From Mercantilism to Free Trade: Online Appendix

Didac Queralt
Carlos III University, Carlos III-Juan March Institute of Social Sciences
dqueralt@march.es

January 2015

Online Appendix

The online Appendix includes

1. Proofs
2. Costly and Imperfect Monopoly Enforcement extension.
3. An alternative investment goal for fiscal capacity.
4. The relaxation of the monopoly assumption in favor of an oligopoly market.
5. Further data details.
6. Robustness test for Hypothesis 1.

I. Proofs

Proof of *Proposition 1*

Proof. The ruler utility function in (7) is a concave function of τ, $\partial V/\partial\tau > 0$, $\partial V/\partial^2\tau < 0$. $V(\tau)$ is maximized for τ^u, the stock of fiscal capacity necessary to implement the unconstrained tax rate. Suppose $\tau_p^* = \tau^u$ and A_l. Then, $V(\tau_p^*)$ defines a horizontal curve in the $V - \tau$ space. When $\tau \to 0$, $\tau_e^* = 0$. As long as (10) is satisfied, $V(\tau_e^*) < V(\tau_p^*)$. When $\tau \to \tau^u$, $\tau_e^* = \tau^u$. Since $A_h > A_l$, then $V(\tau_e^*) > V(\tau_p^*)$. Thus, by the Intermediate Value Theorem, it must be true that there exists an unique $\hat{\tau}$, $0 < \hat{\tau} < \tau^u$ such that $\forall\tau \leq \hat{\tau}$, $V(\tau_p^*) \geq V(\tau_e^*)$. ■

Proof of *Lemma 1* The ruler utility is a function of wages and per capita public spending. A sufficient condition for the existence of the ruler's inter-temporal dilemma is $G(I = 1) \leq G(I = 0)$ and $w(I = 1) \leq w(I = 0)$. That is the case if $\bar{\delta} \leq \alpha + (1 - \alpha)/\rho$.

Proof. Per capita public spending is defined by $(1-\sigma)G/L$, with $G \equiv T = \tau p x$. If no investment in fiscal capacity takes place, $\sigma = 0$, and τ, p and x are defined by *Proposition 1*. If investment takes place, $\sigma > 0$, and τ, p and x are defined by (18), (15) and (16), respectively. Upon substitution,

$$\begin{aligned} G^*(I=1) = &\quad (1-\sigma)(1-\alpha-\tfrac{1-\alpha}{\rho(1-\sigma)})A_l L[\alpha\delta(\alpha+\tfrac{1-\alpha}{\rho(1-\sigma)})]^{\frac{1}{1-\alpha}} \\ G^*(I=0) = &\quad (1-\alpha-\tfrac{1-\alpha}{\rho})A_l L[\alpha+\tfrac{1-\alpha}{\rho})]^{\frac{1}{1-\alpha}} \end{aligned} \tag{1}$$

For $G(I=1) < G(I=0)$, it must be true that

$$(1-\sigma)\delta^{\frac{\alpha}{1-\alpha}}\left[\frac{\rho(1-\sigma)-1}{\rho-1}\right] < \left[\frac{\alpha+\frac{1-\alpha}{\rho}}{\alpha+\frac{1-\alpha}{\rho(1-\alpha)}}\right]^{\frac{\alpha}{1-\alpha}} \tag{2}$$

For all $\alpha \in (0,1)$, $\delta < 1$ and $\rho > 1$, the left-hand side of condition (2) is a decreasing convex function of σ that cuts the vertical axis ($\sigma = 0$) at $\delta^{\frac{\alpha}{1-\alpha}}$, and cuts the horizontal axis at $\sigma = 1$. The right-hand side of condition (2) is a negative concave function that cuts the vertical axis at $1 > \delta^{\frac{\alpha}{1-\alpha}}$, and the horizontal curve at $\sigma = 1$. Thus, $\forall \sigma \in (0,1)$, $G(I=1) < G(I=0)$.

Wages depend on δ and τ. When no technology innovation takes place, $\delta = 1$; when innovation takes place, $\delta < 1$. Given $\tau(I=1)$ and $\tau(I=0)$, as defined by (18) and (9),

$$\begin{aligned} w^*(I=1) = &\quad \tfrac{1-\alpha}{\alpha}(\alpha\delta)^{\frac{\alpha}{1-\alpha}}[\alpha+\tfrac{1-\alpha}{\rho(1-\sigma)}]^{\frac{\alpha}{1-\alpha}} \\ w^*(I=0) = &\quad \tfrac{1-\alpha}{\alpha}(\alpha)^{\frac{\alpha}{1-\alpha}}[\alpha+\tfrac{1-\alpha}{\rho}]^{\frac{\alpha}{1-\alpha}} \end{aligned} \tag{3}$$

For $w^*(I=1) < w^*(I=0)$, it must be true that

$$\delta < \bar{\delta} = \frac{\alpha+\frac{1-\alpha}{\rho}}{\alpha+\frac{1-\alpha}{\rho(1-\sigma)}} \tag{4}$$

Since $\sigma \leq 1 - 1/\rho$, this implies $\bar{\delta} \leq \alpha + (1-\alpha)/\rho < 1$. If innovation cost satisfies this condition, wages following induced-innovation are lower than those without investment in fiscal capacity. ■

Proof of *Proposition 2*

Proof. Let $\sigma \in (0, \bar{\sigma})$. Investment is preferred when $\sum_s^2 V_s(I=1,\delta) \geq \sum_s^2 V_s(I=0)$. After some rearrangement, this implies

$$\begin{aligned}
&A_l(\alpha + \tfrac{(1-\alpha)}{\rho(1-\sigma)})^{\frac{\alpha}{1-\alpha}}\left[\tfrac{1-\alpha}{\alpha}\delta^{\frac{\alpha}{1-\alpha}} + \rho(1-\sigma)\big(1-\alpha-\tfrac{(1-\alpha)}{\rho(1-\sigma)}\big)\right] \\
\geq\ & (2A_l - A_h)(\alpha + \tfrac{1-\alpha}{\rho})^{\frac{\alpha}{1-\alpha}}\left[\tfrac{1-\alpha}{\alpha} + \rho\big(1-\alpha-\tfrac{(1-\alpha)}{\rho}\big)\right]
\end{aligned} \tag{5}$$

The left-hand side (LHS) of expression (5) is a negative convex function of σ, whereas its right-hand side (RHS) is a horizontal curve. We need to prove that they cut within $\sigma \in (0, \hat{\sigma})$ for investment to take place.

When $\sigma \to \bar{\sigma}$, $\tau^*(I = 1) \to 0$. Thus, there is no investment and no induced innovation. For the RHS of (5) to be greater than the LHS, all it is required is $A_h < 2A_l$. Normalize A_l to 1, so $A_h < 2$.

When $\sigma \to 0$,

$$A_l(\alpha + \frac{1-\alpha}{\rho})^{\frac{\alpha}{1-\alpha}}\Big[\frac{1-\alpha}{\alpha}\delta^{\frac{\alpha}{1-\alpha}} + (1-\alpha)(\rho-1)\Big] \geq (2A_l - A_h)\Big[(\alpha + \frac{1-\alpha}{\rho})^{\frac{\alpha}{1-\alpha}} + (1-\alpha)(\rho-1)\Big] \tag{6}$$

For the LHS of (6) to be greater than the RHS, it must be the case that $A_h/A_l > (2 - \delta^{\frac{\alpha}{1-\alpha}})$. With A_l normalized to 1,

$$A_h > (2 - \delta^{\frac{\alpha}{1-\alpha}})$$

Notice that $\delta^{\frac{\alpha}{1-\alpha}} < 1$ and $A_h > 1$. Hence, $(2 - \delta^{\frac{\alpha}{1-\alpha}}) < A_h < 2$ is non-empty.

If $A_h \in [2 - \delta^{\frac{\alpha}{1-\alpha}}, 2]$, by the Intermediate Point Theorem there exists a unique $\hat{\sigma} \in (0, \bar{\sigma})$ such that $\forall \sigma < \hat{\sigma}$ investment is preferred, and $\forall \sigma \in [\hat{\sigma}, \bar{\sigma})$, no investment ever takes place. ■

Proof of *Lemma 2*

Proof. For $w_s(I)$ and $G_s(I) = t_s p_s x_s$ in *Propositions 1 and 2*, $\sum_s^2 \Pi_s(I = 0) > \sum_s^2 \Pi_s(I = 1, \delta)$

reduces to

$$(2A_l - A_h)\Big(\alpha + \frac{1-\alpha}{\rho}\Big)^{\frac{1}{1-\alpha}}\Big[\frac{1}{\rho} - 1\Big] > A_l\Bigg[\delta\Big(\alpha + \frac{1-\alpha}{\rho}\Big)\Bigg]^{\frac{1}{1-\alpha}}\Big[\frac{1}{\delta\rho} - 1\Big] \tag{7}$$

The RHS of (7) is an positive function of δ, while the LHS of (7) is independent of it. To guarantee that both curves cut in the $\delta \in [0, \bar{\delta}]$ interval, two conditions to be satisfied:

(i) As $\delta \to 0$, $LHS > RHS$. Notice that $RHS(\delta = 0) = 0$, which is clearer when we rearrange the RHS as:

$$A_l\Bigg[\Big(\alpha + \frac{1-\alpha}{\rho}\Big)\Bigg]^{\frac{1}{1-\alpha}}\delta^{\frac{\alpha}{1-\alpha}}\Big[\frac{1}{\rho} - \delta\Big]$$

As the $LHS > 0$, then $LHS > RHS(\delta = 0)$.

(ii) As $\delta \to \bar{\delta}$, the latter being defined by (20), $LHS < RHS$. Notice that $\bar{\delta}$ is largest when $\sigma = \bar{\sigma}$. Plugging $\bar{\sigma}$ into $\bar{\delta}$, and replacing δ for $\bar{\delta}$ in (7), we reach

$$(2A_l - A_h)[\frac{1}{\alpha} - 1] < A_l\Big[\frac{\rho}{\alpha(1 + \alpha(\rho - 1))} - 1\Big] \tag{8}$$

Since $A_l > (2A_l - A_h)$, all we need is the element in brackets multiplying A_l to be greater than the one multiplying $(2A_l - A_h)$, which is always satisfied.

Since $LHS > RHS$ for $\delta = 0$ (lowest), and $LHS < RHS$ for $\delta = \bar{\delta}$ (highest), by the Intermediate Point Theorem, there exists $\tilde{\delta} < \bar{\delta}$ such that, $\forall \delta < \tilde{\delta}$ the producer has an incentive to bribe, and none for $\delta \geq \tilde{\delta}$. ■

Proof of *Proposition 3* First, I prove that a $\tilde{\sigma} \in (0, 1)$ exists such that, $\forall \sigma > \tilde{\sigma}$, the *status quo* (i.e. mercantilism) is preferred. Second, I prove that $\tilde{\sigma} < \hat{\sigma}$, the latter being defined in *Proposition 2*.

Proof. The ruler incentive constraint can be re-expressed as

$$\sum_{s}^{2}\Big(\Pi_s(I = 0) + V_s(I = 0) - \big[\Pi_s(I = 1) + V_s(I = 1)\big]\Big) \geq 0$$

Plugging in all equilibrium values we get

$$\begin{aligned}&(2A_l - A_h)\big(\alpha + \tfrac{(1-\alpha)}{\rho}\big)^{\frac{\alpha}{1-\alpha}} \left\{\alpha\big(\alpha + \tfrac{1-\alpha}{\alpha}\big)L\big(\tfrac{1}{\alpha}+1\big) + \left[\tfrac{1-\alpha}{\alpha} + (1-\alpha)(\rho-1)\right]\right\} \\ \geq\ &(A_l)\delta^{\frac{\alpha}{1-\alpha}}\big(\alpha + \tfrac{1-\alpha}{\rho(1-\sigma)}\big)^{\frac{\alpha}{1-\alpha}} \left\{\alpha\delta\big(\alpha + \tfrac{1-\alpha}{\rho(1-\sigma)}\big)L\big(\tfrac{1}{\delta\alpha}-1\big) + \left[\tfrac{1-\alpha}{\alpha} + (1-\alpha)(\rho(1-\sigma))\right]\right\}\end{aligned} \tag{9}$$

Normalize $L = 1$.

Step 1. Let $\sigma \to \bar{\sigma}$, with $\bar{\sigma}$ defined by (19). Then, $t^* = 0$, which inhibits investment (i.e. $\tau_2 = \tau_1 = \tau$) and, as a direct consequence, induced innovation too (i.e. $A_1 = A_2 = A_l$). For this set of parameters, the ruler would always prefers to stick to the *status quo* (see *Proposition 4*).

Let $\sigma \to 0$; then for the right-hand side of expression (9) to be bigger than the left-hand side, it suffices with $A_h \geq 2 - \delta^{\frac{\alpha}{1-\alpha}}$.

By the Intermediate Value Theorem, there exists a $\tilde{\sigma} \in (0, \bar{\sigma})$ such that, $\forall \sigma > \tilde{\sigma}$, the ruler always prefers the *status quo*.

Step 2. Let's now compare $\tilde{\sigma}$ with $\hat{\sigma}$ (*Proposition 2)*. The latter is implicitly defined in the ruler's original problem

$$\left[\alpha + \frac{1-\alpha}{\rho(1-\sigma)}\right]^{\frac{\alpha}{1-\alpha}} = \frac{(2A_l - A_h)(\alpha + \frac{1-\alpha}{\rho})^{\frac{\alpha}{1-\alpha}}\left[\frac{1-\alpha}{\alpha} + \rho\big(1-\alpha-\frac{(1-\alpha)}{\rho}\big)\right]}{A_l\delta^{\frac{\alpha}{1-\alpha}}\left[\frac{1-\alpha}{\alpha} + \rho\big(1-\alpha-\frac{(1-\alpha)}{\rho}\big)\right]} \tag{10}$$

while $\tilde{\sigma}$ is implicitly defined in the ruler incentive constraint in (9)

$$\left[\alpha + \tfrac{1-\alpha}{\rho(1-\sigma)}\right]^{\frac{\alpha}{1-\alpha}} = \frac{(2A_l - A_h)\Big(\alpha + \frac{(1-\alpha)}{\rho}\Big)^{\frac{\alpha}{1-\alpha}}\left\{\alpha\Big(\alpha+\frac{1-\alpha}{\alpha}\Big)L\Big(\frac{1}{\alpha}+1\Big) + \left[\frac{1-\alpha}{\alpha}+(1-\alpha)(\rho-1)\right]\right\}}{(A_l)\delta^{\frac{\alpha}{1-\alpha}}\left\{\alpha\delta\Big(\alpha+\frac{1-\alpha}{\rho(1-\sigma)}\Big)L\Big(\frac{1}{\delta\alpha}-1\Big) + \left[\frac{1-\alpha}{\alpha}+(1-\alpha)(\rho(1-\sigma))\right]\right\}} \tag{11}$$

The left-hand side expressions in (10) and (11) are the same one. And this is an increasing function of σ. The right-hand side of both expressions, on the other hand, are independent of σ. Now, I seek to know which of the two horizontal curves cuts first the left-hand side. Let

$$\begin{aligned} M &= \tfrac{1-\alpha}{\alpha} + (1-\alpha)(\rho - 1) \\ N &= \tfrac{1-\alpha}{\alpha} + (1-\alpha)(\rho(1-\sigma)) \\ X &= (1-\alpha)(\alpha + \tfrac{1-\alpha}{\rho}) \\ Y &= (1-\delta\alpha)(\alpha + \tfrac{1-\alpha}{\rho(1-\sigma)}) \end{aligned}$$

Given $M,\ N,\ X,\ Y$, $\tilde{\sigma} < \hat{\sigma}$ whenever

$$F1 = \frac{X+M}{Y+N} > \frac{M}{N} = F2$$

This is true if

$$\begin{aligned} &\left[\tfrac{1-\alpha}{\alpha} + (1-\alpha)(\rho(1-\sigma))\right] \times \left[(1-\alpha)(\alpha + \tfrac{1-\alpha}{\rho})\right] \\ &> \left[\tfrac{1-\alpha}{\alpha} + (1-\alpha)(\rho - 1)\right] \times \left[(1-\delta\alpha)(\alpha + \tfrac{1-\alpha}{\rho(1-\sigma)})\right] \end{aligned}$$

which is true for all $\sigma \in [0,1]$ and $\delta \in [0,1]$ (thus, satisfying the producer participation constraint). Since $F2$ is first-order dominated by $F1$, $\tilde{\sigma} < \hat{\sigma}$ is always true.

■

II. Costly and Imperfect Monopoly Enforcement

The set up in the core text implicitly assumes that the government is capable of enforcing the domestic monopoly at no cost. This is a simplifying assumption. Next, this assumption is relaxed. Monopoly enforcement requires some degree of bureaucratic capacity, which is itself costly. This cost implies that only a share $\kappa \in [0,1]$ of total revenue actually reaches the putative recipient of public spending (i.e. labor). The remaining share, $1-\kappa$, is spent either in public clerks' salaries, customs buildings, or is even captured by corrupt officials. Without loss of generality, $1-\kappa$ can be interpreted as the sunk cost of taxation derived from costly monopoly enforcement.

We seek to investigate whether this sunk cost unravels the mercantilist equilibrium. In order to do that, we must re-express public spending as $G = \kappa T$. Accordingly, the new ruler's problem is

$$\begin{aligned} \max_t V = & \quad \theta\Big[\omega(t, x_j^*|\phi_j) + \rho\frac{\kappa t p_j^* x_j^*}{L}\Big] + (1-\theta)\pi(t, x_j^*|\phi_j) \\ \text{s.t.}\, & t \leq \tau \end{aligned}$$

where κ premultiplies public good provision in the second element in the first bracket. The new unconstrained equilibrium tax rate $t_{\lambda=0,\kappa}$ becomes

$$\frac{\theta\kappa\rho - 1}{\theta(1 + \frac{\kappa\rho}{1-\alpha}) - 1} < 1$$

which requires $\theta\rho\kappa > 1$. Recall, this condition ensures that the ruler is interested in some positive taxation, which is the case when the ruler cares about labor's welfare, or labor attach high valuations to public spending, or both.

Since $\partial(t_{\lambda=0,\kappa}/\partial\kappa) > 0$, the new unconstrained tax rate is lower than in the benchmark case (as defined in Expression 9). The reason is that the inefficiencies in public good provision reduce the value of public spending relative to market-clearing wages, which, recall, decrease in the tax rate.

Given the unconstrained tax rate, it follows from *Proposition 1* that mercantilism will be an equilibrium only if

$$\frac{A_h}{A_l} < \frac{\theta(1+\alpha(\kappa\rho+1))}{\theta(1-\alpha)+\alpha}\left[\frac{\theta(1+\alpha(\kappa\rho+1))}{\theta(1-\alpha+\kappa\rho)-(1-\alpha)}\right]^{\frac{\alpha}{1-\alpha}} \tag{12}$$

that is, if the technology distance between the incumbent and would-be producer is limited. The right-hand side of (12) is increasing in κ. Thus, the larger the inefficiencies in providing public goods are, the less likely condition (12) is met. This is due to the marginal rate of substitution between wages and public spending. The inefficiencies reduce the marginal gain of public goods relative to wages, making *free entry* more appealing for a welfare utility-maximizing ruler. When $\kappa \longrightarrow 0$, inefficiencies are pervasive and condition (12) is never met.[1] When $\kappa \longrightarrow 1$, inefficiencies are virtually inexistent and condition (12) is more easily met. By continuity, there is a $\bar{\kappa}$ such that, for $\kappa < \bar{\kappa}$ *free entry* is always preferred, and for $\kappa \geq \bar{\kappa}$, mercantilism is always preferred. In other words, as long as the costs of enforcing entry barriers (or, alternatively, the inefficiencies of public

[1] In fact, this is already true for $\kappa \longrightarrow 1/\theta\rho$.

good provision) are contained, the mercantilist equilibrium exists.

Notice that this extension implicitly suggests which sectors should be more prone to strike a mercantilist agreement: those which are easier to tax, that is, those that have higher κ, which speaks to Gehlbach (2008).

Imperfect monopoly enforcement could be modeled in a similar fashion: when a monopoly is imperfectly enforced, the size of the monopoly market is reduced by a factor $\epsilon < 1$. A share $1 - \epsilon$ of the intermediate market is now in hands of fringe producers, which are assumed to operate the same old technology ϕ_l (otherwise Schumpeterian competition would drive one or the others out of business). With imperfect monopoly enforcement, the monopolist producer's profit is

$$\pi = (1-t)\epsilon px - \epsilon x = \epsilon[(1-t)px - x] \tag{13}$$

that is, the monopolist earns only a share ϵ of the original level, but production decreases proportionally as well. Thus, the monopolist's profit is a share ϵ of the perfectly enforcement scenario's in (2). Since ϵ pre-multiplies π, it drops the maximization problem, meaning that the reaction function $x(\epsilon)^*$ is the same as in Expression 3; hence, also equilibrium prices and wages, defined in (4) and (6), respectively. What changes? The tax revenue. Since only the monopolist producer (the one with the charter) pays taxes —fringe producers do not by definition—, tax revenue becomes a share ϵ of the original one: that is, $T = \epsilon\tau px$. Replace ϵ for κ, and we are back to the Costly Monopoly Enforcement extension. This implies that, provided that monopoly enforcement imperfections are not pervasive (i.e. $\epsilon \to 0$), there is room for mercantilism, as historical evidence suggests.

III. Evaluation of an Alternative Investment Goal

Why is the investment goal τ^u and not $\hat{\tau}$, as defined in Expression 9 and *Proposition 1*, respectively? $\hat{\tau}$ is not explicitly defined. That makes results less intuitive, but they are equivalent. That is, there still exists a non-empty interval of investment costs, $\sigma \in (0, \hat{\hat{\sigma}})$, for which investment in fiscal capacity takes place. This online Appendix sketches the existence of this interval and compares it to the one defined by *Proposition 2*. Suppose the investment goal is $\hat{\tau} < \tau^u$, and the investment costs σ_j is proportional to the investment goal. Thereby, $\sigma_{\hat{\tau}} < \sigma_{\tau^u}$. From (18) we know $\partial t^*/\partial \sigma < 0$, then $t_1^*(\sigma_{\hat{\tau}}) > t_1^*(\sigma_{\tau^u})$.

Wages are a negative function of taxes. Upon investment in fiscal capacity, $w_1^*(t_1^*(\sigma_{\hat{\tau}})|I=1) > w_1^*(t_1^*(\sigma_{\tau^u})|I=1)$. Public spending G is increasing in t^*, thus $G(t_1^*(\sigma_{\hat{\tau}})|I=1) > G(t_1^*(\sigma_{\tau^u})|I=1)$. In words, when the investment goal is $\hat{\tau}$ instead of τ^u, period 1 equilibrium wage is lower but equilibrium per capita public spending is higher.

1. Given the investment goal $\hat{\tau}$, period 1 wages $w_1^*(t_1^*(\sigma_{\hat{\tau}}))$ and public spending $G(t_1^*(\sigma_{\hat{\tau}}))$, when does the ruler invest in fiscal capacity? Suppose all the conditions in *Proposition 2* are met. Then, there exists a unique SPNE such that for all $\sigma < \hat{\hat{\sigma}}$ and $\hat{\hat{\sigma}} \in (0,1)$ investment is preferred. The proof is similar to that of *Proposition 2*.

2. Provided $\hat{\hat{\sigma}}$ exists, how does it compare to $\hat{\sigma}$, as defined in *Proposition 2*? Answer: $\hat{\hat{\sigma}} < \hat{\sigma}$

   Proof. Let $w_j^s(I)$ and $G_j^s(I)$ be the indirect utility of wages and per capita public spending following investing in fiscal capacity, $I \in \{0,1\}$, with goal $j \in \{l,h\}$, where l denotes lower investment goal $\hat{\tau}$, and h the higher investment goal τ^u, and period $s \in \{1,2\}$. Investment takes place whenever

$$w_j^1(\sigma_j^1|I=1) + G_j^1(\sigma_j^1|I=1) + w_j^2(t_j^2|I=1) + G_j^2(t_j^2|I=1) \geq 2\Big[w(I=0) + G(I=0)\Big] \tag{14}$$

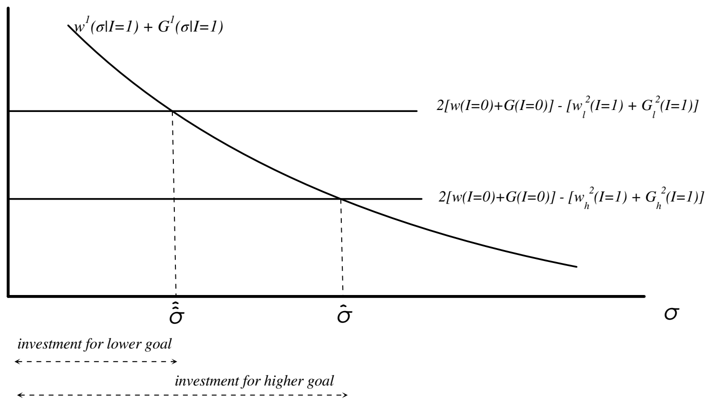
   where $t_j^2 = \hat{\tau}$ for the lower goal and τ^u for the higher. From *Proposition 1*, we know that $w(t) + G(t)$ is increasing in $t \in (0, \tau^u)$. Thus, upon investment in fiscal capacity, $w_l^2 + G_l^2 < w_h^2 + G_h^2$. We can now rearrange (14) as

$$w_j^1(\sigma_j^1|I=1) + G_j^1(\sigma_j^1|I=1) \geq 2\Big[w(I=0) + G(I=0)\Big] - \Big[w_j^2(t_j^2|I=1) + G_j^2(t_j^2|I=1)\Big] \tag{15}$$

   The left-hand side of (15) is a decreasing monotone function of σ. Since $w_l^2 + G_l^2 < w_h^2 + G_h^2$, it must be the case that the right-hand side of (15) cuts the left-hand side at a higher value of σ whenever the ruler pursues the highest goal. That is, $\hat{\hat{\sigma}} < \hat{\sigma}$. Figure Appendix-1 offers an illustration of the Proof. ■

This result implies that the parameter space of positive investment for the lower goal, $\hat{\tau}$, is smaller than the one for the higher goal, τ^u. The reason lies in the marginal gain of period 1 investment.

Figure Appendix-1: Intervals of Investment Cost of Fiscal Capacity for which Investment actually takes place for the lower and higher investment goals, $\hat{\tau}$ and τ^u, respectively.



Since the latter is relatively smaller for the lower goal, the incentives to invest also weaken. Altogether, focusing on the higher investment goal τ^u sets a more conservative scenario as it expands the parameter space of fiscal capacity investment.

IV. Mercantilism and Oligopolies

The mercantilism model in the main text assumes a monopoly market in the intermediate sector. However, the historical evidence suggests that mercantilism might be implemented in an oligopoly market (e.g. Nye (2007)). Next I model this possibility. For ease of exposition I assume a duopoly scenario, the simplest oligopoly. The results do hold for more populated versions. However, there is an obvious limit: the market has to be somehow uncompetitive so that firms gain positive profit that can be taxed in return for protection. Likewise, the more competitive the market is, the higher the transaction costs of collecting taxes. From online Appendix II, we know that high transaction costs makes mercantilism less appealing for the ruler.

This extension is more intuitively executed if the technology gap between firms is set at the marginal cost of production ϕ_j instead of quality A_j of the intermediate good. The different marginal costs associated with old and new technologies naturally reflect onto the equilibrium prices, which also capture the change in the market structure upon entry of a superior firm: namely, Schumpeterian competition transforms the oligopoly market into a monopoly market (potentially raising prices). Importantly, the isomorphism between the sources of heterogeneity across firms (marginal costs or quality of the intermediate good) is discussed in fn. 14 in the main text.[2]

Suppose that both incumbent producers, the duopolists, operate an old technology with high marginal cost, ϕ_h. The would-be entrant producers operates a new technology with low marginal costs, ϕ_l. The timing is the same. First, the incumbent firms set optimal production, and then the ruler decides whether to raise barriers or open the economy. The game is solved by backwards induction. Since the relaxation of the monopolist assumption only affects the protectionist regime, we only have to evaluate optimal production upon barriers being raised.

Suppose barriers are up. Total production of intermediate product in the duopoly x^d is the sum of individual production x_1 and x_2. The price of intermediate duopoly p^d is still determined by the productivity of the intermediate product

$$p^d = L^{1-\alpha}(x^d)^{\alpha-1} \tag{16}$$

with total duopoly production $x^d = x_1 + x_2$. For marginal cost of production ϕ_h, Firm 1 problem becomes

$$\max_{x_1} \pi_1 = (1 - t^d)L^{1-\alpha}x_1\Big[(x_1 + x_2^*)^{\alpha-1} - \phi_h\Big] \tag{17}$$

where x_2^* denotes the anticipated equilibrium production of Firm 2, and $t^d \in [0, 1]$ the tax rate imposed on the duopolists. Firm 1 problem is solved for x_1 as implicitly defined by

$$(1 - t^d)L^{1-\alpha}(x_1 + x_2^*)^{-2+\alpha}(\alpha x_1 + x_2^*) = \phi_h \tag{18}$$

Expression 18 implies x_1^* is a negative function of x_2^*, ranging from $x_1^* = 0$ to $x_1^* = L(\alpha(1 - t^d)/\phi_h)^{1/(1-\alpha)}$, the monopolist production, xm, given by (3) in the main text.

Since both firms face similar production costs, the reaction function of Firm 2 is symmetrical. Thus, x_2^* is implicitly defined by

$$(1 - t^d)L^{1-\alpha}(x_2 + x_1^*)^{-2+\alpha}(\alpha x_2 + x_1^*) = \phi_h \tag{19}$$

[2] Recall that in the original set up, the intermediate good price is independent of the quality of the product. Still, the final market producers prefer the more productive intermediate good, as final production is increasing in quality, $Y(A_j)$. That assumption is enough to model Schumpeterian competition when we work with monopolies, and it simplifies algebra too. But when we work with oligopolies, we need prices to reflect the market structure, as they change in case of entry: from oligopoly to monopoly pricing.

By symmetry, (19) defines x_2^* as a negative function of x_1^*. Since both firms are analogous, by the Cournot Theorem we know that $0 < x_1^* = x_2^* < x^m$, with total duopolistic production $x^d = x_1^* + x_2^* > x^m$.

Given $x_1^* = x_2^*$ we can express the FOC in (18) as

$$(1-t^d)L^{1-\alpha}(2x_1)^{-2+\alpha}(x_1(1+\alpha)) = \phi_h \tag{20}$$

and solve for x_1:

$$x_1^* = \left[\frac{(1-t^d)(1+\alpha)L^{1-\alpha}2^{\alpha-2}}{\phi_h}\right]^{\frac{1}{1-\alpha}} \tag{21}$$

Since $x_1^* = x_2^*$, total duopolist production

$$(x^d)^* = L\left[\frac{(1-t^d)(1+\alpha)}{2\phi_h}\right]^{\frac{1}{1-\alpha}} \tag{22}$$

Given $(x^d)^*$ and inverse demand $p(x^d)^*$, the welfare utility maximizing ruler optimizes the tax rate paid by each duopolist if barriers are raised in exchange for higher tax compliance

$$(t_m^d)^* = \frac{(1-\alpha)\Big[\theta(2\rho-1)-1\Big]}{\theta(2\rho+1-\alpha)-(1-\alpha)} \tag{23}$$

where subscript m denotes the trade regime, *mercantilism* or free *entry*, and the superscript denotes the market structure, *duopoly* vs *monopoly.* Notice that (23) is always constrained between 0 and 1, and is increasing botwh in θ and ρ. This tax rate is smaller than $(t_m^m)^*$, the tax rate when protection in adopted in a monopolist market and defined in (9).[3] Duopolists make smaller profit than the monopolist and, as a direct consequence, they cannot be taxed as much as the latter.

Upon entry, the market becomes monopolistic. Thus, prices might increase relative to the duopolist scenario, making protection unnecessary. This is not the case if the would-be entrant is competitive enough. Specifically,

$$\frac{\phi_h}{\phi_l} > \frac{1+\alpha}{2\alpha}\frac{1-t_m^d}{1-t_m^m} \tag{24}$$

guarantees that, upon entry, the price offered by the new firm beats that of the incumbent produc-

[3] This can be proved with a little algebra.

ers.[4] When this condition is satisfied, the duopolists have an interested in protection even if that implies higher taxes (i.e. they accept the conditions of mercantilism).

Given $x_m^d(t_m^d)^*$, a welfare utility maximizing ruler decides whether to raise barriers and enforce $(t_m^d)^*$ as defined by (23) or allow *free entry*, with $(t_e^m)^* = \tau$ and payoffs as defined by *Proposition 1.*

Proposition 1. *Suppose the fiscal capacity constraint in (8) binds. Then*

- *If*
$$\frac{\phi_h}{\phi_l} < \frac{(1+\alpha)\big[\theta(\rho+1-\alpha)-(1-\alpha)\big]}{\alpha\big[\theta(2\rho+1-\alpha)-(1-\alpha)\big]} \tag{25}$$
then, protection is preferred to free entry for all $\tau \in [0, t_{\lambda=0}]$

- *If*
$$\frac{\phi_h}{\phi_l} > \frac{(1+\alpha)\theta(1+\alpha(\rho-1))}{\alpha\big[\theta(2\rho-1-\alpha)-(1-\alpha)\big]}\left[\frac{\theta(\alpha(\rho-1)+1)}{\theta(1-\alpha)+1}\right]^{\frac{1-\alpha}{\alpha}} \tag{26}$$
then, free entry is preferred to protection for all $\tau \in [0, t_{\lambda=0}]$

- *If*
$$\frac{\theta(\rho+1-\alpha)-(1-\alpha)}{\theta(1+\alpha(\rho-1))} \leq \frac{\phi_h}{\phi_l} \leq \left[\frac{\theta(\alpha(\rho-1)+1)}{\theta(1-\alpha)+1}\right]^{\frac{1-\alpha}{\alpha}} \tag{27}$$
then, there exists a $\hat{\tau}_d < t_{\lambda=0}$ *such that, for all* $\tau \leq \hat{\tau}_d$*, a unique SPNE exists in which the ruler adopts entry barriers and the duopolist pay* $(t_m^d)^* > \tau$*, as defined in (23); and for all* $\tau > \hat{\tau}_d$*, free entry is allowed, entry takes place, and the tax rate is set to exhaust the stock of fiscal capacity* $(t_e^m)^* = \tau$*.*

First, *Proposition 4* states that when the technological distance between the duopolist and the new entrant is very low, the gains of entry (better technology) do not compensate its costs (monopolist prices increase relative to duopoly). Accordingly, the *status quo* (i.e., protection) is preferred. Intuitively, in an oligopolistic scenario the ruler is more demanding with the new entrant's technology than she is in the original monopoly set up. Second, *Proposition 4* states that whenever the technological distance between the duopolist and the new entrant is very large, the gains of entry cannot be compensated by an increase in taxation by the duopolist. Accordingly,

[4]This conditions comes from comparing equilibrium prices of the duopolist vs the monopolist, given ϕ_j.

entry is preferred. Third, when the technological distance between the duopolist and the new entrant is intermediate, protection is preferred to *free entry* only if the stock of fiscal capacity is sufficiently low. Importantly, only when the latter condition is met, protection is exchanged for tax compliance. This is true because the duopolists seek protection from superior competitors (which pay back in taxes) only when (24) is met, and this condition coincides with the lower bound of (27), once we plug in $(t_m^d)^*$ and $(t_m^m)^*$. Notice that Expression 27 is virtually identical to *Proposition 1*. Ultimately, this extension suggests that the assumption of a monopolist producer in the main text is just a simplification. An oligopoly market is consistent with mercantilism.

To proof of *Proposition 4* we follow the same strategy as in *Proposition 1*. Let L be normalized to 1, then protection is preferred to free entry whenever $V_m^d((t_m^d)^*, (x^d)^*|\phi_h) > V_e^m((t_e^m)^*, (x_e^m)^*|\phi_l)$, with $(t_m^d)^*$, $(x^d)^*$, $(x_e^m)^*$ defined in (23), (22) and (3), respectively, and marginal costs $\phi_h > \phi_l$. $V_m^d((t_m^d)^*, (x^d)^*|\phi_h)$ defines a horizontal line in the $V - t$ space. From *Proposition 1*, we know that V_e is increasing in the stock of fiscal capacity τ. Moreover, we know that the tax rate is set to exhaust fiscal capacity under free entry $(t_e^m)^* = \tau$. For the existence of $\hat{\tau}_d$, both curves, V_m^d and V_e, must cut at some $\hat{\tau}_d$ between 0 and $\tau_{\lambda=0}$, the unconstrained tax rate. By continuity of $V_m^d(\cdot)$ and $V_e^m(\cdot)$, this point exists if and only if $V_e^m(\tau \to 0) < V_m^d$ and $V_e^m(\tau = \tau_{\lambda=0}) > V_m^d$.

For $V_e^m(\tau \to 0) < V_m^d((t_m^d)^*, (x^d)^*|\phi_h)$, we first plug equilibrium values and then simplify to

$$\begin{aligned} &\left[\frac{\alpha}{\phi_l}\right]^{\frac{\alpha}{1-\alpha}} [(1-\alpha)(\theta(\tfrac{1}{\alpha}-1)+1)] \\ &< \left[\frac{1+\alpha}{\phi_h}\right]^{\frac{\alpha}{1-\alpha}} \left[\frac{\theta(1+\alpha(\rho-1))}{\theta(2\rho+1-\alpha)-(1-\alpha)}\right]^{\frac{\alpha}{1-\alpha}} \left(\frac{\theta(1-\alpha)(1+\alpha(\rho-1))}{\alpha}\right) \end{aligned} \tag{28}$$

which is true when (25) is *not* met. Otherwise, protection is always preferred.

For $V_e^m(\tau \to \tau_{\lambda=0}) > V_m^d$, we plug equilibrium values and the simplify to

$$\begin{aligned} &\left[\frac{\alpha}{\phi_l}\right]^{\frac{\alpha}{1-\alpha}} \left[\frac{\theta(1+\alpha(\rho-1))}{\theta(\rho+1-\alpha)-(1-\alpha)}\right]^{\frac{\alpha}{1-\alpha}} \left(\frac{\theta(1-\alpha)(1+(\rho-1))}{\alpha}\right) \\ &> \left[\frac{1+\alpha}{\phi_h}\right]^{\frac{\alpha}{1-\alpha}} \left[\frac{\theta(1+\alpha(\rho-1))}{\theta(2\rho+1-\alpha)-(1-\alpha)}\right]^{\frac{\alpha}{1-\alpha}} \left(\frac{\theta(1-\alpha)(1+\alpha(\rho-1))}{\alpha}\right) \end{aligned} \tag{29}$$

which is true when (26) is *not* met. Otherwise, free entry is always preferred.

Conditions 28 and 29 are simultaneoysly met when (27) is met. Then, by he Intermediate Value Theorem, a $\hat{\tau}_d < t_{\lambda=0}$ exists such that for all $\tau < \hat{\tau}_d$, protection of the duopoly is preferred to free entry. This completes the proof of *Proposition 4*.

V. Further Data Details

Fiscal Capacity. Fiscal capacity is proxied by the share of income taxes to total taxation. The ratio is drawn from Flora, Kraus and Pfenning (1983). The income tax is adopted at different dates across Europe. When no income tax exists, the variable is set to 0. The oldest income tax records for Austria, Italy and Denmark are missing. For Austria, the income tax data starts in 1898, 33 years after the income tax was officially adopted. The record for 1898 is 3.4 (as % of total tax revenue). Given this small value, I set all records for Austria from 1865-1897 to 0. The first records for Italy and Denmark, dated 1877 and 1917, respectively, are 17.8 and 14 (as % of total tax revenue). These values are too large to assume that the income tax proceeds were 0 since the time of adoption (1864 in Italy, 1903 in Denmark). We would be ignoring much of the learning curve in income tax collection if we set these values to 0. Thus, I keep them as missing.

Interpolation. Only control variables are interpolated: GDP, Population, Military Mobilization, Urbanization and Schooling rates. This way I minimize the risk of correlation among key variables in the model being driven by artificial data completion.

Austria. Lampe and Sharp's (2013) dataset does not include AVE tariff data for Austria. I retrieve these values from Clemens and Williamson, provided that the country-correlation between the two series for the remaining ten countries is at least .93. I do not use Clemens and Williamson because Belgium, Netherlands and Switzerland are not covered, and data gaps for the remaining countries are bigger.

VI. Robustness Tests

In this section I retest hypothesis 1 by not setting *tax ratios* to 0 for all the years separating 1820 from the adoption date of the income tax. Instead, I leave them as missing. For instance: Norway adopted the income tax in 1892. In the original test, the tax ratio equals 0 between 1820 and 1891. Here, the tax ratio is set to missing. Then, I compute the first difference of tax ratio (the measure of fiscal capacity growth), allowing for positive and negative changes.

Table Appendix-1 suggests that results in Table 2 are not driven by a coding decision. Regardless

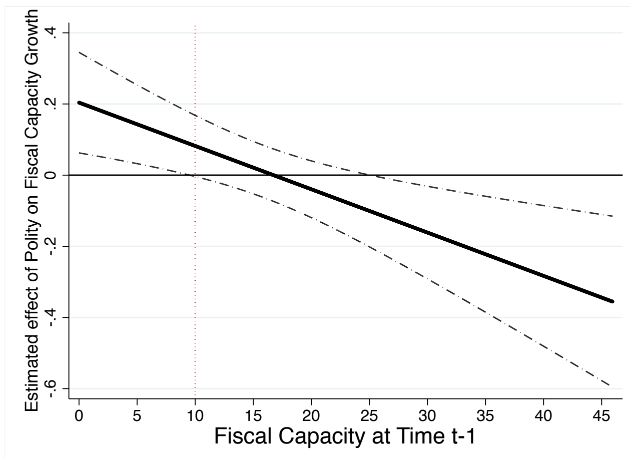
Table Appendix-1: Fiscal Capacity Growth (positive and negative changes) as a function of past realizations of the stock of fiscal capacity and the ruler-labor policy preference alignment (proxied by Polity IV). In this test, the stock of the tax ratio (and thus the dependent variable) is set to missing while the income tax has not yet been adopted.

	Two-way FE		Flex Polynomial	
	(1)	(2)	(3)	(4)
Lagged Fiscal Capacity	-0.146**	-0.130**	-0.116***	-0.088*
	(0.058)	(0.057)	(0.044)	(0.045)
Polity	0.129	0.220**	0.204**	0.247**
	(0.093)	(0.098)	(0.086)	(0.098)
Polity × Lagged Fiscal Capacity	-0.012**	-0.018***	-0.012***	-0.019***
	(0.005)	(0.005)	(0.005)	(0.005)
GDP/cap	0.257	-0.405	0.038	-0.325
	(0.520)	(0.663)	(0.461)	(0.562)
War	-2.332*	-2.168	0.594	-0.689
	(1.338)	(1.399)	(0.827)	(0.832)
AVE tariffs		6.196		-7.633
		(8.730)		(5.392)
Urbanization		-17.005**		-12.002**
		(8.379)		(6.014)
Military size		-0.001		0.006
		(0.006)		(0.004)
Primary Education		-7.133***		-5.584**
		(2.709)		(2.441)
Constant	5.779*	24.967***	-0.991	-44.275**
	(3.196)	(9.081)	(4.717)	(18.680)
Observations	468	443	468	443
R-squared	0.363	0.376	0.140	0.165
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	No	No
Flexibile Polynomial	No	No	Yes	Yes
WW Participant Indicator	Yes	Yes	Yes	Yes

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

of how I code tax ratios for the time span separating income tax adoption from 1820, results hold: when the stock of fiscal capacity is low, fiscal capacity expands provided that ruler and labor preferences (proxied by Polity IV) are aligned. We can conclude this based on Figure Appendix-2, which plots how the Polity score affects Fiscal Capacity Growth as a function of the stock of fiscal capacity. The results are even more favorable to the working hypothesis, as the interval of the past realization of the fiscal capacity for which the marginal effect of Polity is positive and statistically different from 0 is larger.

Figure Appendix-2: Marginal effect of Ruler-Labor Policy Preference Alignment (proxied by Polity IV) on Fiscal Capacity Growth as a function of the stock of fiscal capacity at time $t-1$. 90% CI. The stock is set to missing while the income tax has not yet been adopted.



References

Flora, Peter, Franz Kraus and Winfried Pfenning. 1983. *State, Economy, and Society in Western Europe 1815-1975*. London: Macmillan Press.

Gehlbach, Scott. 2008. *Representation Through Taxation: Revenue, Politics, and Development in Postcommunist States*. Cambridge: Cambridge University Press.

Lampe, Markus and Paul Sharp. 2013. "Tariffs and income: A time series analysis for 24 countries." *Cliometrica* 7(3):207–235.

Nye, John VC. 2007. *War, Wine, and Taxes: The Political Economy of Anglo-French Trade, 1689-1900*. New Jersey: Princeton University Press.