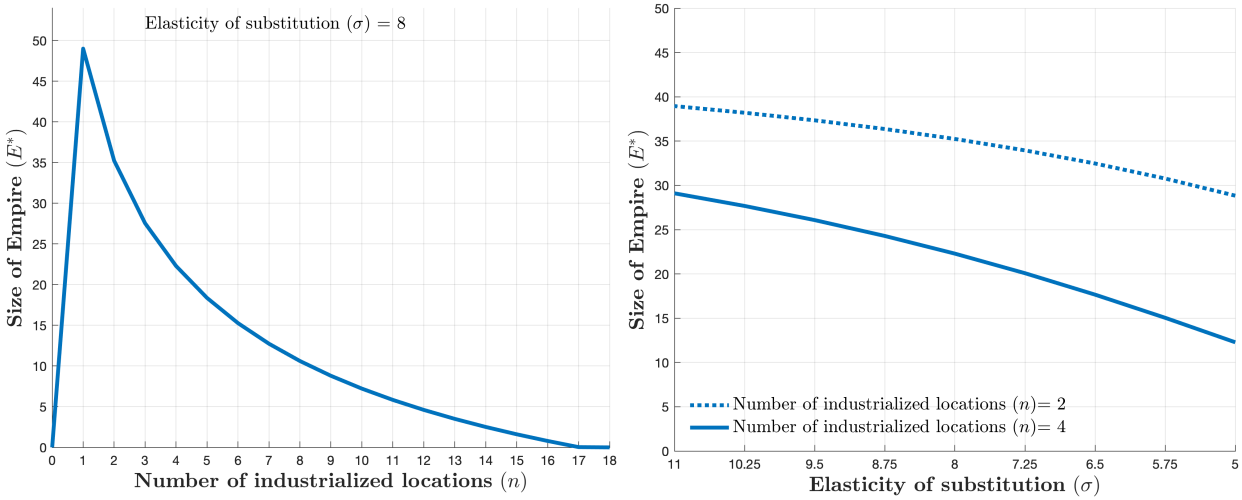


Figure C.2: **Number of Countries Comparative Statics**

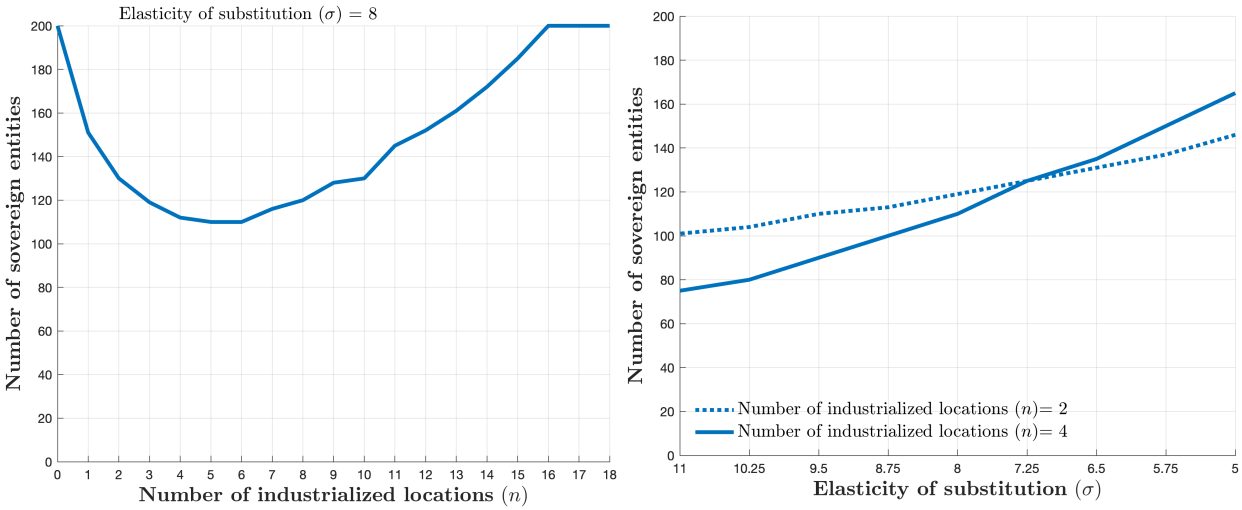


Figure C.3: **Income Dispersion Comparative Statics**

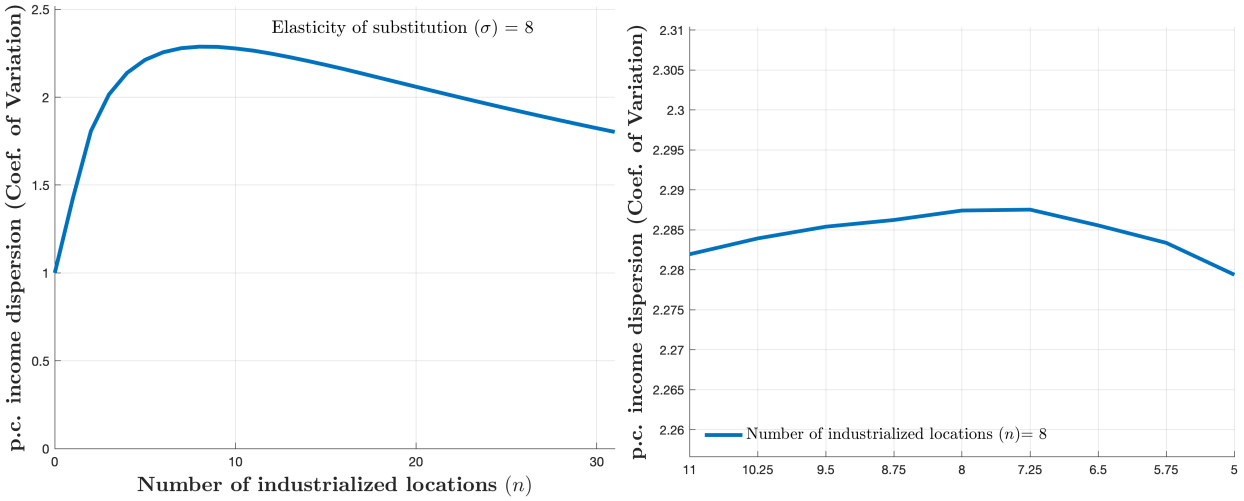


Figure C.4: **GL Index of Intra-industry Trade Comparative Statics**