NBER WORKING PAPER SERIES

TECHNOLOGICAL RIVALRY AND OPTIMAL DYNAMIC POLICY IN AN OPEN ECONOMY

Yan Bai
Keyu Jin
Dan Lu

Working Paper 31703
http://www.nber.org/papers/w31703

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2023, Revised July 2024

We acknowledge the support from the Research Grants Council of Hong Kong under Grant GRF-14505721. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2023 by Yan Bai, Keyu Jin, and Dan Lu. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Technological Rivalry and Optimal Dynamic Policy in an Open Economy
Yan Bai, Keyu Jin, and Dan Lu
NBER Working Paper No. 31703
September 2023, Revised July 2024
JEL No. E23,F12,F63,O38

ABSTRACT

What are a country's policy options concerning the development of emerging technologies in a global economy? To address this question, we investigate optimal dynamic policies in an open economy where technology is endogenously accumulated through R&D innovation. Our key insight is that a country has incentives to influence foreign innovation efforts across sectors and over time—giving rise to optimal policies even when private innovation allocations are (Pareto) efficient. We derive explicit expressions for optimal taxes linked to both an intratemporal and an intertemporal motive to manipulate foreign technology. To affect foreign forward-looking innovation decisions, Home uses committed future trade policies (Ramsey) or innovation policies (Markov). For intratemporal price manipulation, Home government imposes higher tariffs in sectors where it has a comparative advantage.

Yan Bai
Department of Economics
University of Rochester
216 Harkness Hall
Rochester, NY 14627
and NBER
yanbai06@gmail.com

Keyu Jin
London School of Economics
k.jin@lse.ac.uk

Dan Lu
Department of Economics
9/Fl., Esther Lee Building
The Chinese University of Hong Kong
danlyu@cuhk.edu.hk

1 Introduction

Trade disputes are often not about trade, but about technological rivalry. Throughout history, we have seen how development in leading-edge sectors has spurred national competition, protectionism, sometimes under the guise of trade wars. For example, when the U.S. and Japan developed their semiconductor industry in the 1980s, both countries imposed significant subsidies in the domestic semiconductor sectors. In another example, one of the core disputes in the 2018 U.S.-China trade war was over technology. Along with increasing U.S. restrictions and controls, domestic policies inside China to boost demand and procurement and incentivize research in these goods have soared. At the same time, in the U.S., the passing of the America Competes Act of 2022 earmarks 52 billion dollars of funding for the semiconductors industry. In a world where there is rising demand for technologies such as semiconductors, 5G, a race to gain an edge in emerging technologies such as renewables and AI, a country may have a desire to react.

This paper provides a framework to explore the nexus between trade policies and technological development in the context of globalization and competition. The spirit of the contest is captivatingly explored in Krugman (1994), which argues that developing countries' technological advancement can sometimes harm the interests of advanced economies, by becoming more productive in sectors in which rich countries have a comparative advantage (or conversely, benefit them if rapid productivity growth happens in developing countries' comparative advantage sectors). This simple but powerful argument is based on comparative static analysis over exogenous technology, and does not consider the option that one country may attempt to influence the innovation efforts of another, nor does it consider the costs of innovation. To re-evaluate Krugman's thesis when technology endogenously evolves, and when countries have a set of instruments to influence trade and technology dynamics, we theoretically characterize optimal dynamic taxation in a workhorse Ricardian model with endogenous technology.

We show that a government has an incentive to manipulate innovation efforts in *other* countries across sectors and over time. Our main takeaway and novel finding is that even when markets are efficient and the government has access to a full set of trade policies, the optimal policy involves domestic innovation policies when a government *cannot* commit to

its future policies. This motivation goes beyond conventional industrial policies that justify subsidizing or protecting some sectors at the expense of others on the basis of externalities and spillovers.

We extend the one-sector framework of Eaton and Kortum (2001) to a multi-country, multi-sector model. This model incorporates comparative advantage-based trade and endogenous technology accumulation through R&D and innovation. We study optimal innovation and trade policy in this framework.

The economy's government is benevolent and can choose a set of sector-specific domestic taxes/subsidies on R&D, as well as differential trade policies across sectors and trading partners. When choosing these policies, the government internalizes its choices on trade and technology development in its own country as well as in others. Other economies' government is taken to be passive in the benchmark scenario. We first consider Ramsey optimal policy, where Home can commit to a sequence of taxes, and then compare it to the Markov optimal policy.

We opt for this rich, dynamic model as it meets three criteria. First, it nests an efficient baseline Ricardian economy. In the baseline, private innovation is efficient; hence, it obviates the need for other classic industrial-policy interventions in a closed economy. This helps isolate the new mechanism underpinning optimal policy. Second, the innovation process is such that both sectoral and aggregate trade patterns–and hence optimal policy–can be explicitly characterized. This allows us to compare the results with the standard workhorse Eaton and Kortum (2002) Ricardian model with exogenous technology. Moreover, individual researcher's expected profits are explicitly characterized, as a result, we can formulate the dynamic optimal policies in a more explicit manner. Third, the enriched version of the model can accommodate many additional features of interest, for example, economies of scale, knowledge spillovers, or externalities. The complexity of the problem lies in the fact that these additional elements already give rise to policy interventions in a closed-economy environment, but added to the fact is that they are intertwined with trade policies. Thus, the sharp predictions stemming from our analytically convenient results are particularly appealing given the richness of these mechanisms.

The baseline economy is efficient. Productivity (innovation efficiency) differences across

sectors shape comparative advantage, which determines trade across countries. There are constant returns to scale in production and in innovation, and a free flow of labor between research and production. The model features Bertrand competition between producers for each goods, where each producer competes with all other producers in the world. Although the realized markup of each firm follows a distribution, the aggregate endogenous distribution of markups is Pareto, and the expected profit of a firm is a constant share of sales in the industry. Thus, there are constant and identical aggregate markups in each sector. In addition, there are no other taxes, distortions, or externalities in the baseline case that underlie conventional industrial policy motives.

In our baseline closed-economy setting, the planner would choose the same allocation as the market. Openness itself does not affect the level of private innovation intensity: the increased foreign competition that spurs innovation is exactly offset by the larger foreign market effect that tends to reduce innovation effort. However, in the open economy environment, there is scope for policy corresponding to both an intertemporal and intratemporal motive to manipulate technology. Home's planner understands that it can undertake actions to influence foreign profits and research returns, and hence curb foreign innovation efforts in some sectors or channel them towards those that would benefit Home.

To bring clarity to the disparate set of mechanisms, we build up our results by first zeroing in on a dynamic technology manipulation motive in a one-sector model that is efficient— absent any externalities or intratemporal relative prices. Consider a case where Home is innovating and transitioning to a higher level of technology. We show that Home would want to suppress foreign innovation efforts by curtailing its research returns while raising its cost of innovation. If Home could commit to a schedule of trade policies (a path of export taxes), it would suffice to implement an optimal foreign allocation and relative prices over time. Under these Ramsey policies, Home's export tax rises but becomes flatter over time. This generates a higher Foreign price index today and a relatively lower future price, reducing expected profits while pushing up the interest rate. As a result, Foreign reduces innovation investment and produces more, which in turn makes innovation cheaper for Home.

Our first key result is that in the absence of commitment mechanisms, trade policies

would not be enough. Home would try to influence Foreign's innovation incentives by altering its own innovation efforts. To show this result, we consider Markov policies. In this case, the home government chooses optimal policies every period, taking the world state and future government's optimal policy rules as given. In contrast to Ramsey policies, Markov policies would invoke domestic innovation policies even in the absence of any domestic externalities and even though they distort domestic R&D efforts. The reason is that without commitment, the government always has an incentive to deviate from the promised trade policy in the following period: it imposes a higher export tax to increase Foreign prices today, but cannot commit to lowering it tomorrow. Hence, trade policies can only be used to manipulate contemporaneous prices. To influence future prices, Home changes its own level of technology through innovation subsidy so as to alter the path of foreign interest rates and wages (innovation costs). This, in turn, induces Foreign to innovate less and produce more.

The second key result is that when comparative advantage is introduced, there is a rich set of interweaving intratemporal mechanisms to alter relative prices and innovation. To uncover these different forces, we turn to a multi-sector model in the long run and derive an explicit formula that links a country's optimal tariff in a given sector to the country's comparative advantage: Home would generally want to improve its terms of trade by inducing Foreign to do more research and enhance technology in Foreign's comparative advantage sector (as opposed to Home's). To do so, it imposes a higher tariff in sectors that Home has larger net exports (relative to the foreign production) and raises that tariff when net exports increase in that sector–for example, due to openness or a rise in global demand. By reducing the demand for foreign goods in those sectors, these tariffs can curb foreigners' research efforts in those sectors.

The heterogenous import tariff schedule looks very different from the uniform optimal tariff under the fixed technology case, as analyzed in Costinot, Donaldson, Vogel, and Werning (2015). They show that a country can exercise monopoly power to tilt relative prices in its favor and improve welfare. A country would opt for a higher export tax in sectors with greater comparative advantage (or a higher subsidy in the comparative disadvantage sectors), combined with a uniform (or zero) import tariff. The reason is that there

is more room to manipulate world prices in the comparative advantage sector—achieved through a heterogenous export tax schedule. There is no need to levy differential import tariffs, as a uniform tax is sufficient to bring about a reduction in foreign wages, obviating the need to distort domestic consumption prices further.[1] While this force is also present here, the novel mechanisms in our model give rise to starkly different results: heterogeneous-sector level import tariffs serve the purpose of reducing demand for corresponding foreign goods and hence Foreign's incentives to innovate in these sectors.

Third, our framework provides additional insights when there are more than two countries. Take the example of the US and China competition in a sector such as semiconductors. In the two-country case, our model predicts that the U.S. would want to discourage innovation in China's semiconductor industry as it is the U.S.' comparative advantage. Suppose there is a rise in the global demand for semiconductors, then unilaterally the U.S. would consequently raise its tariffs on Chinese semiconductors, whereas China would lower its tariff on the U.S. semiconductors. However, the two-country case precludes the realistic scenario in which the U.S. and China are both net exporters of semiconductor products to the rest of the world. In the multi-region scenario, both China and the U.S. would unilaterally raise their tariffs levied on each other in the semiconductor industry. But both would also impose a higher tariff on the rest of the world's other sector–textiles, for instance— to induce the other competitor to shift innovation efforts into textiles and away from semiconductors. We explore the long-run optimal policies under a number of scenarios, including a greater demand for certain goods, the developing nation's catching up, a fall in trade costs, etc., in a multi-region long-run model.

The main mechanism that optimal trade and industrial policies are employed to influence a foreign nation's innovation efforts is robust to a range of extensions, for example, allowing for varying returns to scale in innovation. In Section 5, we explore a case with

[1] Bagwell and Staiger (1999) emphasize terms-of-trade manipulation and its implication for the WTO. Trade policy in a partial equilibrium setting is explored in Gros (1987), and Broda, Limao, and Weinstein (2008), which show that industry tariff is related to the foreign export supply elasticity. Demidova and Rodríguez-Clare (2009) characterize optimal tariffs in a small economy, single industry, Melitz-Pareto setting. Ossa (2014) computes optimal tariff across sectors considering traditional, new trade, and political economy motives for protection. Trade policy analyzed in quantitative or new trade theories include Caliendo and Parro (2021); Costinot and Rodríguez-Clare (2014); Costinot et al. (2015); Demidova (2017); Lashkaripour and Lugovskyy (2023). Costinot, Rodríguez-Clare, and Werning (2020) characterizes optimal firm-level trade policy in a two-country-single-sector Melitz model.

decreasing return to scale in innovation. The gap between the private and social return for R&D justifies industrial policies that subsidize or tax innovation when there are externalities. There are also nontrivial interactions between industrial policies and heterogeneous trade policies. For example, if all sectors and countries have the same decreasing returns to scale in innovation, and both Home and Foreign use innovation tax to correct these externalities at the steady state, Markov policy will still employ additional innovation policies. Ramsey policy can simply use a constant innovation tax to target the externality and use a path of trade policies. But without commitment, Home resorts to additional innovation policies to change future marginal innovation costs across sectors in Foreign. This distinguishes dynamic optimal policies from static ones, which use domestic policies to correct domestic externalities and trade policies to manipulate terms of trade. For dynamic optimal policies, even at the steady state, there are future periods, and whether Home can commit to future policies matters.

Our Markov results contrast with optimal policy in Bartelme, Costinot, Donaldson, and Rodriguez-Clare (2019) and Lashkaripour and Lugovskyy (2023),[2] whereby industrial policies are used solely to correct domestic wedges/inefficiency. In both papers, the model is static. By contrast, we focus on optimal dynamic policies with individual investment decisions that can influence technology. Under our general setup, optimal policy depends on the dynamic features, in addition to the gap between the private and social choice of R&D generated in the domestic market.

Our contribution is primarily theoretical, as a first attempt to understand the basic contours of optimal policy when there is technological evolution and competition. As such, our technical contribution is to theoretically characterize optimal dynamic policy and derive general results in a framework with elemental features, while providing explicit formulas that give rise to sharp predictions about the structure of optimal policy in special cases. The emphasis on dynamic terms of trade manipulation is closely related to Costinot, Lorenzoni, and Werning (2014), which proposes a theory of capital controls as dynamic terms-of-trade manipulation in an endowment economy. By contrast, in our dynamic economy with en-

[2] Bartelme et al. (2019) characterize optimal policy for a small open economy in a multi-sector Ricardian model with Marshallian externalities. Lashkaripour and Lugovskyy (2023) study optimal industry and trade policy with scale economies.

dogenous technology, industrial policies serve as intertemporal trade policies. In general, our approach differs from the numerical approaches to computing optimal policy in a particular environment at a moment in time.

Our paper sidesteps from issues like international technology diffusion or policy competition. First, there is already a large and expansive literature on the topic of international technology diffusion in the global economy, but few consider optimal policy in these settings.[3] One exception is Bai, Jin, Lu, and Wang (2024), who studies optimal trade policies in a framework as Buera and Oberfield (2020) with international diffusions through imports. Our paper provides a general setup and solution method that can nest cross-sector and cross-country innovation diffusion. Second, we emphasize a country's incentives on trade and domestic policies and thus assume foreign countries are passive in their policies.

The paper is related to, but has little overlap with, the growth literature emphasizing the importance of R&D on long-run growth. Optimal policies in these contexts depend on assumptions of each theory—featuring either imperfect competition pricing, knowledge spillovers, congestion externalities, or creative destruction. Akcigit, Ates, and Impullitti (2019) explore policies with these features in an open economy with a one-sector model that does not have comparative advantage aspects to trade (and hence heterogeneous tariffs). Liu and Ma (2021) examine optimal R&D policy for a small open economy when there are cross-sector spillovers and externalities without dynamic considerations. In both of these papers, it is the presence of externalities, spillovers, or distortions that justify interventionist policies, whereas optimal dynamic policies in our setting arise from both static and dynamic terms of trade consideration and comparative advantage for an open economy. Furthermore, our model features a dynamic economy with endogenous technology, industrial policies serve as intertemporal trade policies.

Our paper proceeds as follows. Section 2 sets up the multi-country, multi-sector dynamic theoretical framework, while Section 3 zooms in on the dynamic technology manipulation motive in a one-sector model. Section 4 focuses on the intratemporal motive to manipulate

[3]Innovation or international technology diffusion in the global economy include works such as Arkolakis, Ramondo, Rodríguez-Clare, and Yeaple (2018), Atkeson and Burstein (2010), Bloom, Draca, and Van Reenen (2016), Buera and Oberfield (2020), Cai, Li, and Santacreu (2022), Eaton and Kortum (1999), Grossman and Helpman (1990), Grossman and Helpman (1993), Hsieh, Klenow, and Nath (2019), Perla, Tonetti, and Waugh (2021), and Somale (2021).

technology, and Section 5 concludes with a comparison of various policies in the full-fledged dynamic model with externalities.

2 Theoretical Framework

2.1 Model

The baseline model extends the endogenous technology model in Eaton and Kortum (2001), henceforward EK2001, to one that features multiple sectors and countries and derives unilateral optimal R&D and trade policies therein.

The world has N countries and J sectors. We assume country 1 is the Home country and study its unilateral optimal innovation and trade policies. We adopt a semi-primal approach where the Home government chooses R&D investment directly and implements this optimal choice with an innovation policy. The form of trade policies is specified explicitly, i.e., country-sector-specific export taxes $\tau^x_{i,j,t}$ and tariffs $\tau^m_{i,j,t}$ on the goods in sector $j \in [1, J]$ of country $i \in [2, N]$. In this section, we characterize the market equilibrium given Home's trade policies.

All consumers' discount factor is ρ. Country n has a measure L_n of labor, which can freely flow into the production sector as a worker or the research sector as an innovator. Consumer preference in each country n is

$$\sum_{t=0}^{\infty} \rho^t u\left(C_{nt}\right), \tag{1}$$

where the per period utility $u(\cdot)$ satisfies the standard Inada condition. Final goods is a Cobb-Douglas function across the consumption of different sector $j \in J$ goods $C_n = \Pi_{j\in J}\left(C_{n,j}\right)^{\beta_j}$, where β_j is constant and reflects the share of sector j. Within a sector, there is a continuum of varieties of consumption goods, which are aggregated with a Cobb-Douglas function $C_{n,j} = \exp \int_0^1 \ln c_{nj}(\omega) d\omega$. All goods are tradable with an iceberg trade cost d_{ni} between country n and i. Within each country, there are perfect financial markets. Across countries, consumers can only trade goods, and trade is balanced every period.

Innovation incentive and research decision. We start by explaining innovation efforts

within each sector, as in the one-sector economy model of Eaton and Kortum (2001). All countries n are capable of producing any variety ω of good with technology $q_n(\omega)$, the distribution of which is endogenous and depends on the number of researchers and research productivity.

Researchers in country n and sector j draw ideas about how to produce goods in this sector. Each idea consists of the realization of two random variables. One is the good ω to which the idea applies, drawn from the uniform distribution over $[0, 1]$. The other is the efficiency $q(\omega)$, drawn from a Pareto distribution with a parameter θ. The number of ideas per researcher is Poisson distributed with parameter $\alpha_{n,j}$, which reflects how effective the researchers are in the country n sector j's innovation process—-or, innovation efficiency.[4]

Let the measure of researchers in the country n sector j at period t be $L^r_{n,j,t}$, and the cumulative stock of ideas be $T_{n,j,t}$. Under a unit interval of varieties, the number of ideas for producing a specific good is Poisson distributed with parameter $T_{n,j,t}$. Ideas retire with probability δ and hence the evolution of the stock of ideas $T_{n,j,t}$ is[5]

$$T_{n,j,t} = (1-\delta)T_{n,j,t-1} + \alpha_{n,j}L^r_{n,j,t}. \tag{2}$$

Kortum (1997) proves that when the quality of idea is Pareto distributed, the distribution of technology efficiency frontier is a Frechet distribution with parameter $T_{n,j,t}$ and θ.

Firms engage in Bertrand competition: the lowest-cost producer of each good in each market claims the entire market for that good, charging a markup just enough to keep the second-lowest-cost producer out of the market. In equilibrium, the distribution of the markup is Pareto with the parameter θ. Since all firms selling in the market charge a markup drawn from the same distribution, total profits $\zeta_{n,j,t}$ of firms at period t earned by firms in the market are a constant share of total sales in that market.

[4]In our baseline, the average number of ideas $\{\alpha_{n,j}\}$ in each sector and country are constant. One can let them endogenously depend on decreasing return to scale in innovation or domestic diffusion, for example, $\alpha_{n,j,t} = \bar{\alpha}_{n,j}(L^r_{n,j,t})^{\varepsilon-1}(T_{n,j,t-1})^{\eta}$ with a decreasing return to scale parameter $\varepsilon < 1$ and intertemporal diffusion η. Alternatively, the number of ideas can depend on foreign diffusion $\alpha_{n,j,t} = \bar{\alpha}_{n,j}(L^r_{n,j,t})^{\varepsilon-1}(T_{m,j,t-1})^{\eta}$ with $m \neq n$. Furthermore, it can be used to consider policies with input-output innovation structure across sectors, indicated by j and k, where $\alpha_{n,j,t} = \bar{\alpha}_{n,j}\prod_k (T_{n,k,t-1})^{\omega_{jk}}$.

[5]Note that our model is isomorphic to the detrended semi-endogenous growth model in Eaton and Kortum (2001), which assumes the growth rate of α_n is exogenously driven by population growth g_L. Since policies would not affect the exogenous growth rate, we use the detrended model.

Specifically, the profit $\zeta_{n,j,t}$ of a firm in sector j of country n has two key pieces. First, it depends on the total expenditure of each country i on sector j's goods. Let x_{it} denote country i's total spending at time t. Then, $\beta_j x_{it}$ is its expenditure on sector j, and it is also the expenditure per variety in sector j given Cobb-Douglas preferences. Thus, profits earned by either domestic or foreign firms who sell in that market is $\beta_j x_{it}/(1+\theta)$. Second, the profit of a firm in country n depends on whether it can excel other firms in each market. The probability that a researcher in n and j draws a q that is the lowest price in market i at t is $\pi_{in,j,t}/T_{n,j,t}$: a firm innovates and surpasses the current set of ideas in its own country with probability $1/T_{n,j,t}$ at time t, but then it needs to be the cheapest source of a particular good in country i, with probability $\pi_{in,j,t}$.

Hence, the expected profit earned by a firm in country $n \neq 1$ is given by

$$\zeta_{n,j,t} = \frac{1}{T_{n,j,t}} \frac{1}{1+\theta} \left[\frac{1}{1+\tau^m_{n,j,t}} \pi_{1n,j,t} \beta_j x_{1t} + \sum_{i \neq 1} \pi_{in,j,t} \beta_j x_{it} \right] = \frac{1}{\theta} \frac{w_{nt} L^p_{n,j,t}}{T_{n,j,t}}, \tag{3}$$

where the expected profit takes into account Country 1's import tariffs $\tau^m_{n,j,t}$ on Country n's goods. The second equality holds because in equilibrium $1/(1+\theta)$ fraction of total sales goes to profit, and $\theta/(1+\theta)$ share goes to production labors. We can, therefore, replace the total sales of country n inside the bracket with $\frac{1+\theta}{\theta} w_{nt} L^p_{n,j,t}$ and obtain the second equality. Similarly, the expected profit earned by a firm in country 1 (Home country) at time t is:

$$\zeta_{1,j,t} = \frac{1}{T_{1,j,t}} \frac{1}{1+\theta} \left[\pi_{11,j,t} \beta_j x_{1t} + \sum_{i \neq 1} \frac{1}{1+\tau^x_{i,j,t}} \pi_{i1,j,t} \beta_j x_{it} \right] = \frac{1}{\theta} \frac{w_{1t} L^p_{1,j,t}}{T_{1,j,t}}, \tag{4}$$

which shows that the export tax of Home would affect Home firms' profit.

We can write the expected discounted value of an idea as

$$v_{n,j,t} = \sum_{s=t}^{\infty} [\rho(1-\delta)]^{s-t} \frac{u_{ns}}{u_{nt}} \frac{P_{nt}}{P_{ns}} \zeta_{n,j,s}. \tag{5}$$

where future profits are discounted with each country n's own endogenous discount factor $\rho^{s-t} \frac{u_{ns}}{u_{nt}} \frac{P_{nt}}{P_{ns}}$ and idea surviving rate $1-\delta$, with u_{ns} denoting marginal utility of consumption and P_{ns} consumer price at period s.

A researcher is motivated by the possibility of coming up with an idea with value. Free mobility across sectors ensures that the present value of the expected profits of being a researcher is equal to the wage of being a worker in the production sector w, i.e., $\alpha_{n,j} v_{n,j,t} = w_{nt}$. This determines the level of R&D conducted. Workers engaged in research do not know how good their ideas will be ex-ante. Since each idea is worth $v_{n,j,t}$ in expectations, the total value of research output at time t is $\alpha_{n,j} L^r_{n,j,t} v_{n,j,t}$. The average value of a researcher is $\alpha_{n,j} v_{n,j,t}$. Total number of research workers is $L^r_{n,j,t} = r_{n,j,t} L_{n,j,t}$, where $r_{n,j,t}$ is the equilibrium share of research workers—or, research intensity. Thus,

$$\begin{aligned} \text{if } \alpha_{n,j} v_{n,j,t} &= w_{nt}, \quad \text{then } r_{n,j,t} \in [0,1], \\ \text{if } \alpha_{n,j} v_{n,j,t} &< w_{nt}, \quad \text{then } r_{n,j,t} = 0, \\ \text{if } \alpha_{n,j} v_{n,j,t} &> w_{nt}, \quad \text{then } r_{n,j,t} = 1. \end{aligned} \tag{6}$$

We now define the world market equilibrium. Variables with a prime denote variables in the next period.

Definition 1 (World Market Equilibrium)**.** *The world market equilibrium consists of an allocation of labor and consumption* $\left\{L^r_{n,j,t}, L^p_{n,j,t}, C_{n,j,t}\right\}$*, technology* $\left\{T_{n,j,t}\right\}$*, expenditures* $\left\{x_{nt}\right\}$*, prices* $\left\{P_{n,t}\right\}$*, and wages* $\left\{w_{nt}\right\}$ *such that consumers maximize expected discounted utility, firms maximize profits, and the free entry and market clearing conditions hold, for any period* t*, taking as given Home government's policies* $\left\{\tau^m_{n,j,t}, \tau^x_{n,j,t}\right\}$.[6]

1. *Free entry conditions for researchers in country* n *sector* j*:*[7]

$$\frac{w_{nt}}{\alpha_{n,j}} = \frac{1}{\theta} \frac{w_{nt} L^p_{n,j,t}}{T_{n,j,t}} + \rho(1-\delta) \frac{u_{n,t+1}/P_{n,t+1}}{u_{nt}/P_{nt}} v_{n,j,t+1}, \quad (\forall n \in N, \forall j \in J) \tag{7}$$

[6]Note again that we adopt a semi-primal approach. In the market equilibrium, Home innovation doesn't face any innovation policy; in the next section, the Home government chooses Home R&D directly. We then implement this optimal R&D choice with an innovation policy.

[7]We derive the equation (7) in two steps. First, we write recursively the expected discounted value of an idea (5) as,

$$v_{n,j,t} = \frac{1}{\theta} \frac{w_{nt} L^p_{n,j,t}}{T_{n,j,t}} + \rho(1-\delta) \frac{u_{nt+1}}{u_{nt}} \frac{P_{nt}}{P_{nt+1}} v_{n,j,t+1},$$

which uses the profit definition from (3). Second, we use the free entry condition (6) and write $w_{nt}/\alpha_{n,j} = v_{n,j,t}$ with $v_{n,j,t}$ from the above equation.

2. *Evolution of technology*

$$T_{n,j,t} = \alpha_{n,j} L^r_{n,j,t} + (1-\delta) T_{n,j,t-1}, \quad (\forall n \in N, \forall j \in J), \tag{8}$$

3. *Goods market clearing conditions*

$$\frac{1+\theta}{\theta} w_{1t} L^p_{1,j,t} = \beta_j \left[\pi_{11,j,t} x_{1t} + \sum_{i \neq 1} \frac{1}{1+\tau^x_{i,j,t}} \pi_{i1,j,t} x_{it} \right], \tag{9}$$

$$\frac{1+\theta}{\theta} w_{nt} L^p_{n,j,t} = \beta_j \left[\frac{1}{1+\tau^m_{n,j,t}} \pi_{1n,j,t} x_{1t} + \sum_{i \neq 1} \pi_{in,j,t} x_{it} \right], \tag{10}$$

where the expenditures are given by

$$x_{1t} = \frac{1+\theta}{\theta} w_{1t} \sum_j L^p_{1,j,t} + \sum_{i \neq 1}^{N} \sum_{j=1}^{J} \beta_j \frac{\tau^x_{i,j,t}}{1+\tau^x_{i,j,t}} \pi_{i1,j,t} x_{it} + \sum_{i \neq 1}^{N} \sum_{j=1}^{J} \beta_j \frac{\tau^m_{i,j,t}}{1+\tau^m_{i,j,t}} \pi_{1i,j,t} x_{1t} \tag{11}$$

$$x_{nt} = \frac{1+\theta}{\theta} w_{nt} \sum_j L^p_{n,j,t} \quad \textit{(for } n \neq 1\textit{)} \tag{12}$$

and the trade shares satisfy

$$\pi_{11,j,t} = \frac{T_{1,j,t} w_{1t}^{-\theta}}{T_{1,j,t} w_{1t}^{-\theta} + \sum_{n \neq 1} T_{n,j,t} (w_{nt}(1+\tau^m_{n,j,t}) d_{1n})^{-\theta}}, \tag{13}$$

$$\pi_{i1,j,t} = \frac{T_{1,j,t} (w_{1t}(1+\tau^x_{i,j,t}) d_{i1})^{-\theta}}{T_{1,j,t} (w_{1t}(1+\tau^x_{i,j,t}) d_{i1})^{-\theta} + \sum_{n \neq 1} T_{n,j,t} (w_{nt} d_{in})^{-\theta}}, \tag{14}$$

$$\pi_{in,j,t} = \frac{T_{n,j,t} (w_{nt} d_{in})^{-\theta}}{T_{1,j,t} (w_{1t}(1+\tau^x_{i,j,t}) d_{i1})^{-\theta} + \sum_{m \neq 1} T_{m,j,t} (w_{mt} d_{im})^{-\theta}}, \tag{15}$$

$$\pi_{1i,j,t} = \frac{T_{i,j,t} (w_{it}(1+\tau^m_{i,j,t}) d_{1i})^{-\theta}}{T_{1,j,t} w_{1t}^{-\theta} + \sum_{n \neq 1} T_{n,j,t} (w_{nt}(1+\tau^m_{n,j,t}) d_{1n})^{-\theta}}. \tag{16}$$

Note that $\sum_n \pi_{in,j,t} = 1$ *for any country* i.

The consumer price of Home is given by

$$P_{1t} = \Pi_j \left[T_{1,j,t} w_{1t}^{-\theta} + \sum_{n \neq 1} T_{n,j,t} (w_{nt}(1+\tau^m_{n,j,t}) d_{1n})^{-\theta} \right]^{-\beta_j/\theta},$$

and the consumer price of country n is

$$P_{nt} = \Pi_j \left[T_{1,j,t}(w_{1t}(1+\tau^x_{n,j,t})d_{n1})^{-\theta} + \sum_{m\neq 1} T_{m,j,t}(w_{mt}d_{nm})^{-\theta} \right]^{-\beta_j/\theta}.$$

4. *The labor market clearing conditions for each country n*

$$\sum_j \left(L^r_{n,j,t} + L^p_{n,j,t} \right) = L_n. \tag{17}$$

Note that in the world market equilibrium, the market clearing conditions and definition of expenditures imply that the balanced trade conditions hold.

Proposition 1. *Under zero export taxes and import tariffs, the world market equilibrium is efficient and a world social planner would choose the same innovation allocation as in the market equilibrium.*

Proof. See Appendix A. Note that the model has no distortions, externalities, or international spillovers. Specifically, innovation has a constant return across time $\alpha_{n,j}$, and the allocation between researchers and workers is efficient. Under the Bertrand competition, the endogenous markup of each firm follows a distribution that is invariant over time and does not depend on the destination to which the firm sells. Hence, there are identical and constant aggregate markups across all sectors.

Our model with constant $\alpha_{n,j}$ and technology depreciation δ features a steady state in the long run, where the technologies become constant for each country and each sector. Our model is isomorphic to the detrended semi-endogenous growth model in Eaton and Kortum (2001), which assumes the growth rate of α_n is exogenously driven by population growth. The two models would prescribe the same optimal policies in both the transition and the long run.

Define the research intensity in sector j country n as $r_{n,j,t} \equiv \frac{L^r_{n,j,t}}{L_{n,j,t}}$, i.e., the fraction of researchers over the total used in this sector. With this definition, we can rewrite the technology evolution (8) as $T_{n,j,t} = (1-\delta)T_{n,j,t-1} + \alpha_{n,j} r_{n,j,t} L_{n,j,t}$. At the steady state, the research intensity, labor, and technologies in each sector are all constant, in particular, $r_{n,j,t} = r_{n,j}$ at the steady state.

Proposition 2. *At the steady state, the research intensity $r_{n,j}$ is the same as in the closed economy for all sectors j in country n.*

Proof. See Appendix B. At the steady state, the optimal research intensity $r_{n,j}$ in each sector j does not depend on the country's size, research productivity, or trade openness. The reason is that accessing foreign markets increases the potential profits, but competition from foreign inventions decreases them. These two effects exactly cancel out, and the level of openness does not affect research intensity. However, the research level $r_{n,j}L_{n,j}$ would depend on size, research productivity, and openness. Thus, given the same level of research intensity $r_{n,j}$, more labor reallocated to the comparative advantage sector increases the total number of researchers in that sector, and hence its technology $T_{n,j}$.

Proposition 3 (*Tax Neutrality*)**.** *Consider two sets of Home government trade policies* $\Gamma = \{(1+\tau^m_{n,j,t}, 1+\tau^x_{n,j,t}) : \forall n \neq 1, j, t\}$ *and* $\check{\Gamma} = \left\{ \left(\lambda_t(1+\tau^m_{n,j,t}), \frac{1+\tau^x_{n,j,t}}{\lambda_t} \right) : \forall n \neq 1, j, t \right\}$ *for any sequence of positive constant* $\{\lambda_t\}_{t=0}^{\infty}$*. The welfare and allocation of the world market equilibrium under* Γ *and* $\check{\Gamma}$ *are the same. In addition, the prices* $w_{1t}(\Gamma) = w_{1t}(\check{\Gamma})$, $P_{1t}(\Gamma) = P_{1t}(\check{\Gamma})$, $w_{nt}(\Gamma) = w_{nt}(\check{\Gamma})/\lambda_t$, $P_{nt}(\Gamma) = P_{nt}(\check{\Gamma})/\lambda_t$ *and expenditure* $x_{nt}(\Gamma) = x_{nt}(\check{\Gamma})/\lambda_t$ *for* $n > 1$ *and any* t*.*

Proof. See Appendix C. Proposition 3 shows tax neutrality holds as in Lerner (1936) and Costinot and Werning (2019). As a result, *one* of the tax levels is indetermined. In the two-country and one-sector case, there is only one tariff and one export tax. Hence, we can normalize either export tax or import tariff to zero, and tax neutrality implies that import tariff and export tax are equivalent (thus Lerner symmetry). Furthermore, in the models with multiple countries or sectors, the import tariff and export tax within each country or industry are no longer equivalent. In this case, we can only set the tariff/export tax in one specific country and one specific sector equal to zero.

2.2 Optimal trade and innovation policies

The Home government chooses optimal unilateral trade policies and domestic *R&D* policies by maximizing Home households' lifetime utility (1). Foreigners are taken to be passive. Trade policy instruments are restricted to the country-sector level, comprising country-sector-specific import tariffs $\tau^m_{n,j}$ and export taxes $\tau^x_{n,j}$ directed at country $n \neq 1$. We

derive optimal domestic innovation summarized by $L^r_{1,j}$ and then examine how to implement it with domestic $R\&D$ policies. The government rebates the tax income to households in a lump-sum fashion.

We consider two types of government policies: Ramsey and Markov policies. The Ramsey government has the ability to commit to all its future policies at the beginning of time and chooses *a sequence of policies* to maximize utility— taking into consideration the private responses to the policies. In contrast, the Markov government cannot commit to future policies. In this case, the Home government chooses *current-period policies*, which are constrained to depend only on the value of the current period's state. Private individuals react by taking current and future policies as given. The government cannot commit but correctly anticipates how future policies will depend on current ones through the state of the economy.

Ramsey policy Under Ramsey optimal policies, the Home government *can* commit to the path of policies. In period 0, taking as given the initial levels of technologies $\left\{T_{n,j,0}\right\}$, Home government chooses a sequence of $\left\{L^r_{1,j,t}, \tau^x_{n,j,t}, \tau^m_{n,j,t}\right\}_{t=0}^{\infty}$ to maximize the present value of utility

$$\max_{\left\{L^r_{1,j,t}, \tau^x_{n,j,t}, \tau^m_{n,j,t}\right\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \rho^t \frac{C_{1t}^{1-\sigma}}{1-\sigma} \tag{18}$$

subject to the world market equilibrium characterized by foreign worker-researcher choices:

$$\frac{w_{nt}}{\alpha_{n,j}} \frac{u_{nt}}{P_{nt}} = \sum_{s=t}^{\infty} [\rho(1-\delta)]^{s-t} \frac{u_{ns}}{P_{ns}} \left(\frac{1}{\theta} \frac{w_{ns} L^p_{n,j,s}}{T_{n,j,s}} \right), \quad \text{for any } t, j, n \neq 1 \tag{19}$$

along with the resource constraints (8) - (17), and the normalization of Home price $P_{1t} = 1$ for any period t.

The Ramsey problem has no obvious recursive formulation since the foreign worker-research constraint (19) in each period t involves variables from all future periods $s \geq t$. To solve this problem, we extend the recursive contract approach of Marcet and Marimon (2019). The key to this approach is to define as part of the state the history of all multipliers on foreign worker-research constraints. We explain the approach briefly here and present

the full recursive problem in Appendix F.

Specifically, let $\rho^t \gamma^v_{n,j,t}$ be the multiplier on foreign worker-research constraint (19). We can write the Lagrangian as (18) plus the sum of terms relating to the worker-research constraints of the form

$$\rho^t \gamma^v_{n,j,t} \left[\sum_{s=t}^{\infty} [\rho(1-\delta)]^{s-t} \frac{u_{ns}}{P_{ns}} \left(\frac{w_{ns} L^p_{n,j,s}}{\theta T_{n,j,s}} \right) - \frac{w_{nt}}{\alpha_{n,j}} \frac{u_{nt}}{P_{nt}} \right]$$

for all t, countries $n > 1$, and sectors $j \in J$, plus standard terms relating to the resource constraints (8)-(17).

Define $\Gamma^v_{n,j,t} = \sum_{s=0}^{t} (1-\delta)^{t-s} \gamma^v_{n,j,s}$. We can regroup terms and write the Lagrangian as

$$\sum_{t=0}^{\infty} \left\{ \frac{C_{1t}^{1-\sigma}}{1-\sigma} + \sum_{n \neq 1} \sum_{j}^{J} \left(\Gamma^v_{n,j,t} \frac{w_{nt} L^p_{n,j,t}}{\theta T_{n,j,t}} - \gamma^v_{n,j,t} \frac{w_{nt}}{\alpha_{n,j}} \right) \frac{u_{nt}}{P_{nt}} \right\} \tag{20}$$

plus standard terms relating to the resource constraints. Here $\Gamma^v_{n,j,t}$ summarizes all the past multipliers, and we can write it recursively as

$$\Gamma^v_{n,j,t} = (1-\delta)\Gamma^v_{n,j,t-1} + \gamma^v_{n,j,t}, \qquad \text{for any } n \neq 1 \text{ and } j. \tag{21}$$

Therefore, the cumulative multiplier $\Gamma^v_{n,j,t}$ captures Home government's past promises on policies that affect foreign innovation incentives through constraint (19).

We can now write the Ramsey problem recursively with the state variable as $\left\{ \Gamma^v_{n,j,-1}, T_{n,j,-1} \right\}$,

$$\begin{aligned} (\textit{Ramsey}) \quad L\left(\left\{ \Gamma^v_{n,j,-1}, T_{n,j,-1} \right\} \right) = \inf_{\gamma^v_{n,j}} \sup_{\{L^r_{1,j}, \tau^x_{n,j}, \tau^m_{n,j}\}} & \frac{C_1^{1-\sigma}}{1-\sigma} + \sum_{n \neq 1} \sum_{j}^{J} \left(\Gamma^v_{n,j} \frac{w_n L^p_{n,j}}{\theta T_{n,j}} - \gamma^v_{n,j} \frac{w_n}{\alpha_{n,j}} \right) \frac{u_n}{P_n} \\ & + \cdots + \rho L\left(\left\{ \Gamma^v_{n,j}, T_{n,j} \right\} \right) \end{aligned}$$

where $\Gamma^v_{n,j} = (1-\delta)\Gamma^v_{n,j,-1} + \gamma^v_{n,j}$ and the omitted terms are related to the resource constraints. Appendix F shows the proof of the mapping between the original Ramsey problem and this recursive problem, and the derivations of the optimal conditions for the Ramsey policy.

Markov policy The non-commitment property of Markov policy allows us to write the problem in a recursive way. Here, we omit the time variable t for the current period and utilize a prime symbol to represent future values.

The Home government determines researchers in each sector j, taking into account foreign private innovation decisions and equilibrium production and trade. There are number $N \times J$ state variables, i.e., the technologies $\left\{T_{n,j,-1}\right\}$ for all country n and sector j. Specifically, the government chooses $L^r_{1,j}$ with $j \in \{1,2,..,J\}$, country-sector-specific import tariff $\tau^m_{n,j}$ and export taxes $\tau^x_{n,j}$ toward country $n > 1$ to solve the following problem:

$$(Markov) \quad V\left(\left\{T_{n,j,-1}\right\}\right) = \max_{\left\{L^r_{1,j}, \tau^x_{n,j}, \tau^m_{n,j}\right\}} \frac{C_1^{1-\sigma}}{1-\sigma} + \rho\left[V\left(\left\{T_{n,j}\right\}\right)\right] \tag{22}$$

subject to the world market equilibrium characterized by

$$\frac{w_n}{\alpha_{n,j}} = \frac{w_n L^p_{n,j}}{\theta T_{n,j}} + \rho(1-\delta)\frac{u'_n/P'_n}{u_n/P_n} v'_{nj}, \qquad (\gamma^r_{n,j} \quad J \times (N-1)) \tag{23}$$

$$\frac{1+\theta}{\theta} w_1 L^p_{1,j} = \beta_j \left[\pi_{11,j} x_1 + \sum_{i \neq 1} \frac{1}{1+\tau^x_{i,j}} \pi_{i1,j} x_i\right], \quad (\gamma^L_{1,j} \quad J) \tag{24}$$

$$\frac{1+\theta}{\theta} w_n L^p_{n,j} = \beta_j \left[\frac{1}{1+\tau^m_{n,j}} \pi_{1n,j} x_1 + \sum_{i \neq 1} \pi_{in,j} x_i\right], \quad (\gamma^L_{n,j} \quad J \times (N-1)) \tag{25}$$

the evolution of technology (8), the definition of expenditures (A.16)-(12), the labor market clearing conditions (17), and the normalization of home price $P_1 = 1$.

The major difference between the Markov and Ramsey governments is that the Ramsey government decides on the entire path of policies honored in the future. The Markov government only chooses current policies $\left\{L^r_{1,j}, \tau^x_{n,j}, \tau^m_{n,j}\right\}$, but it understands that these policies affects the technology choice $\left\{T_{n,j}\right\}$, which in turn affect future marginal utility u'_n, prices P'_n, and value $v'_{n,j}$ showing up in the worker-researcher constraints of foreign countries (23). Hence the Markov government affects future allocations and policies indirectly through current policies. Optimal conditions for the Markov policy are derived in Appendix G.

Note that it is difficult to establish the existence and uniqueness of Ramsey and Markov policies due to the non-convexity property. The challenge lies in the foreign worker-

researcher constraints, which involve future allocations and values of foreign researchers, as illustrated by equation (19) or (23). It is important to note that this challenge is prevalent in much of the existing literature with limited enforcement, for example, Kehoe and Perri (2002). Furthermore, establishing the uniqueness of optimal policies is challenging even for a static model, as pointed out by Lashkaripour and Lugovskyy (2023).[8] We assume there exists a unique solution for Ramsey or Markov problem. We then characterize the optimal policies using first-order conditions.

2.3 Home government incentive

Before characterizing the optimal Ramsey and Markov policies, we discuss intuitively Home government's incentives to use trade and innovation policies. As shown in Proposition 2, the world market equilibrium is efficient, and a world planner has no incentives to use any trade or innovation policies to distort private decisions. Home government, however, would like to impose unilateral policies to manipulate both static and dynamic terms of trade. The reason is that individual households (workers and researchers) take market prices as given and do not internalize the impact of their demand or research effort on the country's comparative advantages and relative prices, both across time and across sectors.

The key equation in our model is researchers' free entry condition, i.e.,

$$\frac{w_{nt}}{\alpha_{n,j}} = \sum_{s=t}^{\infty} [\rho(1-\delta)]^{s-t} \frac{u_{ns}/P_{ns}}{u_{nt}/P_{nt}} \zeta_{n,j,s}. \tag{26}$$

where we combine the researcher's value (5) and the free-entry condition (6). The left-hand side of the equation captures the opportunity cost of innovating, i.e., losing wage income, and the right-hand side is the value of innovating, i.e., earning lifetime profits. We can define country n's interest rate between period s and n as the inverse of the endogenous discount factor for any period $s > t$,

$$R^n_{t,s} \equiv \frac{u_{nt}/P_{nt}}{\rho^{s-t} u_{ns}/P_{ns}}. \tag{27}$$

[8] As stated in Lashkaripour and Lugovskyy (2023), '...the pseudo-uniqueness of the optimal policy formula is different from the uniqueness of the optimal policy equilibrium. Establishing the latter is a daunting task well beyond the scope of this paper.'

When the interest rate is high, the value of innovating falls, and so does the innovation incentive.

One factor that determines the value of innovating at country n is its consumer price P_{nt}, which has two opposing effects on innovation. Mechanically, as shown in Eq.27, high P_{nt} directly lowers interest rates $R^n_{t,s}$ and thus raises innovation incentives at period t. Intuitively, the substitution effect plays a role here: when goods become more expensive in period t, households lower their consumption in that period and save for the future. As a result, the interest rate falls. However, through the income effect, higher P_{nt} also lowers consumption in period t and leads to higher marginal utility u_{nt}, which tends to increase interest rates $R^n_{t,s}$ and lowers the expected value of innovation.

Under the CRRA preference (1), the interest rate (27) can be rewritten as $R^n_{t,s} = \frac{x_{nt}^{-\sigma} P_{nt}^{\sigma-1}}{\rho^{s-t} x_{ns}^{-\sigma} P_{ns}^{\sigma-1}}$. When $\sigma > 1$, the income effect dominates, and thus higher P_{nt} raises the interest rate and lowers innovation. When $\sigma = 1$, the income and substitution effects cancel out, and P_{nt} has no direct impact on the interest rate and innovation incentives of country n.

The period t price, P_{nt}, also affects all the past innovation incentives for any period $s < t$. We can rewrite the interest rate (Eq.27) for period $s < t$: $R^n_{s,t} = \frac{u_{ns}/P_{ns}}{\rho^{t-s} u_{nt}/P_{nt}}$. In this case, higher P_{nt} lowers period t consumption and raises marginal utility u_{nt}. If this income effect dominates, the interest rate falls in period s, leading to higher innovation in period $s < t$.

Hence, higher P_{nt} at period t could *lower* the innovation incentives in period t, but it would also *raise* innovation in all the past periods $s < t$. To increase P_{nt}, Home government can raise export tax on country n. It is clear that there exists time inconsistency. At period t, Home government can announce a lower export tax in a future period $s > t$ to induce lower innovation in period t. However, if the Home government cannot commit to future policies, such an announcement can not induce a lower innovation in the current period t.

In general, besides the export taxes, the Home government can also use import tariffs or domestic innovation policy to affect country n's innovation incentives in sector j. Higher import tariff $\tau^m_{n,j,t}$ directly lowers sector j's innovation profit $\zeta_{n,j,s}$ in country n, as shown in Eq.3. Moreover, to reduce foreign innovation, Home government can boost its own innovation in sector j through domestic subsidy and raise competition globally. All three

policies are feasible. What are the optimal ones? The answer depends on whether the government can commit: Ramsey versus Markov policies.

2.4 Reference: Optimal Policies under Exogenous Technology

As a reference, we first characterize the optimal policies when technology is exogenous. In this case, the worker-researcher constraint drops from the equilibrium, the goods markets are competitive, and all after-tax sales go to labor income.[9] See Appendix D for the setup and the derivations of the optimal conditions. Home government solves a static problem and chooses trade policies.

Proposition 4 (Exogenous Technology). *When technology is exogenous, optimal trade policies consist of country-specific import tariffs and country-sector-specific export taxes:*

$$\tau_{n,j}^m = \bar{\tau}_n^m \equiv -\frac{\gamma_n}{u_{c_1}}, \qquad 1+\tau_{n,j}^x = \frac{1+\theta(1-\pi_{n1,j})}{\theta\sum_{i\neq 1}(1+\bar{\tau}_i^m)\pi_{ni,j}}, \tag{28}$$

where γ_n are the multipliers on country n's goods market clearing conditions and satisfy the system of equations A.12. We define the country-specific tariff $-\gamma_n/u_{c_1}$ for each country n as $\bar{\tau}_n^m$. In the case of two countries $N=2$, optimal policies imply $\tau_{2,j}^m = 0$ and $1+\tau_{2,j}^x = \frac{1+\theta\pi_{22,j}}{\theta\pi_{22,j}}$.

Proof: See Appendix D.

When T is exogenous, optimal tariffs are uniform across sectors, and one overall level is not uniquely pinned down according to the tax neutrality in Proposition 3. With multiple countries and exogenous technology, Home would want to use tariffs to manipulate relative wages across foreign countries. Thus, tariffs vary across importers but are uniform across sectors within an importer. We show in Appendix D that these γ_n multipliers or import tariffs are simultaneously pinned down by the system of first-order conditions on wages A.12 across N countries.

[9] Under perfect competition, the world market equilibrium is characterized by the goods market clearing conditions: for each country n,

$$w_n L_n = \sum_{m\neq 1}^{N}\sum_{j=1}^{J}\pi_{mn,j}\beta_j x_m + \sum_{j=1}^{J}\frac{1}{1+\tau_{n,j}^m}\pi_{1n,j}\beta_j x_1.$$

Home government takes this condition as a constraint when solving for the optimal policies. We let the multiplier for this constraint as γ_n for country $n>1$.

Export taxes are used to exploit Home country's monopoly power, and they tend to increase with $\pi_{n1,j}$, i.e., the share of good j that country n imports from Home. In other words, the export tax for a specific country-sector increases as the market power of that sector in the destination country increases. This schedule of trade policies is consistent with the one proposed by Costinot et al. (2015), where the government can manipulate relative prices in its favor by limiting the supply of its export goods.

With endogenous technology, the Home government still has standard incentives to use tariffs and country-sector-specific export taxes to affect the static terms of trade. Moreover, it has incentives to manipulate foreign innovations both across time and across sectors. To isolate these two extra incentives, we study the dynamic technology manipulation in Section 3 and the incentive to manipulate comparative advantage in Section 4.

3 Dynamic Technology Manipulation

We first study an intertemporal motive to influence foreign innovation in one's favor. To isolate this force, we first consider a two-country-one-sector case without domestic externalities. The key takeaway is that the Markov government, unable to commit to its future trade policies, deploys domestic innovation policies in order to manipulate foreign innovation efforts over time. These results stand in sharp contrast to those under Ramsey policies. Appendix E provides a simplified two-period model to further demonstrate our results.

3.1 Ramsey policies

Due to tax neutrality, we can normalize import tariffs to zero in this two-country, one-sector model. The key feature of the optimal Ramsey policies is that the Home government does not need to undertake any innovation policies, given that it can commit to a sequence of export taxes.

Recall that γ_2^v is the multiplier on the worker-research constraint (19) and Γ_2^v the multiplier recording the past commitment of policies with $\Gamma_2^v = (1-\delta)\Gamma_{2,-1}^v + \gamma_2^v$. The following proposition characterizes the optimal Ramsey policies.

Proposition 5 (Ramsey Policy). *Under the two-country, one-sector model, the optimal Ramsey*

policy is characterized by zero tariffs and the export tax and domestic innovation formula as follows:

$$1+\tau^{x,R}=\frac{1+\theta\pi_{22}}{\theta\pi_{22}+\left(\gamma_2^v\frac{w_2}{\alpha_2}-\Gamma_2^v\frac{w_2L_2^p}{\theta T_2}\right)(\sigma-1)\frac{u_{c_2}}{P_2}\frac{1}{u_{c_1}x_2}}, \tag{29}$$

$$\frac{w_1}{\alpha_1}=\frac{w_1L_1^p}{\theta T_1}+\rho(1-\delta)\frac{u'_{c_1}/P'_1}{u_{c_1}/P_1}\frac{w'_1}{\alpha_1}. \tag{30}$$

Furthermore, $\gamma_2^v=\Gamma_2^v=0$ *at the steady state.*

Proof. See Appendix H.

It is easy to see that the Ramsey innovation condition (30) is the same as that in the private market equilibrium (7) with $v'_1=w'_1/\alpha_1$ for Home country. Hence, the Ramsey government does not impose any taxes to impact Home individuals' choices to become a worker or a researcher. Namely, the government would not impose any domestic industry policies. The reason is that with commitment, the path of export taxes is sufficient to implement the intertemporal allocation.

Compare the export tax formula (29) with the one under exogenous technology in Proposition 4: $1+\tau^x=\frac{1+\theta\pi_{22}}{\theta\pi_{22}}$, which captures the static terms of trade consideration. The extra term in the denominator of formula (29), $(\sigma-1)\left(\gamma_2^v\frac{w_2}{\alpha_2}-\Gamma_2^v\frac{w_2L_2^p}{\theta T_2}\right)\frac{u_{c_2}}{P_2}/(u_{c_1}x_2)$, reflects a dynamic technology consideration. This term depends not only on the current multiplier γ_2^v but also on the cumulative past commitments Γ_2^v, with the latter dampening the impact of the current multiplier γ_2^v and showing up as a negative term.

As we discussed in section 2.3, higher export tax τ_t^x at period t makes it more expensive for foreign consumers to acquire home goods, leading to higher foreign consumer price P_{2t}. When the income effect dominates, $\sigma>1$, higher P_{2t} makes the foreign country more beneficial to produce and less desirable to innovate. Export tax, therefore, affects the worker-researcher constraint (19). As a result, the multiplier γ_2^v of this constraint enters the export tax formula.

Furthermore, higher P_{2t} at period t induces more innovation in the past period $s<t$. When forecasting higher prices in the future, more agents would like to do research in period s. Hence, export tax τ_t^x affects all the past worker-researcher constraints (19) in period $s<t$. This is why the past cumulative multiplier Γ_{2t}^v enters the tax formula.

However, Γ^v_{2t} enters with a negative sign since the impacts τ^x_t on past innovations are opposite to the current innovation.

The sign of $\gamma^v_{2,t}$ depends on whether the home government aims to reduce or raise foreign innovations. A negative sign implies that Home prefers lower foreign innovation at period t. In other words, from the Home perspective, Foreign's innovation benefits, i.e., the right-hand side of (19), are too large relative to its innovation cost. To lower these foreign innovation benefits, Home raises the current export tax above the static terms-of-trade consideration but lowers future export tax below its static consideration $1/(\theta\pi_{22})$.

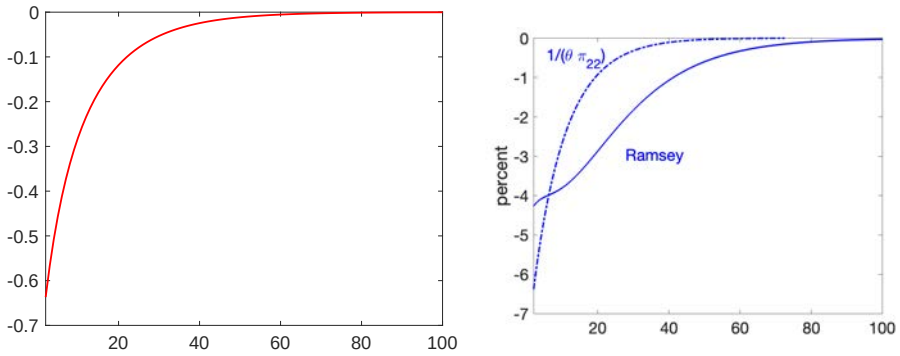
Numerical example To further understand the Ramsey policy, we consider an example where the Home government's technology T starts below the steady state, while country 2's technology is at its steady state. Figure 1 shows the evolution of technology and optimal policies employed relative to their steady-state levels during a transition path.[10]

Over time, country 1's technology increases until it reaches the steady-state level, as shown in Figure 1 (a). According to Proposition 5, the Ramsey government only utilizes export taxes, depicted in Figure 1 (b). The export tax has two salient features. First, it increases when T_1 increases over time. Second, it is higher than the static case ($1/(\theta\pi_{22})$) for low levels of T_1, and lower in later periods. Moreover, the gap between export tax and $1/(\theta\pi_{22})$ shrinks and eventually goes to zero when T_1 reaches the steady state.

These two features reinforce our previous discussions in Proposition 5. When T_1 increases over time, the home country gains monopoly power and is able to impose a higher export tax. This motive is clear from the rising $1/(\theta\pi_{22})$, capturing the static terms of trade effect. Both the initial higher and later lower export taxes than $1/(\theta\pi_{22})$ raise foreign interest rates in the early periods. This reflects the home government's motives to lower foreign innovation efforts and increase foreign production in the early periods when the home country allocates more labor to do research. In the long run, Ramsey policies have no additional dynamic incentive to alter the terms of trade, and the export tax is the same as the static one.

[10] In this example, $\theta = 2, \sigma = 2.5, \beta = 0.9, \delta = 0.02, \alpha_1 = 1, \alpha_2 = 0.9, d = 1.1$.

Figure 1: Optimal Ramsey Polices: Two-Country One-Sector



(a) T_1 (b) Export tax

3.2 Markov policies

We normalize Home's import tariff as zero, as in the Ramsey case. In each period, Markov government optimizes over current export taxes while directly choosing domestic innovation and allocations, respecting the world market equilibrium and future Markov policies. Recall that γ_2^r is the multiplier on country 2's worker-researcher choice condition (23).

Proposition 6 (Markov Policy). *Under the two-country, one-sector model, the optimal Markov policy is characterized by zero tariffs and the export tax and domestic innovation formula as follows:*

$$1+\tau^{x,M} = \frac{1+\theta\pi_{22}}{\theta\pi_{22}+\gamma_2^r(\sigma-1)\rho(1-\delta)\frac{u'_{c_2}/P'_2}{u_{c_2}/P_2}\frac{w'_2}{\alpha_2}\frac{1}{u_{c_1}x_2}}, \tag{31}$$

$$\frac{w_1}{\alpha_1} = \frac{1}{\theta}\frac{w_1 L_1^p}{T_1} + \rho(1-\delta)\frac{u'_{c_1}/P'_1}{u_{c_1}/P_1}\frac{w'_1}{\alpha_1} + \underbrace{\gamma_2^r\frac{\theta}{1+\theta}\frac{1}{u_{c_1}}\frac{\rho(1-\delta)}{u_{c_2}/P_2}\frac{\partial\left(\frac{u'_{c_2}}{P'_2}\frac{w'_2}{\alpha_2}\right)}{\partial T_1}}_{\textit{innovation wedge}}. \tag{32}$$

In addition, $\gamma_2^r = 0$ at the steady state.

Proof. See Appendix H.

What affects Foreign's innovation incentives today is the return to innovation and the relative cost of innovation. As we explain below, export taxes affect contemporaneous prices, while innovation policies are deployed to impact future prices—given the lack of commitment—and the two policy tools in the above equations combine to determine current-period foreign innovation efforts.

Export tax Relative to the exogenous case in Proposition 4, there is an extra term in the denominator of the export tax formula, $\gamma_2^r(\sigma-1)\rho(1-\delta)\frac{u'_{c_2}/P'_2}{u_{c_2}/P_2}\frac{w'_2}{\alpha_2}\frac{1}{u_{c_1}x_2}$, reflecting the dynamic technology consideration. The intuition is similar to the Ramsey case. The export tax imposed by the Home government affects foreign price P_2 and thus the foreign interest rate, which in turn determines the foreign worker-research constraint. Hence, the multiplier γ_2^r on this constraint shows up in the export tax formula. Furthermore, such impact depends on the income versus substitution effect, and thus $(\sigma-1)$ also shows up. When $\sigma=1$, the income and substitution effect cancel out, P_2 has no direct impact on the interest rate and innovation incentives of country 2, and the export tax formula is the same as the static one. However, the level of optimal export taxes under endogenous technology might be different from that of static ones.

The optimal Markov export tax formula of (31) looks similar to the one in Ramsey (29). To compare, we plug in the worker-researcher constraint (23) to Eq.31, and the Markov export tax becomes

$$1+\tau^{x,M}=\frac{1+\theta\pi_{22}}{\theta\pi_{22}+\left(\gamma_2^r\frac{w_2}{\alpha_2}-\gamma_2^r\frac{w_2}{\theta}\frac{L_2^p}{T_2}\right)(\sigma-1)\frac{1}{u_{c_1}x_2}}. \tag{33}$$

As is clear, the multiplier $\gamma_2^v u_{c_2}/P_2$ in the Ramsey corresponds to the multiplier γ_2^r in the Markov case.

However, taxes under Markov and Ramsey contain an important difference. Markov export taxes (33) depend only on the current multiplier γ_2^r associated with the foreign innovation incentive, while Ramsey export taxes $\tau^{x,R}$ depend not only on the current multiplier γ_2^v but also on past cumulative commitments $\Gamma_{2,-1}^v$. Furthermore, the past commitments $\Gamma_{2,-1}^v$ dampen the impact of the current multiplier γ_2^v on the export tax, showing up as the negative term in the denominator of Eq.29. This comparison demonstrates how the commitment feature affects optimal trade policies.

Innovation policy Because Foreign's future prices matter for current-period innovation, and because the Home country cannot commit to a path of trade policies to affect foreign future wages and prices, domestic innovation policies must be deployed. By dictating its

own next-period technology T_1, Home can influence Foreign's innovation efforts today. To see this—in the 'Euler' equation of innovation choices (32), there is an extra *innovation wedge* term when compared to the private market equilibrium. There are two key elements in this wedge term: (1) the multiplier γ_2^r, and (2) the derivative of the foreign marginal utility of innovation to T_1, denoted as $der \equiv \partial \left(\frac{u'_{c_2}}{P'_2} \frac{w'_2}{\alpha_2} \right) / \partial T_1$ for the purpose of discussion.

A higher T_1 lowers Home's future export prices, which in turn pushes down foreign price P'_2 and raises foreign consumption c'_2. As a result, foreign country has a lower marginal utility of consumption in the future, implying a higher foreign interest rate, as defined in (27). When the income effect dominates, $der \leq 0$ and the high interest rate depresses foreign innovation incentives. This affects the tightness of the foreign worker-researcher constraint (23)—the reason why the multiplier of this constraint γ_2^r shows up in the wedge.

The sign of the multiplier γ_2^r once again matters, and captures whether Home would like to reduce or raise Foreign's innovation. For instance, when T_1 is low relative to its innovation efficiency α_1, Home would like to have more research and higher consumption. Hence, Home wants Foreign to innovate less and produce more to increase Home's real consumption today. In this case, γ_2^r tends to be negative and, along with a negative *der* term, produces a positive innovation wedge. The higher the innovation wedge, the larger the Home innovation incentive as the right-hand side of the Eq.32 reflects the benefit as a researcher. The innovation wedge behaves like a subsidy to Home's innovation efforts. We can define the innovation subsidy τ_d as the innovation wedge relative to Home firms' current profit:

$$\tau_d = \frac{1}{\frac{1}{\theta} \frac{w_1 L_1^p}{T_1}} \gamma_2^r \frac{\theta}{1+\theta} \frac{1}{u_{c_1}} \frac{\rho(1-\delta)}{u_{c_2}/P_2} \frac{\partial \left(\frac{u'_{c_2}}{P'_2} \frac{w'_2}{\alpha_2} \right)}{\partial T_1}.$$

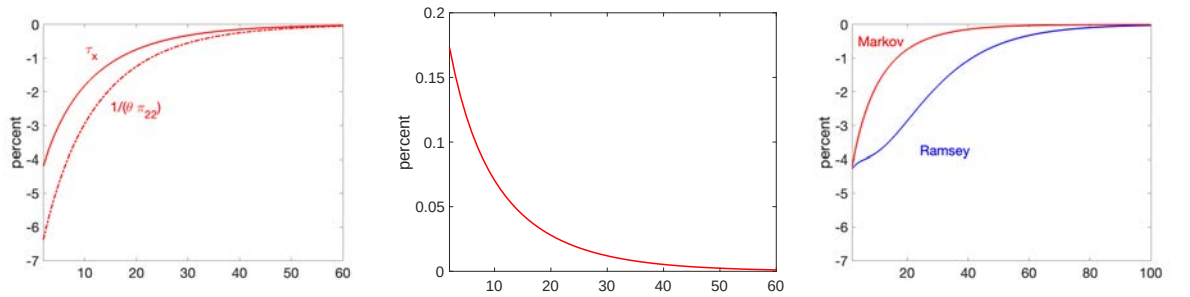
Note that γ_2^r becomes zero at the steady state. In this case, the optimal export tax is the same as in the exogenous case, and there are no innovation subsidies. The intuition is straightforward: when the technology reaches the steady state, the government can no longer use policies to affect foreign interest rates. In addition, there is no domestic externality or comparative advantage to create policy incentives.

Numerical example The Markov government utilizes both export taxes and innovation subsidies, as depicted in Figure 2, panel (a) and (b). The export tax has two salient features. First, it increases when T_1 increases over time. Second, it consistently exceeds the level in the static model of $1/(\theta\pi_{22})$. Moreover, the gap between export tax and $1/(\theta\pi_{22})$ shrinks and eventually goes to zero when T_1 increases over time.

These two features reinforce our previous discussions in Proposition 6. When T_1 increases over time, the home country gains monopoly power and is able to impose a higher export tax. This motive is clear from the rising $1/(\theta\pi_{22})$, capturing the static terms of trade effect. A higher export tax than $1/(\theta\pi_{22})$ reflects the home government's motives to lower foreign innovation efforts by raising foreign interest rates.

In addition, Home increases subsidies on domestic research efforts to further deter Foreign's. As technology progresses towards its long run value, the need to manipulate foreign innovation incentives diminishes and eventually disappears. As a result, the export tax equals $1/(\theta\pi_{22})$, and innovation subsidies become zero, in the long run.

Figure 2: Optimal Markov Polices: Two-Country, One-Sector



(a) Export tax (b) Innovation subsidy (c) Export tax: Markov & Ramsey

Figure 2(c) plots Markov export taxes against Ramsey export taxes. The path of export taxes tends to be flatter under Ramsey than under Markov. This discrepancy arises from the ability of a Ramsey government to manipulate not only the current foreign price P_2 through the current τ^x, but also future prices P_2' through commitments to future $\tau^{x\prime}$. Consequently, the government is able to avoid excessive increases in the current τ^x, which could lead to undesirable distortions. In contrast, the Markov government cannot commit to a sequence of export taxes to influence the dynamic terms of trade, and it is obliged to implement distortionary domestic industry policies such as innovation policies.

In the long run, neither Markov nor Ramsey policies have any additional incentive to alter the terms of trade. Both end up with the same steady state of technology and export taxes. Note that the steady-state innovation and trade policy becomes the same as in a static model only in the special case with one sector and no externalities. We proceed to discuss this in detail in the following section and explore why the multi-country and multi-sector model illuminates additional insights.

4 Comparative Advantage: Multi-Country, Multi-Sector Model

When comparative advantage is introduced, intratemporal considerations to manipulate foreign technology become more intricate. To make these forces transparent, we derive optimal taxes in a long-run model with multiple sectors and countries, before incorporating potential externalities and considering transition paths. In contrast to a uniform tariff applied across all sectors under exogenous technology, conventionally emphasized, optimal policies with endogenous technology consist of heterogeneous tariffs and export taxes across sectors.

Proposition 7. *With endogenous technology accumulation, at the steady state in the baseline model, the Markov and Ramsey governments choose the same trade policies. Moreover, neither government uses domestic innovation policies.*

Proof. See Appendix I. At the steady state in the baseline model, technologies and interest rates across countries stabilize. Consequently, a Markov government lacks the incentive to implement domestic policies that manipulate foreign research efforts through interest rates. As a result, both the Markov and Ramsey governments opt for the same trade policies, leading to the same allocations in both cases.

Proposition 8 (**Multi-Country Multi-Sector Endogenous Technology**)**.** *In the multi-country, multi-sector steady-state case with endogenous technology,*

1. *optimal policies do not distort domestic innovation (zero innovation taxes), but consist of heterogenous import tariffs and export taxes across countries and sectors.*

2. *On average, tariffs are higher for sectors with relatively higher net exports, for any sector j:*

$$\sum_{n\neq 1}(\tau^m_{n,j}+\xi_n)\frac{1+\theta}{\theta}w_n L^p_{n,j}=\frac{1+\theta}{\theta}w_1 L^p_{1,j}-\beta_j x_1, \tag{34}$$

Within a country, sector-specific import tariffs satisfy

$$\frac{\tau^m_{n,j}}{1+\tau^m_{n,j}}\frac{\beta_j\pi_{1n,j}x_1}{\frac{1+\theta}{\theta}w_n L^p_{n,j}}=\frac{\sum_{i\neq 1}\frac{\tau^x_{i,j}}{1+\tau^x_{i,j}}\beta_j\pi_{in,j}x_i}{\frac{1+\theta}{\theta}w_n L^p_{n,j}}+\delta_n, \tag{35}$$

where ξ_n and δ_n are endogenous and country-specific.[11]

3. *Home optimal export taxes satisfy*

$$1+\tau^x_{n,j}=\frac{1+\theta(1-\pi_{n1,j})}{\theta\sum_{i\neq 1}(1+\tau^m_{i,j})\pi_{ni,j}}. \tag{36}$$

Proof: see Appendix J.

Note that the proposition prescribes the formulae for optimal import tariffs and export taxes at the steady state. The levels of these taxes are jointly determined across countries and sectors. The reason is that tariffs on one foreign country affect *other* foreign countries' sectoral profits and incentives to do innovation—and hence technology, which in turn influence Home's incentives to impose tariffs/export taxes for other countries.

Proposition 8 shows that tariffs are now sector-specific rather than uniform as in the exogenous technology case. With multiple countries, Home's tariffs are on average higher for sectors with relatively higher net exports ($\frac{1+\theta}{\theta}w_1 L^p_{1,j}-\beta_j x_1$): Eq.34 shows that for that sector, the weighted average of Home's tariff across countries is higher. For a specific country/sector, the optimal tariff on country n is increasing in the ratio of country n's exports to other countries relative to its exports to Home, as in Eq.35. That is, if Home imports relatively more from country n, it would want to charge a relatively lower tariff, all else equal. This tariff also depends on other countries' export taxes in sector j and other general equilibrium effects, such as the impact on other countries' labor reallocation

[11] $\xi_n=\sum_{i\neq 1}^N\sum_k^J\tau^m_{i,k}\beta_k\pi_{ni,k}-\sum_k^J\beta_k\frac{\tau^x_{n,k}}{1+\tau^x_{n,k}}\pi_{n1,k}$ and $\delta_n=\frac{1}{1+\theta}\left[\theta\sum_j\tau^m_{n,j}\frac{\frac{1+\theta}{\theta}w_n L^p_{n,j}}{x_n}+\frac{\mu_n L^p_n}{u_{c_1}x_n}-1\right]$.

across sectors, as captured in Eq.35. This general equilibrium effect ceases to operate when a certain pair of goods is exclusively produced by the home country and another specific country. The following corollary outlines the optimal policies in this scenario.

Corollary 1. *In the case with two countries or the case with multiple countries but only Home and one foreign country (country 2) producing sector j and j' goods and the rest of the world having $\alpha_{ROW,j} = 0$ and $\alpha_{ROW,j'} = 0$, Home optimal tariffs satisfy,*

$$\tau^m_{2,j} - \tau^m_{2,j'} = \frac{1}{\theta}\left(\frac{\frac{1+\theta}{\theta}w_1 L^p_{1,j} - \beta_j x_1}{\frac{1+\theta}{\theta}w_2 L^p_{2,j}} - \frac{\frac{1+\theta}{\theta}w_1 L^p_{1,j'} - \beta_{j'} x_1}{\frac{1+\theta}{\theta}w_2 L^p_{2,j'}}\right). \tag{37}$$

Thus, Home's import tariff on sector j increases with its net exports normalized by production in country 2, relative to sector j'.

Proof. See Appendix K. With only Home and another country producing j and j' goods, the general formula (35) becomes (37), where import tariffs are directly linked to the net exports. Corollary 1 clearly demonstrates the Home government's incentive to reduce foreign innovation efforts in sectors where Home has a comparative advantage. If Home exports more of the sector's goods, it would prefer to levy a higher tariff; conversely, if Home imports a relatively greater amount of the goods, it would charge a relatively lower tariff. As for the level of optimal tariffs, multiple considerations are taken into account: the impact of its tariffs on Foreign and Home's own labor and production, Foreign profits that affect its technology, Home's own tariff revenue, consumption price, and terms of trade through changes in wages. Despite the complexity, the sectoral optimal tariff can be summarized by Home's net export normalized by foreign sector output.

In general, we can show that the optimal tariffs satisfy

$$\tau^m_{n,j} = -\frac{\gamma^L_{n,j}}{u_{c_1}} = -\frac{\gamma^r_{n,j}}{(1+\theta)T_{n,j}u_{c_1}} + Const_n, \tag{38}$$

where $\gamma^L_{n,j}$ is the multiplier on the goods market clearing condition (25) and $\gamma^r_{n,j}$ is the multiplier on the foreign worker-researchers constraint (23). The multiplier $\gamma^r_{n,j}$ reflects the impact of a foreign country's excess incentive of innovation on Home welfare, and $\gamma^L_{n,j}$ reflects the impact of excess demand of a foreign good on Home welfare. This equation

demonstrates that Home would like to use tariffs to affect Foreign labor and production across sectors, hence profit and incentive for innovation.

Appendix I proves the optimal innovation policy at the steady state of the baseline model. The Home government chooses to not distort its own R&D, but would like to use heterogenous tariffs to influence Foreign's innovation, as shown in Eq.38. The Home government would want to use export taxes to exploit the country's monopoly power, in conjunction with heterogenous import tariffs across sectors to influence Foreign's innovation. Moreover, in the multi-country case, these taxes depend on how much other countries export to country n ($\pi_{ni,j}$), reflecting the degree of monopoly power Home has on country n in sector j.

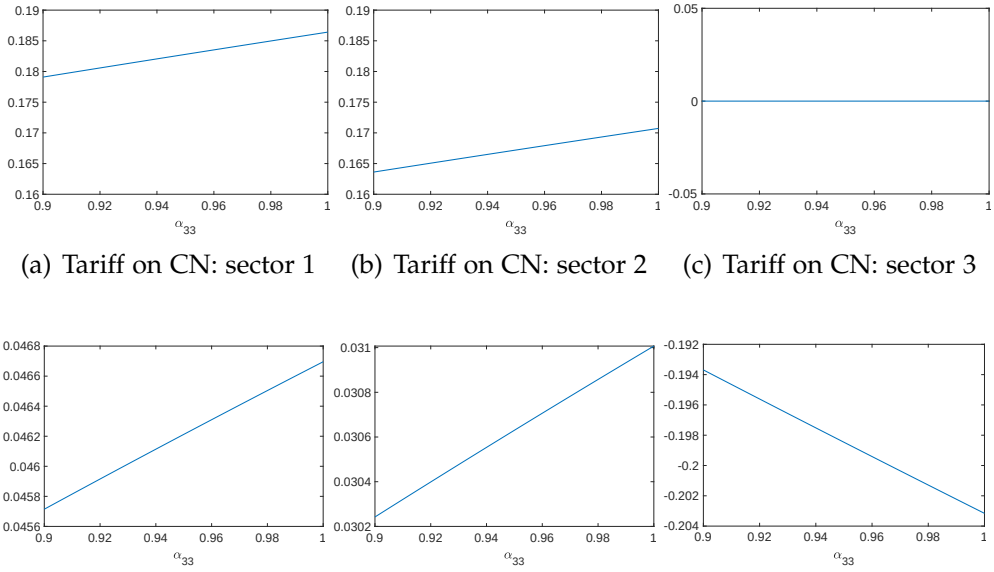
Lastly, it is important to emphasize the difference between the two-country and multi-country cases. In a two-country context, one country's net exports in a sector are necessarily the other country's net imports. However, in the case of multiple countries, there could be sectors in which both Home and Foreign are net exporters to the rest of the world. This distinction has important implications for optimal policies. In the two-country case, only one country has the incentive to levy a tariff in a given sector j. However, in the multi-country scenario, it is possible that both Home and Foreign are net exporters of certain goods, and *both* countries' unilateral policies would favor high tariffs in those sectors. Moreover, an event that raises both Home and Foreign's net exports of good j, for instance, a fall in trade costs or a rise in global demand in sector j, will also lead to an increase in the unilateral tariffs they levy on each other in that sector.

Numerical example The interlinking relationships and spillover of tariffs across countries and sectors may appear to be complex but follow the basic principle outlined in the above proposition. That is— the within-country tariff ranking has to do with the country's comparative advantage. That strict ranking disappears when we deviate from the special case, but still follows some basic trade patterns. To gain some intuition, and to investigate predictions of theory when cast in realistic scenarios, we consider an example where a third country becomes larger. Appendix L also considers a few other cases.

There are three sectors (S1, S2, and S3) and three economies: the United States (US),

China (CN), and the rest of the world (ROW). The innovation efficiencies in the three sectors accord with US $(1, 0.9, 0.9)$, CN $(0.9, 0.9, 1)$, ROW $(0, 0, \alpha_{33})$ with α_{33} varying from 0.9 to 1. Since ROW's innovation efficiency is zero and cannot produce goods in S1 and S2, it imports them from the United States and China. Furthermore, S2 is China's comparative advantage (CA) sector, and S1 is U.S.'s comparative advantage sector.

Figure 3: Optimal Policies for the United States when Global Demand Rises



(a) Tariff on CN: sector 1 (b) Tariff on CN: sector 2 (c) Tariff on CN: sector 3

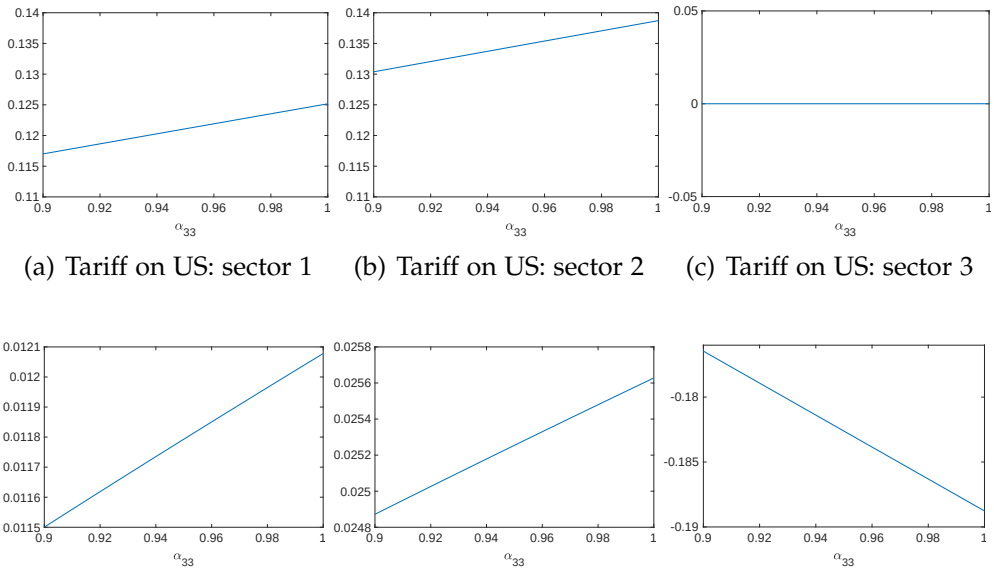
(d) US net export: sector 1 (e) US net export: sector 2 (f) US net export: sector 3

Note: The upper panel plots the U.S. optimal tariff on imports from China in sector 1, sector 2, and sector 3 when sector 3 in ROW becomes more efficient, i.e. higher α_{33}. The lower panel plots the corresponding US's net exports in sectors 1, 2, and 3.

To start with, we increase α_{33} in the first example to make ROW larger. This led to a surge in the global demand for goods in S1 and S2. Note that the U.S. always imposes tariffs on S1, which is its CA sector. Larger global demand for S1 and S2 boosts exports in these sectors, and thus the U.S. increases tariffs on both sectors. Moreover, the tariff on S2 increases by more as the net export in this sector rises faster.

Figure 3 shows that the U.S. increases tariffs on China's sector 2 goods more than tariffs on sector 1 goods. In addition, the U.S. imposes higher tariffs on ROW's sector 3 goods, aiming to induce Chinese labor to shift from sector 1 and 2 to sector 3. However, the impact of this policy on the rise of Chinese labor in sector 3 is not significant. The same optimal policy applies symmetrically to China, which levies a higher tariff on the US goods in sectors 1 and 2 but higher in sector 2, see Figure 4. The larger the rise in global demand,

Figure 4: Optimal Policies for China when Global Demand Rises



(a) Tariff on US: sector 1 (b) Tariff on US: sector 2 (c) Tariff on US: sector 3

(d) CN net export: sector 1 (e) CN net export: sector 2 (f) CN net export: sector 3

Note: The upper panel plots China's optimal tariff on imports from the U.S. in sectors S1, S2, and S3 when S3 in ROW becomes more efficient, i.e. higher α_{33}. The lower panel plots China's net exports in the three sectors.

the larger the tariffs levied by both China and the U.S. to endogenously improve each country's own technology in that sector.

We also explore optimal policies under several alternative scenarios in Appendix L. Specifically, we consider a case when there is an increase in China S1's innovation efficiency α_{11}. In this case, China levies a relatively higher tariff on imports from U.S. sector 1 to discourage innovation in the U.S., see Figure A-3. In addition, we study a case when there is a rise in the importance of sector 1 goods. Figure A-4 demonstrates that both the US and China would raise tariffs on importing sector 1 goods from the other country.

5 Optimal Policies with Externality

Having established separately the two key motives of government policies—to manipulate dynamic technology accumulation and intra-period terms of trade—we proceed to examine policies during transition when these two effects are simultaneously present. We conclude the analysis of the full-fledged multi-country-multi-sector model by introducing an externality in technology accumulation. This externality serves two purposes. First, it

has crucial implications for optimal policies. For a Markov government, even at the steady state, a conventional tax to target the externality is insufficient to implement optimal policies. There would need to be additional industrial policies even in the long run. Second, it helps with computation in dealing with corner solutions.

The domestic externality is hereby modeled with returns to scale in the innovation effort. Specifically, the innovation efficiency in country n sector j depends on a constant $\overline{\alpha}_{n,j}$ and the research effort in this sector, $\alpha_{n,j,t} = \overline{\alpha}_{n,j}(L^r_{n,j,t})^{\varepsilon-1}$, where $\varepsilon < 1$ captures the externality. If $\varepsilon = 1$, the economy returns to the efficient baseline case with constant $\overline{\alpha}_{n,j}$.

We assume that all countries can use the Pigouvian innovation tax to correct the domestic externality. In addition, foreign countries are still passive and do not use policies other than this Pigouvian tax. In this case, the worker-researcher choice constraint in country $n \neq 1$ becomes

$$\frac{w_n}{\overline{\alpha}_{n,j}(L^r_{n,j})^{\varepsilon-1}} = \varepsilon \frac{w_n L^p_{n,j}}{\theta T_{n,j}} + \tilde{\rho}_n v'_{n,j}, \qquad (\gamma^r_{n,j}), \tag{39}$$

where the parameter ε in front of $\frac{w_n L^p_{n,j}}{\theta T_{n,j}}$ reflects that foreign country n has used Pigouvian innovation tax.

Proposition 9 (Ramsey Policy). *The Ramsey government chooses export tax, import tariff, and domestic innovation subsidy $\tau^{d,R}_j$ as*

$$1 + \tau^{x,R}_{n,j} = \frac{1 + \theta(1 - \pi_{n1,j})}{\theta \sum_{i \neq 1}(1 + \tau^{m,R}_{i,j})\pi_{ni,j} + (\sigma - 1)\sum_k^J \left(\gamma^v_{n,k} \frac{w_n}{\overline{\alpha}_{n,j}(L^r_{n,j})^{\varepsilon-1}} - \varepsilon \Gamma^v_{n,k} \frac{w_n L^p_{n,k}}{\theta T_{n,k}} \right) \frac{u_{c_n}}{P_n} / (u_{c_1} x_n)} \tag{40}$$

$$\tau^{m,R}_{n,j} = -\frac{\gamma^L_{n,j}}{u_{c_1}} = -\frac{\varepsilon \Gamma^v_{n,j}}{(1+\theta) T_{n,j} u_{c_1} P_n / u_{c_n}} + Const_n, \tag{41}$$

$$\tau^{d,R}_j = \varepsilon - 1, \tag{42}$$

where $\gamma^v_{n,j}$ and $\gamma^L_{n,j}$ are the multipliers for the worker-researcher condition and labor market clearing condition, respectively, and $\Gamma^v_{n,k} = (1-\delta)\Gamma^v_{n,k,-1} + \gamma^v_{n,k}$. In addition, in the steady state, $\sum_k^J \gamma^v_{n,k} \frac{w_n}{\overline{\alpha}_{n,j}(L^r_{n,j})^{\varepsilon-1}} = 0$ and $\sum_k \varepsilon \Gamma^v_{n,k} \frac{w_n L^p_{n,k}}{\theta T_{n,k}} = 0$, for any foreign country n, even with $\varepsilon \neq 1$.

Proof. See Appendix M.

For the Ramsey government, a constant tax rate ($\varepsilon < 1$ hence a negative subsidy) on innovation is sufficient to correct the externalities. While industrial policies fix externalities, optimal trade policies address terms of trade/technological competition as before, for both static and dynamic considerations. The Ramsey government also considers its past promises on trade policies with $\Gamma^v_{n,k}$ showing up in the export tax formula.

The optimal tariffs exhibit a similar structure to that of the baseline case, related to the multipliers of foreign worker-researcher constraints. Ramsey tariffs depend on all past multipliers $\Gamma^v_{n,j}$. Clearly, tariffs determine a country's cross-sector innovation benefit: in the dynamic setting, innovation benefits in period s are also part of innovation benefits for periods $m < s$. The Ramsey government internalizes the impact of period s tariffs on all the past worker-researcher tradeoffs, and thus $\Gamma^v_{n,j}$ shows up in the tariff formula.

With externalities, labor allocation also affects innovation efficiency $\overline{\alpha}_{n,j}(L^r_{n,j})^{\varepsilon-1}$. Note that $\varepsilon < 1$ means that innovation taxes lower the number of researchers and endogenously strengthen the comparative advantage sectors and hence raise the incentives to conduct trade policies.

Most importantly, through labor allocations, tariffs determine a country's cross-sector average cost of innovation $\frac{w_n}{\overline{\alpha}_{n,j}(L^r_{n,j})^{\varepsilon-1}}$, which affects the tightness of the constraint $\gamma^v_{n,j}$. In sum, the Ramsey government does not use additional domestic innovation policies other than the constant Pigouvian tax, relying solely on a path of trade policies.

Proposition 10 (Markov Policy). *The Markov policy is characterized by the following export tax, import tariff, and innovation subsidy $\tau^{d,M}_j$, for $n > 1$ any sector j*

$$1+\tau^{x,M}_{n,j} = \frac{1+\theta(1-\pi_{n1,j})}{\theta\sum_{i\neq 1}(1+\tau^{m,M}_{i,j})\pi_{ni,j} + \left(\sum_k \gamma^r_{n,k}(\sigma-1)\tilde{\rho}_n v_{n,k}\right)/(u_{c_1}x_n)}, \tag{43}$$

$$\tau^{m,M}_{n,j} = -\frac{\gamma^L_{n,j}}{u_{c_1}} = -\frac{\varepsilon\gamma^r_{n,j}}{(1+\theta)T_{n,j}u_{c_1}} + Const_n, \tag{44}$$

$$\tau^{d,M}_j = (\varepsilon-1) + \underbrace{\frac{\varepsilon}{\frac{1}{\theta}\frac{w_1 L^p_{1,j}}{T_{1,j}}}\frac{\theta}{1+\theta}\frac{\rho(1-\delta)}{u_{c_1}}\sum_{n\neq 1}^{N}\sum_k^J \frac{1}{u_{c_n}/P_n}\gamma^r_{n,k}\frac{\partial\left(u'_{c_n}\frac{w'_n/P'_n}{\overline{\alpha}_{n,k}(L^{r\prime}_{n,k})^{\varepsilon-1}}\right)}{\partial T_{1,j}}}_{\text{additional subsidies}}. \tag{45}$$

where $\gamma^r_{n,j}$ and $\gamma^L_{n,j}$ are the multipliers for the worker-researcher condition and labor market clearing condition, respectively.

Proof. See Appendix N.

The optimal export tax $\tau^x_{n,j}$ in this full model encapsulates comparative advantage—the static terms of trade effect—as in the long-run case (36). In addition, it captures the home government's incentive to manipulate country n's technology accumulation in all sectors, similar to the one-sector case in (31). This dynamic incentive is captured by the multiplier $\gamma^r_{n,k}$ on the worker-researcher constraints, for all sector k in country n.

In our numerical example shown below, we find that in the long run, the second term $\left(\sum_k \gamma^r_{n,k}(\sigma-1)\tilde{\rho}_n v_{n,k}\right)/(u_{c_1}x_n)$ in the denominator of (43) is zero even with externalities. This implies that export taxes have the same formula as ones without externalities in equation (36). The reason is that the government uses export taxes to manipulate dynamic terms of trade, directly through P_n. In the steady state, the relative prices across countries remain constant. Hence, the dynamic aspect of the export tax disappears, and the second term in the denominator becomes zero, even with externalities.

The optimal import tariff $\tau^m_{n,j}$ has a similar formula to the one in the long run. The externality shows up directly in (44) simply because we allow foreign countries to correct the externalities through innovation taxes, as shown in Eq.39. Home government chooses country-sector-specific tariffs to affect both innovation effort and innovation efficiency across sectors. First, as in the case without externality, the government uses these tariffs to influence the relative prices to induce the desired innovation efforts across different sectors in a foreign country. Second, with decreasing returns to scale in innovation, the change in innovation effort further influences innovation efficiency $\overline{\alpha}_{n,j}(L^r_{n,j})^{\varepsilon-1}$ directly.

The optimal innovation effort satisfies the Euler equation with an extra wedge relative to the private market equilibrium. The innovation consists of two wedges. The first term $(\varepsilon - 1)$ is constant and reflects the standard incentive of industrial policies, i.e., correcting the externality. When $\varepsilon = 1$, the first term disappears, and the formula returns to the efficient case. The second term is related to dynamic manipulation, as shown in the multipliers $\gamma^r_{n,k}$ and the derivatives of future discounted value to $T_{1,j}$. Hence, the government internalizes the impact of its technology in sector j on the technology accumulation in any sector k in

other countries.

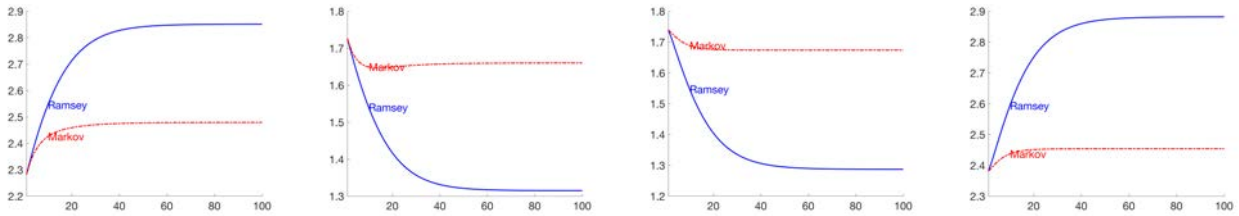
Notably, when $\varepsilon < 1$ for some k, the wedge and the innovation subsidy are not zero, even in the long run. In contrast, when there are no externalities, the wedge equals zero at steady state, as shown in Proposition 8. The reason is that the Markov government cannot commit to using future tariffs to adjust the next period's $L^{r\prime}_{n,k}$, which affects the technology of country n even when all T reaches the long-run steady state. It is obliged to implement extra distortionary innovation policies even at the SS, reflected by $\frac{\partial\left(u'_{cn}\frac{w'_n/P'_n}{\bar{\alpha}_{n,k}(L^{r\prime}_{n,k})^{\varepsilon-1}}\right)}{\partial T_{1,j}}$ in the extra innovation policy term. At the steady state, the Ramsey government can use committed tariffs to affect future labor allocation hence innovation efficiency in the Foreign and thus does not impose the same policies as Markov in the case of externalities. The Markov government cannot use future tariffs to affect Foreign's future labor allocation and future efficiency across sectors even in the steady state. Home use distortionary innovation policy to change their own technology, which in turn affects Foreign's future labor allocation and hence future efficiency across sectors.

Computation and numerical example It is worth noting the computational challenges involved in solving this multi-country and multi-sector model. The Markov problem includes a large number of state variables, specifically, $N \times J$. In a two-country, two-sector model, for instance, there would be four state variables. But then the Ramsey planner needs to consider all past cumulative multipliers $\Gamma^v_{n,j}$, which raises the total number of state variables to $N \times J + (N-1) \times J$. Another problem is that without decreasing return to scale or any adjustment costs, during transition, innovation tends to run into corner solutions, where $L^r_{n,j}$ be binding at zero for some sectors and countries. The equilibrium under decreasing return to scale circumvents the issue of corner solutions but still requires guessing the derivatives of future allocations and prices with respect to current technology choices. To check the robustness of our computations, we solve the models with both global methods under sparse grids and perturbation methods. The two methods give identical results, and they are consistent with our theoretical results.

Consider an example with two countries, Home (country 1) and Foreign (country 2), and two sectors. Home has the comparative advantage in sector 1 in that its α_{11}/α_{12} is higher

than foreign α_{21}/α_{22}. In period 1, α_{11} increases and stays at this high level.[12]

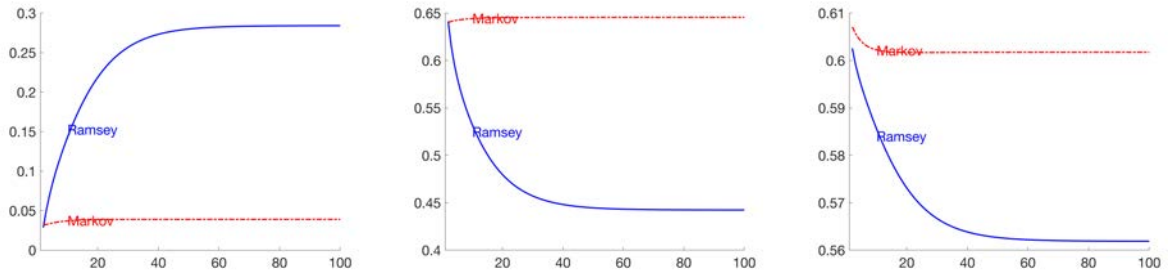
Figure 5: Technology under Ramsey vs Markov: multi-sector with externality



(a) $T_{1,1}$ (b) $T_{1,2}$ (c) $T_{2,1}$ (d) $T_{2,2}$

Figures 5 - 7 compare Markov versus Ramsey during the transition path and at the steady state. Both cases start with the same technology levels, i.e., the steady state of Markov technologies before the change of α_{11}. Figure 5 plots the evolution of technologies $\{T_{1,1}, T_{1,2}, T_{2,1}, T_{2,2}\}$. The Markov and Ramsey equilibrium share some salient features. When α_{11} increases, on impact, the return of researchers in sector 1 of Home country becomes higher, which draws more researchers into sector 1. As a result, $T_{1,1}$ increase over time, and $T_{1,2}$ falls at the Home country. The technology in Country 2 moves in the opposite direction, as the demand for sector 1 goods decreases. Markov differs from Ramsey in that the increase in $T_{1,1}$ in Markov is smaller. The technology gap between sectors within a country is also smaller in Markov in the long run. These are due to the different policies Ramsey and Markov use in the long run when there are externalities.

Figure 6: Trade policies under Ramsey vs Markov: multi-sector with externality

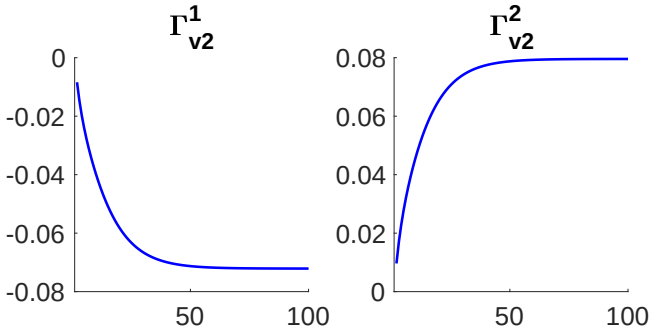


(a) $\tau^m_{2,1}$ (b) $\tau^x_{2,1}$ (c) $\tau^x_{2,2}$

Figure 6 compares trade policies under Ramsey and Markov. We normalize the tariff

[12] In this example, $\theta = 2$, $\sigma = 2.5$, $\rho = 0.9$, $\delta = 0.02$, $d = 1.4$, and $\varepsilon = 0.99$ for all countries and sectors. In addition, $\alpha_{22} = 1$ and $\alpha_{12} = \alpha_{21} = 0.9$ at all the time. At period 0, $\alpha_{11} = 0.95$, and it increases to 1 at period 1 and stays at this value.

Figure 7: Ramsey multiplier: multi-sector with externality



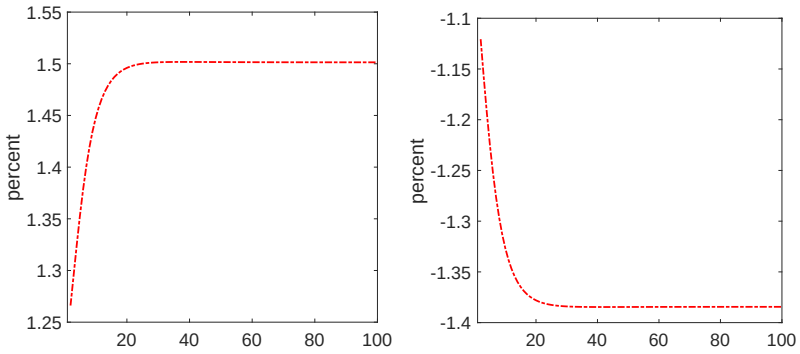
on sector 2 imports to zero. Given that Home has the comparative advantage in Sector 1, Home would like to impose a tariff on this sector, see Figure 6 (a). Over time, Sector 1 becomes more productive and experiences larger exports, leading to higher import tariffs in both Markov and Ramsey. However, the increase in tariff in Ramsey is larger and more persistent than that in Markov. With this large and persistent increase in import tariff, Ramsey government changes the export taxes according to equation (40).

To understand Figure 6, let us revisit the impact of tariffs on innovating incentives. With or without externality, higher tariffs in sector 1 reduce the sector's profit and discourage innovation or research $L^r_{2,1}$. When $\varepsilon < 1$, an opposing force is also present. Decreasing return to scale ($\varepsilon < 1$) together with lower $L^r_{2,1}$ boosts the innovation efficiency $\alpha_{21}(L^r_{2,1})^{\varepsilon-1}$. As a result, the opportunity cost of innovating $w_2/[\alpha_{21}(L^r_{2,1})^{\varepsilon-1}]$ falls, encouraging innovation in sector 1. These two effects, the lower benefit versus the lower opportunity cost, work against each other and dampen the government's incentive for innovation when $\varepsilon < 1$. This explains why the tariff increase in Markov is small.

The two effects are also present in Ramsey. However, the Ramsey government considers the impact of the current tariff on all past benefits of foreign innovation, while the opportunity cost only shows up in one period. As a result, the larger benefit induces higher tariffs in Ramsey. Figure 7 shows that the Ramsey policy wants Foreign to do less innovation in sector 1 ($\Gamma^v_{2,1} < 0$) and more in sector 2, and for this reason, levies a tariff on sector 1 imports.

Unable to commit to raising tariffs in the future, the Markov government resorts to innovation policies. Figure 8 exhibits the additional innovation subsidies deployed in Markov, which is the second term in Eq.45. The classic industry policy would prescribe an innova-

Figure 8: Additional innovation subsidies under Markov: multi-sector with externality



(a) Innovation subsidy: sector 1 (b) Innovation subsidy: sector 2

tion tax, independent across sectors and from trade, when the externality parameter $\varepsilon < 1$ is constant and the same across sectors. In contrast, even with constant ε, the innovation subsidies in Markov problem vary across sectors and time. The first reason is that the Home government subsidizes sector 1 during the transition. Second, even if the Foreign government has used industry policies to correct the externality, Home still uses extra innovation policies at the steady state. The reason is that Home can affect foreign's future marginal innovation cost $\frac{w'_n}{\alpha_{n,k}(L^{r\prime}_{n,k})^{\varepsilon-1}}$ by affecting $L^{r\prime}_{n,k}$ through T_1.

In summary, when externalities are present, both Ramsey and Markov utilize domestic innovation policies in conjunction with heterogeneous export taxes and import tariffs across sectors. However, their policies diverge at the steady state. Ramsey, after imposing the conventional externality tax, relies solely on trade policies. In contrast, Markov requires both trade policies and additional innovation policies to achieve the desired outcomes.

6 Conclusion

Our intention in this paper has been to examine optimal policies for countries when new technologies emerge or when the rest of the world grows. The question can be fully answered only by examining a dynamic model, where there are multiple goods in a multiple-region world economy. Two important motives for governments—a dynamic manipulation of foreign technology and an intra-period terms of trade effect–underlie optimal policies. Markov optimal policies of a country invoke innovation and trade policies, even when private innovation allocations are efficient. Innovation policies can be used to manipulate

the benefit and cost of foreign innovation investment in the absence of commitment. In contrast to the Markov government, Ramsey optimal policies are able to avoid deploying innovation policies that distort its own innovation investment by invoking trade policies (heterogenous export tax and tariff across sectors and over time). The full-fledged model is important but entails non-trivial technical challenges. And yet, our model yields explicit expressions for optimal policies and more general results which make the mechanisms transparent. These results stand in sharp contrast to the standard models with exogenous technology, where optimal policies call for uniform tariffs across sectors. Future work can take up an analysis of a more quantitative nature, engage with other models of technology competition, or with the various state-of-the-art developments in trade theory.

References

Akcigit, Ufuk, Sina Ates, and Giammario Impullitti. 2019. "Innovation, trade policy, and globalization." *Globalization* :09–30.

Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple. 2018. "Innovation and production in the global economy." *American Economic Review* 108 (8):2128–73.

Atkeson, Andrew and Ariel Tomas Burstein. 2010. "Innovation, firm dynamics, and international trade." *Journal of political economy* 118 (3):433–484.

Bagwell, Kyle and Robert W Staiger. 1999. "An economic theory of GATT." *American Economic Review* 89 (1):215–248.

Bai, Yan, Keyu Jin, Dan Lu, and Hanxi Wang. 2024. "Optimal Trade Policy with International Technology Diffusion." Tech. rep., National Bureau of Economic Research.

Bartelme, Dominick G, Arnaud Costinot, Dave Donaldson, and Andres Rodriguez-Clare. 2019. "The textbook case for industrial policy: Theory meets data." Tech. rep., National Bureau of Economic Research.

Bloom, Nicholas, Mirko Draca, and John Van Reenen. 2016. "Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity." *The review of economic studies* 83 (1):87–117.

Broda, Christian, Nuno Limao, and David E Weinstein. 2008. "Optimal tariffs and market power: the evidence." *American Economic Review* 98 (5):2032–65.

Buera, Francisco J and Ezra Oberfield. 2020. "The global diffusion of ideas." *Econometrica* 88 (1):83–114.

Cai, Jie, Nan Li, and Ana Maria Santacreu. 2022. "Knowledge diffusion, trade, and innovation across countries and sectors." *American Economic Journal: Macroeconomics* 14 (1):104–45.

Caliendo, Lorenzo and Fernando Parro. 2021. "Trade policy." .

Costinot, Arnaud, Dave Donaldson, Jonathan Vogel, and Iván Werning. 2015. "Comparative advantage and optimal trade policy." *The Quarterly Journal of Economics* 130 (2):659–702.

Costinot, Arnaud, Guido Lorenzoni, and Iván Werning. 2014. "A theory of capital controls as dynamic terms-of-trade manipulation." *Journal of Political Economy* 122 (1):77–128.

Costinot, Arnaud and Andrés Rodríguez-Clare. 2014. "Trade theory with numbers: Quantifying the consequences of globalization." In *Handbook of international economics*, vol. 4. Elsevier, 197–261.

Costinot, Arnaud, Andrés Rodríguez-Clare, and Iván Werning. 2020. "Micro to macro: Optimal trade policy with firm heterogeneity." *Econometrica* 88 (6):2739–2776.

Costinot, Arnaud and Iván Werning. 2019. "Lerner symmetry: A modern treatment." *American Economic Review: Insights* 1 (1):13–26.

Demidova, Svetlana. 2017. "Trade policies, firm heterogeneity, and variable markups." *Journal of International Economics* 108:260–273.

Demidova, Svetlana and Andrés Rodríguez-Clare. 2009. "Trade policy under firm-level heterogeneity in a small economy." *Journal of International Economics* 78 (1):100–112.

Eaton, Jonathan and Samuel Kortum. 1999. "International technology diffusion: Theory and measurement." *International Economic Review* 40 (3):537–570.

———. 2001. "Technology, trade, and growth: A unified framework." *European economic review* 45 (4-6):742–755.

———. 2002. "Technology, geography, and trade." *Econometrica* 70 (5):1741–1779.

Gros, Daniel. 1987. "A note on the optimal tariff, retaliation and the welfare loss from tariff wars in a framework with intra-industry trade." *Journal of international Economics* 23 (3-4):357–367.

Grossman, Gene M and Elhanan Helpman. 1990. "Trade, innovation, and growth." *The American economic review* 80 (2):86–91.

———. 1993. *Innovation and growth in the global economy*. MIT press.

Hsieh, Chang-Tai, Peter J Klenow, and Ishan B Nath. 2019. "A Global View of Creative Destruction." Tech. rep., National Bureau of Economic Research.

Kehoe, Patrick J and Fabrizio Perri. 2002. "International business cycles with endogenous incomplete markets." *Econometrica* 70 (3):907–928.

Kortum, Samuel S. 1997. "Research, patenting, and technological change." *Econometrica: Journal of the Econometric Society* :1389–1419.

Krugman, Paul. 1994. "Does third world growth hurt first world prosperity?" *Harvard Business Review* 72 (4):113–121.

Lashkaripour, Ahmad and Volodymyr Lugovskyy. 2023. "Profits, scale economies, and the gains from trade and industrial policy." *American Economic Review* 113 (10):2759–2808.

Lerner, Abba P. 1936. "The symmetry between import and export taxes." *Economica* 3 (11):306–313.

Liu, Ernest and Song Ma. 2021. "Innovation Networks and Innovation Policy." Tech. rep., National Bureau of Economic Research.

Marcet, Albert and Ramon Marimon. 2019. "Recursive contracts." *Econometrica* 87 (5):1589–1631.

Ossa, Ralph. 2014. "Trade wars and trade talks with data." *American Economic Review* 104 (12):4104–46.

Perla, Jesse, Christopher Tonetti, and Michael E Waugh. 2021. "Equilibrium technology diffusion, trade, and growth." *American Economic Review* 111 (1):73–128.

Somale, Mariano. 2021. "Comparative advantage in innovation and production." *American Economic Journal: Macroeconomics* 13 (3):357–96.

Online Appendix to "Technological Rivalry and Optimal Dynamic Policy in an Open Economy"

by Yan Bai, Keyu Jin, and Dan Lu

This appendix is organized as follows.

A. Proof for Proposition 1: market equilibrium efficiency

B. Proof for Proposition 2: market equilibrium at steady state

C. Proof for Proposition 3: tax neutrality

D. Proof for Proposition 4: optimal policies with exogenous technology

E. A Simplified Two-Period Model

F. Characterization of Ramsey Policy in the Baseline

G. Characterization of Markov Policy in the Baseline

H. Proof for Proposition 5 and 6: optimal dynamic policies one-sector two-country

I. Proof for Proposition 7

J. Proof for Proposition 8: optimal dynamic policies with multiple countries at SS

K. Proof for Corollary 1: optimal dynamic policies with two countries at SS

L. Numerical examples

M. Proof for Proposition 9: optimal Ramsey policies

N. Proof for Proposition 10: optimal Markov policies

A Proof for Proposition 1: market equilibrium efficiency

This section shows that with international transfers across countries, a world social planner would choose the same innovation as in the private market equilibrium. Let the transfers be $\{\kappa_n\}$ with

$\sum_{n=1}^{N} \kappa_n = 0$. The world planner solves the following problem:

$$V\left(\left\{T_{n,j,-1}\right\}\right) = \max_{\left\{L^r_{n,j}, L^p_{n,j} T_{n,j}, w_n, x_n\right\}} \sum_{n=1}^{N} \lambda_n \frac{\left(x_n / P_n\right)^{1-\sigma}}{1-\sigma} + \rho E\left[V\left(\left\{T_{n,j}\right\}\right)\right]$$

subject to:

$$\begin{aligned}
& T_{n,j} = \alpha_{n,j} L^r_{n,j} + (1-\delta) T_{n,j,-1} \quad (\gamma^T_{n,j} \quad J \times N) \\
& \frac{1+\theta}{\theta} w_n L^p_{n,j} = \beta_j \sum_m \pi_{mn,j} x_m \quad (\gamma^L_{n,j} \quad J \times N) \\
& \sum_j \left(L^r_{n,j} + L^p_{n,j}\right) = L_n \quad (\mu_n \quad N) \\
& x_n = \frac{1+\theta}{\theta} w_n \sum_j L^p_{n,j} + \kappa_n \quad (\gamma_{xn} \quad N) \\
& \sum_{n=1}^{N} \kappa_n = 0 \quad (\gamma_k)
\end{aligned}$$

where

$$\begin{aligned}
P_m &= \Pi_j \left[T_{1,j} (w_1 d_{m1})^{-\theta} + \sum_{n \neq 1} T_{n,j} (w_n d_{mn})^{-\theta} \right]^{-\frac{\beta_j}{\theta}} \\
\pi_{11,j} &= \frac{T_{1,j} (w_1)^{-\theta}}{T_{1,j} (w_1)^{-\theta} + \sum_{n \neq 1} T_{n,j} (w_n d_{1n})^{-\theta}} \\
\pi_{m1,j} &= \frac{T_{1,j} (w_1 d_{m1})^{-\theta}}{T_{1,j} (w_1 d_{m1})^{-\theta} + \sum_{n \neq 1} T_{n,j} (w_n d_{mn})^{-\theta}} \\
\pi_{mn,j} &= \frac{T_{n,j} (w_n d_{mn})^{-\theta}}{T_{1,j} (w_1 d_{m1})^{-\theta} + T_{n,j} (w_n d_{mn})^{-\theta} + \sum_{i \neq \{1,n\}} T_{i,j} (w_i d_{mi})^{-\theta}} \\
\pi_{1m,j} &= \frac{T_{m,j} (w_m d_{1m})^{-\theta}}{T_{1,j} (w_1)^{-\theta} + T_{m,j} (w_m d_{1m})^{-\theta} + \sum_{n \neq \{1,m\}} T_{n,j} (w_n d_{1n})^{-\theta}}
\end{aligned}$$

Note that $\sum_n \pi_{mn,j} = 1$ for any m.

Given all the market clearing conditions, the following balanced trade conditions hold.

$$\sum_{m=1}^{N} \sum_{j=1}^{J} \beta_j \pi_{nm,j} x_n = \sum_{m=1}^{N} \sum_{j=1}^{J} \beta_j \pi_{mn,j} x_m + \kappa_n.$$

First order conditions

Over x_n

$$\lambda_n \frac{u_{cn}}{P_n} + \sum_{i=1}^{N} \sum_j \gamma_{i,j}^L \beta_j \pi_{ni,j} - \gamma_{xn} = 0 \tag{A.1}$$

over $L_{n,j}^r$

$$\mu_n = \gamma_{n,j}^T \alpha_{n,j} \tag{A.2}$$

over $L_{n,j}^p$

$$\mu_n = \left(\gamma_{xn} - \gamma_{n,j}^L\right) \frac{1+\theta}{\theta} w_n$$

hence all sectors have the same γ_n^L, and

$$\mu_n = \left(\gamma_{xn} - \gamma_n^L\right) \frac{1+\theta}{\theta} w_n$$

furthermore, use (A.2), we get

$$\gamma_{n,j}^T = \left(\gamma_{xn} - \gamma_n^L\right) \frac{1+\theta}{\theta} \frac{w_n}{\alpha_{n,j}} \tag{A.3}$$

over $T_{n,j}$

$$-\gamma_{n,j}^T + \rho E\left[\gamma_{n,j}^{T\prime}\left(1-\delta\right)\right] + \sum_{i=1}^{N} \lambda_i u_{ci} \frac{x_i}{P_i} \frac{\beta_j}{\theta} \frac{\pi_{in,j}}{T_{n,j}} + \sum_{i=1}^{N} \gamma_{i,j}^L \beta_j \sum_{m=1} \frac{\partial \pi_{mi,j}}{\partial T_{n,j}} x_m = 0$$

Plugging in derivatives and use all sectors have the same γ_n^L,

$$\begin{aligned}
&-\gamma_{n,j}^T + \rho E\left[\gamma_{n,j}^{T\prime}\left(1-\delta\right)\right] + \sum_{i=1}^{N} \lambda_i u_{ci} \frac{x_i}{P_i} \frac{\beta_j}{\theta} \frac{\pi_{in,j}}{T_{n,j}} \\
&+ \sum_{i\neq n}^{N} \gamma_i^L \beta_j \sum_{m=1} \left(-\pi_{mi,j}\pi_{mn,j}\left(T_{n,j}\right)^{-1}\right) x_m + \gamma_n^L \beta_j \sum_{m=1} \left(\pi_{mn,j}(1-\pi_{mn,j})\left(T_{n,j}\right)^{-1}\right) x_m = 0
\end{aligned}$$

organize

$$\gamma_{n,j}^T = \rho\left(1-\delta\right) E\left[\gamma_{n,j}^{T\prime}\right] + \frac{1}{T_{n,j}} \sum_{i=1}^{N} \lambda_i \frac{u_{ci}}{P_i} x_i \frac{\beta_j}{\theta} \pi_{in,j} - \frac{1}{T_{n,j}} \left[\sum_{i=1}^{N} \gamma_i^L \beta_j \sum_{m=1} \pi_{mi,j}\pi_{mn,j} x_m\right] + \frac{1}{T_{n,j}} \left[\gamma_n^L \beta_j \sum_{m=1} \pi_{mn,j} x_m\right]$$

Using market clearing conditions $\frac{1+\theta}{\theta} w_n L_{n,j}^p = \beta_j \sum_m \pi_{mn,j} x_m$,

$$\gamma_{n,j}^T = \frac{1}{T_{n,j}} \gamma_n^L \frac{1+\theta}{\theta} w_n L_{n,j}^p + \rho\left(1-\delta\right) E\left[\gamma_{n,j}^{T\prime}\right] + \frac{1}{T_{n,j}} \sum_{i=1}^{N} \lambda_i \frac{u_{ci}}{P_i} x_i \frac{\beta_j}{\theta} \pi_{in,j} - \frac{1}{T_{n,j}} \left[\sum_{i=1}^{N} \gamma_i^L \beta_j \sum_m \pi_{mi,j}\pi_{mn,j} x_m\right]$$

Plugging in $\gamma_{n,j}^T$ (Eq.A.3), finally the FOC for $T_{n,j}$ becomes

$$\left(\gamma_{xn}-\gamma_n^L\right)\frac{1+\theta}{\theta}\frac{w_n}{\alpha_{n,j}}=\frac{1}{T_{n,j}}\gamma_n^L\frac{1+\theta}{\theta}w_nL_{n,j}^p+\rho\left(1-\delta\right)E\left[\left(\gamma'_{xn}-\gamma_n'^L\right)\frac{1+\theta}{\theta}\frac{w'_n}{\alpha_{n,j}}\right]+\frac{1}{T_{n,j}}\sum_{i=1}^N\lambda_i\frac{u_{ci}}{P_i}x_i\frac{\beta_j}{\theta}\pi_{in,j}$$
$$-\frac{1}{T_{n,j}}\left[\sum_{i=1}^N\gamma_i^L\beta_j\sum_m\pi_{mi,j}\pi_{mn,j}x_m\right] \tag{A.4}$$

Over w_n

$$-\sum_{j=1}^J\sum_{i=1}^N\lambda_iu_{ci}\frac{x_i}{P_i}\beta_j\frac{\pi_{in,j}}{w_n}-\gamma_n^L\sum_{j=1}^J\frac{1+\theta}{\theta}L_{n,j}^p+\gamma_{xn}\sum_{j=1}^J\frac{1+\theta}{\theta}L_{n,j}^p+\sum_{i=1}^N\sum_j^J\gamma_i^L\beta_j\sum_m\frac{\partial\pi_{mi,j}}{\partial w_n}x_m=0$$

plugging in derivatives,

$$-\sum_{j=1}^J\sum_{i=1}^N\lambda_iu_{ci}\frac{x_i}{P_i}\beta_j\frac{\pi_{in,j}}{w_n}-\gamma_n^L\sum_{j=1}^J\frac{1+\theta}{\theta}L_{n,j}^p+\gamma_{xn}\sum_{j=1}^J\frac{1+\theta}{\theta}L_{n,j}^p$$
$$+\sum_{i=1}^N\sum_j^J\gamma_i^L\beta_j\sum_{m=1}\theta\pi_{mi,j}\pi_{mn,j}w_n^{-1}x_m-\sum_j^J\gamma_n^L\beta_j\sum_{m=1}^N\theta\pi_{mn,j}w_n^{-1}x_m=0$$

reorganize,

$$\sum_{j=1}^J\sum_{i=1}^N\lambda_iu_{ci}\frac{x_i}{P_i}\beta_j\pi_{in,j}+\gamma_n^L(x_n-\kappa_n)-\gamma_{xn}(x_n-\kappa_n)-\sum_{i=1}^N\sum_j^J\gamma_i^L\beta_j\sum_{m=1}\theta\pi_{mi,j}\pi_{mn,j}x_m+\theta\sum_j^J\gamma_n^L\beta_j\sum_{m=1}^N\pi_{mn,j}x_m=0$$

use market clearing conditions,

$$\sum_{j=1}^J\sum_{i=1}^N\lambda_iu_{ci}\frac{x_i}{P_i}\beta_j\pi_{in,j}+\gamma_n^L(x_n-\kappa_n)-\gamma_{xn}(x_n-\kappa_n)-\theta\sum_{m=1}^N\sum_{i=1}^N\gamma_i^L\sum_j^J\beta_j\pi_{mi,j}\pi_{mn,j}x_m+\theta\gamma_n^L(x_n-\kappa_n)=0 \tag{A.5}$$

With international transfers, no wedge on innovation The FOC for transfers implies

$$\gamma_{xi}=\gamma_{xn}=\gamma_x.$$

Guess with transfers κ_i, $\gamma_i^L=0$ for any i, then we verify the FOCs hold with this guess.

The FOC for x_n (Eq.A.1) becomes

$$\lambda_i\frac{u_{ci}}{P_i}=\gamma_x$$

In this case

$$\sum_{j=1}^{J}\sum_{i=1}^{N}\lambda_i u_{ci}\frac{x_i}{P_i}\beta_j\pi_{in,j}=\gamma_x(x_n-\kappa_n)$$

and FOCs on wages (Eq.A.5) become

$$\gamma_n^L(x_n-\kappa_n)-\theta\sum_{m=1}^{N}\sum_{i=1}^{N}\gamma_i^L\sum_{j}^{J}\beta_j\pi_{mi,j}\pi_{mn,j}x_m+\theta\gamma_n^L(x_n-\kappa_n)=0$$

which holds when $\gamma_i^L=0$.

Now we check FOCs of Ts, with $\gamma_{xi}=\gamma_x$ and $\gamma_i^L=0$, equation (A.4) becomes:

$$\gamma_x\frac{1+\theta}{\theta}\frac{w_n}{\alpha_{n,j}}=\rho\left(1-\delta\right)E\left[\gamma_x'\frac{1+\theta}{\theta}\frac{w_n'}{\alpha_{n,j}}\right]+\frac{1}{\theta}\frac{1+\theta}{\theta}\gamma_x\frac{w_nL_{n,j}^p}{T_{n,j}}$$

further reorganize it,

$$\frac{w_n}{\alpha_{n,j}}=\frac{1}{\theta}\frac{w_nL_{n,j}^p}{T_{n,j}}+\rho\left(1-\delta\right)E\left[\frac{\frac{u_{ci}'}{P_i'}}{\frac{u_{ci}}{P_i}}\frac{w_n'}{\alpha_{n,j}}\right].$$

It is the same as the private innovation equation. Hence, there are no wedges in the innovation, and the world social planner chooses the same innovation allocation as in the private market equilibrium.

B Proof for Proposition 2: market equilibrium at steady state

At the steady state, the evolution of technology implies,

$$\delta T_{n,j}=\alpha_{n,j}L_{n,j}^r=\alpha_{n,j}r_{n,j}L_{n,j}, \tag{A.6}$$

where the second equality uses the definition of research intensity, $r_{n,j}=L_{n,j}^r/L_{n,j}$, i.e., the share of research labor in sector j of country n. The free-entry condition implies that private research efforts satisfy

$$\frac{w_n}{\alpha_{n,j}}=\frac{1}{\theta}\frac{w_nL_{n,j}^p}{T_{n,j}}+\tilde{\rho}_n v_{n,j}'. \tag{A.7}$$

Combining the equation (A.6), (A.7), the definition of $\tilde{\rho} = \rho(1-\delta)\frac{u_n' P_n}{u_n P_n'}$, and $u_n' = u_n$, $P_n' = P_n$ at the steady state, we find that the steady-state research intensity for any country n and sector j is given by

$$r_{n,j}^{ss} = \frac{\delta}{\theta\left[1-\rho(1-\delta)\right]+\delta}.$$

Hence, the research intensity is constant across sectors and countries. At the steady state, technology satisfies $T_{n,j} = \frac{\alpha_{n,j} r_{n,j} L_{n,j}}{\delta}$.

C Proof for Proposition 3: tax neutrality

This section provides the proof of tax neutrality in our model (as in Lerner (1936), Costinot and Werning (2019)). Given $\Gamma = \{(\tau_{n,j,t}^m + 1, \tau_{n,j,t}^x + 1) : \forall n \neq 1, j, t\}$ and $\check{\Gamma} = \{(\check{\tau}_{n,j,t}^m + 1, \check{\tau}_{n,j,t}^x + 1) : \forall n \neq 1, j, t\}$, let $\mathcal{E}(\Gamma)$ denote the set of $\{T_{n,j,t}, \pi_{1n,j,t}, \pi_{mn,j,t}, \frac{w_{nt}}{P_{nt}}, \frac{x_{nt}}{P_{nt}}, L_{n,j,t}^p, L_{n,j,t}^r\}$ that form an equilibrium. We say that from Γ to $\check{\Gamma}$ is neutral if $\mathcal{E}(\Gamma)$=$\mathcal{E}(\check{\Gamma})$. This captures neutrality because the equilibrium allocations and welfare under Γ and $\check{\Gamma}$ are the same.

We assume $\check{\tau}_{n,j,t}^m + 1 = \lambda_t(\tau_{n,j,t}^m + 1)$ and $\check{\tau}_{n,j,t}^x + 1 = \lambda_t(\tau_{n,j,t}^x + 1)$. We guess the allocations $\{\check{T}_{n,j,t}, \check{\pi}_{1n,j,t}, \check{\pi}_{mn,j,t}, \frac{\check{w}_{nt}}{\check{P}_{nt}}, \frac{\check{x}_{nt}}{\check{P}_{nt}}, \check{L}_{n,j,t}^p, \check{L}_{n,j,t}^r\}$ in the new equilibrium have the following relations with allocations in the old equilibrium: $\check{T}_{n,j,t} = T_{n,j,t}, \check{\pi}_{1n,j,t} = \pi_{1n,j,t}, \check{\pi}_{mn,j,t} = \pi_{mn,j,t}, \check{L}_{n,j,t}^p = L_{n,j,t}^p, \check{L}_{n,j,t}^r = L_{n,j,t}^r, \check{P}_{1t} = P_{1t} = 1, \check{w}_{1t} = w_{1t}, \check{x}_{1t} = x_{1t}$, and $\check{P}_{nt} = \frac{P_{nt}}{\lambda_t}, \check{w}_{nt} = \frac{w_{nt}}{\lambda_t}, \check{x}_{nt} = \frac{x_{nt}}{\lambda_t}$, for $n > 1$. We then verify that the new equilibrium conditions are satisfied.

Free entry conditions for researchers

$$\frac{\check{w}_{nt}}{\alpha_{n,j}} = \frac{1}{\theta}\frac{w_{nt} L_{n,j,t}^p}{\lambda_t T_{n,j,t}} + \rho(1-\delta)\frac{u_n' P_n/\lambda_t}{u_n P_n'/\lambda_{t+1}}\frac{w_{n,t+1}}{\lambda_{t+1}} = \frac{w_{nt}}{\lambda_t \alpha_{n,j}}$$

Evolution of technology

$$\check{T}_{n,j,t} = \alpha_{n,j} L_{n,j,t}^r + (1-\delta) T_{n,j,t-1} = T_{n,j,t}$$

Price index

$$\check{P}_{1t} = \Pi_j \left[T_{1,j,t} w_{1t}^{-\theta} + \sum_{n\neq 1} T_{n,j,t} \left(\frac{w_{nt}}{\lambda_t}\lambda_t (1+\tau_{n,j,t}^m) d_{1n}\right)^{-\theta} \right]^{-\frac{\beta_j}{\theta}} = P_{1t} = 1$$

$$\check{P}_{mt} = \Pi_j \left[T_{1,j,t} \left(w_{1t}\frac{1+\tau_{m,j,t}^x}{\lambda_t} d_{m1}\right)^{-\theta} + \sum_{n\neq 1} T_{n,j,t} \left(\frac{w_{nt}}{\lambda_t} d_{mn}\right)^{-\theta} \right]^{-\frac{\beta_j}{\theta}} = \frac{P_{mt}}{\lambda_t}$$

Market clearing conditions

$$\frac{1+\theta}{\theta}\check{w}_{1t}\check{L}^{p}_{1,j,t} = \beta_j\left[\pi_{11,j,t}x_{1t} + \sum_{m\neq 1}\frac{\lambda_t}{1+\tau^{x}_{m,j,t}}\pi_{m1,j,t}\frac{x_{mt}}{\lambda_t}\right] = \frac{1+\theta}{\theta}w_{1t}L^{p}_{1,j,t}$$

$$\frac{1+\theta}{\theta}\check{w}_{nt}\check{L}^{p}_{n,j,t} = \beta_j\left[\frac{1}{\lambda_t(1+\tau^{m}_{n,j,t})}\pi_{1n,j,t}x_{1t} + \sum_{m\neq 1}\pi_{mn,j,t}\frac{x_{mt}}{\lambda_t}\right] = \frac{1+\theta}{\theta}\frac{w_{nt}}{\lambda_t}L^{p}_{n,j,t}$$

Trade shares

$$\check{\pi}_{11,j,t} = \frac{T_{1,j,t}w_{1t}^{-\theta}}{T_{1,j,t}w_{1t}^{-\theta} + \sum_{n\neq 1}T_{n,j,t}(\frac{w_{nt}}{\lambda_t}\lambda_t(1+\tau^{m}_{n,j,t})d_{1n})^{-\theta}} = \pi_{11,j,t}$$

$$\check{\pi}_{m1,j,t} = \frac{T_{1,j,t}(w_{1t}\frac{(1+\tau^{x}_{m,j,t})}{\lambda_t}d_{m1})^{-\theta}}{T_{1,j,t}(w_{1t}\frac{(1+\tau^{x}_{m,j,t})}{\lambda_t}d_{m1})^{-\theta} + \sum_{n\neq 1}T_{n,j,t}(\frac{w_{nt}}{\lambda_t}d_{mn})^{-\theta}} = \pi_{m1,j,t}$$

$$\check{\pi}_{mn,j,t} = \frac{T_{n,j,t}(\frac{w_{nt}}{\lambda_t}d_{mn})^{-\theta}}{T_{1,j,t}(w_{1t}\frac{(1+\tau^{x}_{m,j,t})}{\lambda_t}d_{m1})^{-\theta} + T_{n,j,t}(\frac{w_{nt}}{\lambda_t}d_{mn})^{-\theta} + \sum_{i\neq\{1,n\}}T_{i,j,t}(\frac{w_{it}}{\lambda_t}d_{mi})^{-\theta}} = \pi_{mn,j,t}$$

$$\check{\pi}_{1m,j,t} = \frac{T_{m,j,t}(\frac{w_{mt}}{\lambda_t}\lambda_t(1+\tau^{m}_{m,j,t})d_{1m})^{-\theta}}{T_{1,j,t}w_{1t}^{-\theta} + T_{m,j,t}(\frac{w_{mt}}{\lambda_t}\lambda_t(1+\tau^{m}_{m,j,t})d_{1m})^{-\theta} + \sum_{n\neq\{1,m\}}T_{n,j,t}(\frac{w_{nt}}{\lambda_t}\lambda_t(1+\tau^{m}_{n,j,t})d_{1n})^{-\theta}} = \pi_{1m,j,t}$$

Labor market clearing conditions

$$\sum_j\left(\check{L}^{r}_{n,j,t} + \check{L}^{p}_{n,j,t}\right) = \sum_j\left(L^{r}_{n,j,t} + L^{p}_{n,j,t}\right) = L_{nt}$$

Country m's expenditure ($m \neq 1$)

$$\check{x}_{mt} = \frac{1+\theta}{\theta}\frac{w_{mt}}{\lambda_t}\sum_j L^{p}_{m,j,t} = \frac{x_{mt}}{\lambda_t}$$

Country 1's expenditure

$$\begin{aligned}
\check{x}_{1t} &= \frac{1+\theta}{\theta}w_{1t}\sum_j L^{p}_{1,j,t} + \sum_{m\neq 1}^{N}\sum_j\beta_j(1-\frac{\lambda_t}{1+\tau^{x}_{m,j,t}})\pi_{m1,j,t}\left(\frac{x_{mt}}{\lambda_t}\right) + \sum_{i\neq 1}^{N}\sum_j\beta_j(1-\frac{1}{\lambda_t(1+\tau^{m}_{i,j,t})})\pi_{1i,j,t}x_{1t} \\
&= \frac{1+\theta}{\theta}w_{1t}\sum_j L^{p}_{1,j,t} + \sum_{m\neq 1}^{N}\sum_j\beta_j\frac{1-\lambda_t+\tau^{x}_{m,j,t}}{\lambda_t(1+\tau^{x}_{m,j,t})}\pi_{m1,j,t}x_{mt} + \sum_{i\neq 1}^{N}\sum_j\beta_j\frac{\lambda_t-1+\lambda_t\tau^{m}_{i,j,t}}{\lambda_t(1+\tau^{m}_{i,j,t})}\pi_{1i,j,t}x_{1t} \\
&= \frac{1+\theta}{\theta}w_{1t}\sum_j L^{p}_{1,j,t} + \sum_{m\neq 1}^{N}\sum_j\beta_j(\frac{\tau^{x}_{m,j,t}}{1+\tau^{x}_{m,j,t}} + \frac{1-\lambda_t}{\lambda_t})\pi_{m1,j,t}x_{mt} + \sum_{i\neq 1}^{N}\sum_j\beta_j(\frac{\tau^{m}_{i,j,t}}{1+\tau^{m}_{i,j,t}} + \frac{\lambda_t-1}{\lambda_t(1+\tau^{m}_{i,j,t})})\pi_{1i,j,t}x_{1t} \\
&= \frac{1+\theta}{\theta}w_{1t}\sum_j L^{p}_{1,j,t} + \sum_{m\neq 1}^{N}\sum_j\beta_j\frac{\tau^{x}_{m,j,t}}{1+\tau^{x}_{m,j,t}}\pi_{m1,j,t}x_{mt} + \sum_{i\neq 1}^{N}\sum_j\beta_j\frac{\tau^{m}_{i,j,t}}{1+\tau^{m}_{i,j,t}}\pi_{1i,j,t}x_{1t} = x_{1t}
\end{aligned}$$

where the fourth equation used the balanced trade condition for country 1

$$\sum_{m\neq 1}^{N}\sum_{j}\beta_j\pi_{m1,j,t}x_{mt}=\sum_{i\neq 1}^{N}\sum_{j}\beta_j\frac{1}{1+\tau_{i,j,t}^{m}}\pi_{1i,j,t}x_{1t}$$

Hence, the home expenditure equation can be satisfied under the guessed equilibrium allocations.

We confirm that the guessed allocations and prices satisfy all the equilibrium conditions. This completes the proof of tax neutrality in this model.

D Proof for Proposition 4: optimal policies with exogenous technology

We normalize the price of Home's final goods to 1 and denote countries by m, n. The technology of all countries is exogenously given, and the markets are competitive. The Home country government chooses $\left\{\tau_{n,j}^{x},\tau_{n,j}^{m}\right\}$ to maximize the value of utility subject to the private market equilibrium conditions:

$$\max\ u(x_1/P_1)=\frac{(x_1/P_1)^{1-\sigma}}{1-\sigma}$$

subject to for any period:

$$x_1=w_1L_1+\sum_{m\neq 1}^{N}\sum_{j=1}^{J}\beta_j\frac{\tau_{m,j}^{x}}{1+\tau_{m,j}^{x}}\pi_{m1,j}x_m+\sum_{i\neq 1}^{N}\sum_{j=1}^{J}\beta_j\frac{\tau_{i,j}^{m}}{1+\tau_{i,j}^{m}}\pi_{1i,j}x_1,\quad(\gamma_x)$$

$$w_nL_n=\sum_{m\neq 1}^{N}\sum_{j=1}^{J}\pi_{mn,j}\beta_jx_m+\sum_{j=1}^{J}\frac{1}{1+\tau_{n,j}^{m}}\pi_{1n,j}\beta_jx_1,\quad(\gamma_n,n>1,\quad(N-1))$$

where x_m, and $\pi_{mn,j}$ are given by equation (12) - (16), and

$$P_1=\Pi_j\left[T_{1,j}(w_1)^{-\theta}+\sum_{n\neq 1}T_{n,j}(w_n(1+\tau_{n,j}^{m})d_{1n})^{-\theta}\right]^{-\frac{\beta_j}{\theta}}=1.$$

FOC over x_1

$$u_{c_1}+\sum_{n=1}^{N}\sum_{j}\gamma_n\beta_j\frac{1}{1+\tau_{n,j}^{m}}\pi_{1n,j}-\gamma_x\left(1-\sum_{n\neq 1}^{N}\sum_{j=1}^{J}\beta_j\frac{\tau_{n,j}^{m}}{1+\tau_{n,j}^{m}}\pi_{1n,j}\right)=0.\tag{A.8}$$

FOC over $\tau^x_{n,j}$

$$\sum_{i\neq 1}^{N}\gamma_i\left[\theta\pi_{ni,j}\pi_{n1,j}\beta_j x_n\right]+\gamma_x\beta_j\frac{1}{(1+\tau^x_{n,j})}\pi_{n1,j}x_n-\gamma_x\beta_j\frac{\tau^x_{n,j}}{1+\tau^x_{n,j}}\theta\pi_{n1,j}(1-\pi_{n1,j})x_n=0$$

reorganize it, and we have optimal export taxes satisfy

$$1+\tau^x_{n,j}=\frac{\left[1+\theta(1-\pi_{n1,j})\right]}{\theta\sum_{m\neq 1}(1-\gamma_m/\gamma_x)\pi_{nm,j}}$$

FOC over $\tau^m_{n,j}$

$$\begin{aligned}&-u_{c_1}x_1\beta_j\frac{\pi_{1n,j}}{1+\tau^m_{n,j}}-\gamma_n\beta_j\frac{\pi_{1n,j}}{\left(1+\tau^m_{n,j}\right)^2}x_1+\gamma_n\frac{1}{1+\tau^m_{n,j}}\frac{\partial\pi_{1n,j}}{\partial\tau^m_{n,j}}\beta_j x_1+\sum_{i\neq\{1,n\}}\gamma_i\frac{1}{1+\tau^m_{i,j}}\frac{\partial\pi_{1i,j}}{\partial\tau^m_{n,j}}\beta_j x_1\\&+\gamma_x\beta_j\frac{1}{\left(1+\tau^m_{n,j}\right)^2}\pi_{1n,j}x_1+\gamma_x\sum_{i\neq 1}^{N}\beta_j\frac{\tau^m_{i,j}}{1+\tau^m_{i,j}}\frac{\partial\pi_{1i,j}}{\partial\tau^m_{n,j}}x_1=0\end{aligned}$$

plugging in derivatives, we get the optimal import tariff

$$\frac{1}{1+\tau^m_{n,j}}=\frac{u_{c_1}+\gamma_x\left[\theta\pi_{11,j}+\sum_{i\neq\{1,n\}}^{N}(\frac{1-\gamma_i/\gamma_x}{1+\tau^m_{i,j}})\theta\pi_{1i,j}\right]}{(\gamma_x-\gamma_n)\left[1+\theta\left(1-\pi_{1n,j}\right)\right]}\tag{A.9}$$

Combine the FOCs of tariffs (A.9) and of x_1 (A.8), we get

$$\tau^m_{n,j}=-\frac{\gamma_n}{u_{c_1}},\quad \gamma_x=u_{c_1}\tag{A.10}$$

and

$$1+\tau^x_{n,j}=\frac{\left[1+\theta(1-\pi_{n1,j})\right]}{\theta\sum_{i\neq 1}(1+\tau^m_{i,j})\pi_{ni,j}}.\tag{A.11}$$

These imply the optimal trade policies under exogenous T feature country-specific tariffs and country-sector-specific export taxes.

The country-specific tariffs are determined by γ_n, the multiplier on the market clearing conditions. Those conditions pin down relative wages in the equilibrium. Hence, in the government problem, the following FOCs on wages pin down γ_n, hence tariffs.

FOC over w_1

$$
\begin{aligned}
&- u_{c_1} x_1 \sum_j^J \beta_j \frac{\pi_{11,j}}{w_1} + \sum_{n\neq 1}^N \sum_j^J \beta_j \gamma_n \frac{1}{1+\tau_{n,j}^m} \frac{\partial \pi_{1n,j}}{\partial w_1} x_1 + \sum_{n\neq 1}^N \gamma_n \sum_{m\neq 1}^N \sum_j^J \beta_j \frac{\partial \pi_{mn,j}}{\partial w_1} x_m \\
&+ \gamma_x L_1 + \gamma_x \sum_{m\neq 1}^N \sum_j^J \beta_j \frac{\tau_{m,j}^x}{1+\tau_{m,j}^x} \frac{\partial \pi_{m1,j}}{\partial w_1} x_m + \gamma_x \sum_{i\neq 1}^N \sum_j^J \beta_j \frac{\tau_{i,j}^m}{1+\tau_{i,j}^m} \frac{\partial \pi_{1i,j}}{\partial w_1} x_1 = 0
\end{aligned}
$$

plugging in derivatives

$$
\begin{aligned}
&- u_{c_1} x_1 \sum_j^J \beta_j \pi_{11,j} + \sum_{n\neq 1}^N \sum_j^J \beta_j \gamma_n \frac{1}{1+\tau_{n,j}^m} \theta \pi_{1n,j} \pi_{11,j} x_1 + \sum_{n\neq 1}^N \gamma_n \sum_{m\neq 1}^N \sum_j^J \beta_j \theta \pi_{mn,j} \pi_{m1,j} x_m + \gamma_x w_1 L_1 \\
&- \gamma_x \sum_{m\neq 1}^N \sum_j^J \beta_j \frac{\tau_{m,j}^x}{1+\tau_{m,j}^x} \theta \pi_{m1,j} (1-\pi_{m1,j}) x_m + \gamma_x \sum_{i\neq 1}^N \sum_j^J \beta_j \frac{\tau_{i,j}^m}{1+\tau_{i,j}^m} \theta \pi_{1i,j} \pi_{11,j} x_1 = 0
\end{aligned}
$$

FOC over w_n for $n \neq 1$

$$
\begin{aligned}
&- u_{c_1} x_1 \sum_j^J \beta_j \frac{\pi_{1n,j}}{w_n} + \gamma_x \sum_{m\neq 1}^N \sum_j^J \beta_j \frac{\tau_{m,j}^x}{1+\tau_{m,j}^x} \frac{\partial \pi_{m1,j}}{\partial w_n} x_m + \gamma_x \sum_j^J \beta_j \frac{\tau_{n,j}^x}{1+\tau_{n,j}^x} \pi_{n1,j} \frac{x_n}{w_n} + \gamma_x \sum_{i\neq 1}^N \sum_j^J \beta_j \frac{\tau_{i,j}^m}{1+\tau_{i,j}^m} \frac{\partial \pi_{1i,j}}{\partial w_n} x_1 \\
&+ \sum_{i\neq 1} \sum_{m\neq 1}^N \sum_j^J \beta_j \gamma_i \frac{\partial \pi_{mi,j}}{\partial w_n} x_m + \sum_{i\neq 1} \gamma_i \sum_j^J \beta_j \pi_{ni,j} \frac{x_n}{w_n} + \sum_{i\neq 1} \sum_j^J \beta_j \gamma_i \frac{1}{1+\tau_{i,j}^m} \frac{\partial \pi_{1i,j}}{\partial w_n} x_1 - \gamma_n \frac{x_n}{w_n} = 0
\end{aligned}
$$

plugging in derivatives

$$
\begin{aligned}
&- u_{c_1} x_1 \sum_j^J \beta_j \pi_{1n,j} + \gamma_x \sum_{m\neq 1}^N \sum_j^J \beta_j \frac{\tau_{m,j}^x}{1+\tau_{m,j}^x} \theta \pi_{m1,j} \pi_{mn,j} x_m + \gamma_x \sum_j^J \beta_j \frac{\tau_{n,j}^x}{1+\tau_{n,j}^x} \pi_{n1,j} x_n + \gamma_x \sum_{i\neq\{1,n\}}^N \sum_j^J \beta_j \frac{\tau_{i,j}^m}{1+\tau_{i,j}^m} \theta \pi_{1i,j} \pi_{1n,j} x_1 \\
&- \gamma_x \sum_j^J \beta_j \frac{\tau_{n,j}^m}{1+\tau_{n,j}^m} \theta \pi_{1n,j} (1-\pi_{1n,j}) x_1 + \sum_{i\neq\{1,n\}} \sum_{m\neq 1}^N \sum_j^J \beta_j \gamma_i \theta \pi_{mi,j} \pi_{mn,j} x_m - \sum_{m\neq 1}^N \gamma_n \theta \pi_{mn,j} (1-\pi_{mn,j}) x_m + \sum_{i\neq 1} \sum_j^J \beta_j \gamma_i \pi_{ni,j} x_n \\
&+ \sum_{i\neq\{1,n\}} \sum_j^J \beta_j \gamma_i \frac{1}{1+\tau_{i,j}^m} \theta \pi_{1i,j} \pi_{1n,j} x_1 - \gamma_n \sum_j^J \beta_j \frac{1}{1+\tau_{n,j}^m} \theta \pi_{1n,j} (1-\pi_{1n,j}) x_1 - \gamma_n x_n = 0
\end{aligned}
$$

use $\gamma_x = u_{c_1}$, finally the FOCs on w_n satisfy

$$
\begin{aligned}
&- u_{c_1} x_1 \sum_j^J \beta_j \pi_{1n,j} + u_{c_1} \sum_{m\neq 1}^N \sum_j^J \beta_j \frac{\tau_{m,j}^x}{1+\tau_{m,j}^x} \theta \pi_{m1,j} \pi_{mn,j} x_m + u_{c_1} \sum_j^J \beta_j \frac{\tau_{n,j}^x}{1+\tau_{n,j}^x} \pi_{n1,j} x_n \qquad \text{(A.12)} \\
&+ u_{c_1} \sum_{i\neq 1}^N \sum_j^J \beta_j \frac{\tau_{i,j}^m}{1+\tau_{i,j}^m} \theta \pi_{1i,j} \pi_{1n,j} x_1 - u_{c_1} \sum_j^J \beta_j \frac{\tau_{n,j}^m}{1+\tau_{n,j}^m} \theta \pi_{1n,j} x_1 + \sum_{i\neq 1} \sum_{m\neq 1}^N \sum_j^J \beta_j \gamma_i \theta \pi_{mi,j} \pi_{mn,j} x_m \\
&- \sum_{m\neq 1}^N \gamma_n \theta \pi_{mn,j} x_m + \sum_{i\neq 1} \sum_j^J \beta_j \gamma_i \pi_{ni,j} x_n + \sum_{i\neq 1} \sum_j^J \beta_j \gamma_i \frac{1}{1+\tau_{i,j}^m} \theta \pi_{1i,j} \pi_{1n,j} x_1 - \gamma_n \sum_j^J \beta_j \frac{1}{1+\tau_{n,j}^m} \theta \pi_{1n,j} x_1 - \gamma_n x_n = 0
\end{aligned}
$$

Using equations (A.10) and (A.11), the above equations are system equations of γ_n. Q.E.D.

E A Simplified Two-Period Model

In this appendix, we consider a simplified two-period model with two countries and one sector. Home researcher L^r is picked once and left unchanged thereafter. There are two periods. In period 1, the technologies are exogenous and fixed at $(\bar{T}_1, \bar{T}_2)$ for country 1 (Home) and 2 (Foreign). Agents choose to be workers or researchers. Ideas formed by researchers are used for production in the second period. Hence, researchers only collect profit in period 2. In addition, period 1's technology $(\bar{T}_1, \bar{T}_2)$ has full depreciation. Thus, period 2's technologies only depend on the number of researchers in period 1, i.e.,

$$T_n = \alpha_n L_n^r \quad \text{for } n = 1, 2.$$

Lastly, all agents work as workers in period 2.

This simple model shares very similar policy patterns as our benchmark model. In particular, Ramsey's policies only use export taxes and do not distort domestic innovation incentives. In contrast, Markov's policy distorts domestic innovation incentives, and the Markov export tax in the future period (period 2 in this simple model) depends only on the static terms of trade effect.

Let variables with a prime denote the variables in period 2. We define the endogenous discount factors $\tilde{\rho}_n$ for $n = 1, 2$ as

$$\tilde{\rho}_n \equiv \rho \frac{u'_{c,n}/P'_n}{u_{c,n}/P_n} = \rho \frac{(x'_n)^{-\sigma}}{x_n^{-\sigma}} \frac{(P'_n)^{\sigma-1}}{P_n^{\sigma-1}},$$

where the second equality holds under CRRA preference, $u(c) = c^{1-\sigma}/(1-\sigma)$, and the expenditure is defined as $x_n = P_n c_n$.

Ramsey With commitment, the Ramsey government chooses all policies in period 1. The government chooses current and future export taxes $(\tau^x, \tau^{x\prime}, \tau^m, \tau^{m\prime})$ and home researchers (L_1^r), subject to the world market equilibrium. Specifically, the Ramsey government solves the following problem

$$\max_{\tau^x, \tau^{x\prime}, \tau^m, \tau^{m\prime}, L_1^r} \frac{(x_1/P_1)^{1-\sigma}}{1-\sigma} + \rho \frac{(x'_1/P'_1)^{1-\sigma}}{1-\sigma}$$

subject to the worker-researcher choice in country 2,

$$\frac{w_2}{\alpha_2} = \rho \frac{(x_2')^{-\sigma}}{x_2^{-\sigma}} \frac{(P_2')^{\sigma-1}}{P_2^{\sigma-1}} \left(\frac{1}{\theta} \frac{w_2' L_2}{\alpha_2 L_2^r} \right) \qquad (\gamma_{r2}), \tag{A.13}$$

the balanced trade conditions

$$\frac{1}{1+\tau^m} \pi_{12} x_1 = \pi_{21} x_2, \qquad (\gamma_2) \tag{A.14}$$

$$\frac{1}{1+\tau^{m\prime}} \pi_{12}' x_1' = \pi_{21}' x_2', \qquad (\gamma_2') \tag{A.15}$$

where the definitions of expenditures (x_1, x_2, x_1', x_2'), prices (P_1, P_2, P_1', P_2'), and trade shares are affected by Home government's policies and are given by

$$x_1 = \frac{1+\theta}{\theta} w_1 (L_1 - L_1^r) + \frac{\tau^x}{1+\tau^x} \pi_{21} x_2 + \frac{\tau^m}{1+\tau^m} \pi_{12} x_1, \tag{A.16}$$

$$x_2 = \frac{1+\theta}{\theta} w_2 (L_2 - L_2^r), \tag{A.17}$$

$$P_1 = \left[\bar{T}_1 (w_1)^{-\theta} + \bar{T}_2 (w_2 (1+\tau^m) d)^{-\theta} \right]^{-\frac{1}{\theta}}, \tag{A.18}$$

$$P_2 = \left[\bar{T}_1 (w_1 (1+\tau^x) d)^{-\theta} + \bar{T}_2 (w_2)^{-\theta} \right]^{-\frac{1}{\theta}}, \tag{A.19}$$

$$\pi_{12} = \frac{\bar{T}_2 (w_2 (1+\tau^m) d)^{-\theta}}{\bar{T}_1 (w_1)^{-\theta} + \bar{T}_2 (w_2 (1+\tau^m) d)^{-\theta}}, \tag{A.20}$$

$$\pi_{21} = \frac{\bar{T}_1 (w_1 (1+\tau^x) d)^{-\theta}}{\bar{T}_1 (w_1 (1+\tau^x) d)^{-\theta} + \bar{T}_2 (w_2)^{-\theta}}, \tag{A.21}$$

$$x_1' = \frac{1+\theta}{\theta} w_1' L_1 + \frac{\tau^{x\prime}}{1+\tau^{x\prime}} \pi_{21}' x_2' + \frac{\tau^{m\prime}}{1+\tau^{m\prime}} \pi_{12}' x_1', \tag{A.22}$$

$$x_2' = \frac{1+\theta}{\theta} w_2' L_2, \tag{A.23}$$

$$P_1' = \left[(\alpha_1 L_1^r)(w_1')^{-\theta} + (\alpha_2 L_2^r)(w_2' (1+\tau^{m\prime}) d)^{-\theta} \right]^{-\frac{1}{\theta}} = 1, \tag{A.24}$$

$$P_2' = \left[(\alpha_1 L_1^r)(w_1' (1+\tau_x') d)^{-\theta} + (\alpha_2 L_2^r)(w_2')^{-\theta} \right]^{-\frac{1}{\theta}}, \tag{A.25}$$

$$\pi_{12}' = \frac{(\alpha_2 L_2^r)(w_2' (1+\tau^{m\prime}) d)^{-\theta}}{(\alpha_1 L_1^r)(w_1')^{-\theta} + (\alpha_2 L_2^r)(w_2' (1+\tau^{m\prime}) d)^{-\theta}}, \tag{A.26}$$

$$\pi_{21}' = \frac{(\alpha_1 L_1^r)(w_1' (1+\tau^{x\prime}) d)^{-\theta}}{(\alpha_1 L_1^r)(w_1' (1+\tau^{x\prime}) d)^{-\theta} + (\alpha_2 L_2^r)(w_2')^{-\theta}}. \tag{A.27}$$

Assuming a unique solution exists and the first-order conditions are necessary and sufficient, we can characterize the Ramsey policies using first-order conditions (FOC). The Ramsey policies have the following properties.

1. Optimal import tariff and slack of balanced-trade conditions.

We first show that under the optimal import tariff and normalization of tariffs $\tau^m = \tau^{m\prime} = 0$, the balanced-trade constraints (A.14)-(A.15) are slack, i.e., $\gamma_2 = \gamma_2' = 0$.[13]

This implies that we can solve the problem in two steps. In the first step, we work on the formula for optimal policies assuming a partial equilibrium with exogenous wages. In the second step, we study the optimal wages in the general equilibrium and evaluate the equilibrium policies under these equilibrium wages. This two-step method is similar to that used in Costinot et al. (2015) and Lashkaripour and Lugovskyy (2023). This method is applicable to a two-country setup, as the model here, or a multi-country model with the assumption that the relative wages across any two other countries are fixed, as assumed in Lashkaripour and Lugovskyy (2023).

In the Ramsey problem, the first-order condition on τ^m implies

$$u_{c1}\frac{1}{P_1}\frac{\partial x_1}{\partial \tau^m} - u_{c1}\frac{x_1}{P_1^2}\frac{\partial P_1}{\partial \tau^m} + \gamma_2\left[\frac{1}{(1+\tau^m)^2}\pi_{12}x_1 - \frac{1}{1+\tau^m}\frac{\partial \pi_{12}}{\partial \tau^m}x_1\right] = 0. \tag{A.28}$$

Using the balanced trade condition (A.14) to replace $\frac{\tau^m}{1+\tau^m}\pi_{12}x_1$ with $\tau^m\pi_{21}x_2$ in home country's expenditure (A.16), we can rewrite Home expenditure as

$$x_1 = \frac{1+\theta}{\theta}w_1(L_1 - L_1^r) + \frac{\tau^x}{1+\tau^x}\pi_{21}x_2 + \tau^m\pi_{21}x_2.$$

It is easy to show the following derivatives hold,

$$\frac{\partial x_1}{\partial \tau^m} = \pi_{21}x_2, \quad \frac{\partial P_1}{\partial \tau^m} = P_1\frac{\pi_{12}}{1+\tau^m}, \quad \frac{\partial \pi_{12}}{\partial \tau^m} = \frac{\theta\pi_{12}\pi_{11}}{1+\tau^m}.$$

Plugging the above derivatives to FOC (A.28) and imposing $\tau^m = 0$ for normalization, we get

$$\frac{u_{c1}}{P_1}\pi_{21}x_2 - \frac{u_{c1}}{P_1}\pi_{12}x_1 + \gamma_2\left[\pi_{12}x_2 - \theta\pi_{12}\pi_{11}x_2\right] = 0.$$

Again using the balanced trade condition (A.14) and $\tau^m = 0$, the above equation becomes

$$\gamma_2(1-\theta\pi_{11})\pi_{12}x_2 = 0.$$

which holds when $\gamma_2 = 0$. Similarly, we can show that $\gamma_2' = 0$.

[13] Under the tax neutrality (Proposition 3), we can normalize one tax equal to zero, and we choose tariffs $\tau^m = \tau^{m\prime} = 0$.

Intuitively, the level of import tariff can be used to manipulate relative wages through balanced trade conditions. Under tax neutrality (or Lerner symmetry in the two-country cases), we can normalize import tariff to zero. Hence, relative wages are pinned down in this case. As a result, Home government can ignore the balanced trade conditions under zero tariffs.

Below, we will drop the balanced trade conditions (A.14-A.15) and impose $\tau^m = \tau^{m\prime} = 0$ when deriving the optimal export tax and innovation policies.

2. Optimal export tax.

FOC on τ^x reflects the static effect on current export tax income and dynamic effect on foreign interest rate and foreign innovation incentives:

$$\frac{u_{c1}}{P_1} \underbrace{\frac{\partial x_1}{\partial \tau^x}}_{\text{static effect}} \underbrace{- \gamma_{r2} \tilde{\rho}_2 \left(\frac{1}{\theta} \frac{w_2' L_2}{\alpha_2 L_2^r} \right) (\sigma - 1) \frac{1}{P_2} \frac{\partial P_2}{\partial \tau^x}}_{\text{dynamic effect}} = 0. \tag{A.29}$$

The static effect captures how export tax affects Home's current expenditure through tax revenues. The dynamic effect reflects the impact of Home export tax on foreign price level P_2 and, in turn, on foreign endogenous discount factor $\tilde{\rho}_2 = \rho \frac{(x_2')^{-\sigma}}{x_2^{-\sigma}} \frac{(P_2')^{\sigma-1}}{P_2^{\sigma-1}}$. Higher τ^x pushes up P_2 and increases foreign interest rate (when $\sigma > 1$), which lowers foreign innovation incentives. When $\gamma_{r2} > 0$, Home government values high T_2. In this case, high τ^x deters foreign innovation and reduces marginal benefit of export tax. Hence, the dynamic effect shows up in FOC (A.29) as a negative sign when $\gamma_{r2} > 0$ and $\sigma > 1$.

The static effect of higher τ^x can be further decomposed into a direct effect from higher tax revenue and an indirect effect of reducing foreign imports,

$$\frac{\partial x_1}{\partial \tau^x} = \underbrace{\frac{1}{(1+\tau^x)^2} \pi_{21} x_2}_{\text{direct effect on tax rev.}>0} + \underbrace{\frac{\tau^x}{1+\tau^x} \frac{\partial \pi_{21}}{\partial \tau^x} x_2}_{\text{indirect effect: reducing foreign import}<0}.$$

Note that the elasticity of foreign price P_2 over export tax is π_{21}:

$$\frac{\partial \ln P_2}{\partial \ln(1+\tau_x)} = \frac{1+\tau_x}{P_2} \frac{\partial P_2}{\partial \tau_x} = \pi_{21}.$$

Plugging in the derivatives, we obtain the optimal export tax under Ramsey:

$$1 + \tau_R^x = \frac{1 + \theta \pi_{22}}{\theta \pi_{22} + (\sigma - 1) \gamma_{r2} \tilde{\rho}_2 \left(\frac{w_2' L_2}{\theta T_2} \right) \frac{1}{x_2 u_{c1}}}. \tag{A.30}$$

Similarly, we can show that the optimal export tax on $\tau^{x\prime}$ also depends on the static and dynamic terms of trade effect:

$$\rho \frac{u'_{c1}}{P'_1} \frac{\partial x'_1}{\partial \tau^{x\prime}} + \gamma_{r2} \tilde{\rho}_2 \left(\frac{1}{\theta} \frac{w'_2 L_2}{T_2} \right) (\sigma - 1) \frac{1}{P'_2} \frac{\partial P'_2}{\partial \tau^{x\prime}} = 0.$$

Plugging in the derivatives and simplifying, we can show that the optimal Ramsey policy in period 2 is given by

$$1 + \tau_R^{x\prime} = \frac{1 + \theta \pi'_{22}}{\theta \pi'_{22} - (\sigma - 1) \gamma_{r2} \tilde{\rho}_2 \left(\frac{w'_2 L_2}{\theta T_2} \right) \frac{1}{\rho x'_2 u'_{c1}}}. \tag{A.31}$$

Comparing current optimal tax τ_R^x with future one $\tau_R^{x\prime}$, it is clear that both include a term, $(\sigma - 1) \gamma_{r2} \tilde{\rho}_2 \left(\frac{w'_2 L_2}{\theta T_2} \right)$, which considers the dynamic effect of export tax on foreign research effort through its impact on foreign interest rate. When $\sigma > 1$, higher τ^x increases P_2 and *increases* foreign interest rate. In contrast, higher $\tau^{x\prime}$ increase P'_2 and *lowers* foreign interest rate. Suppose Home government wants to decrease foreign research efforts. It can achieve this by increasing τ^x and decreasing $\tau^{x\prime}$, both pushing up foreign interest rates. When Home government likes a lower T_2 and thus $\gamma_{r2} < 0$, the dynamic term $(\sigma - 1) \gamma_{r2} \tilde{\rho}_2 \left(\frac{1}{\theta} \frac{w'_2 L_2}{\alpha_1 L_{2r}} \right) \frac{1}{x_2 u_{c1}}$ increases τ^x above the static incentive $1/(\theta \pi_{22})$, as shown in formula (A.30), but a similar term pushes $\tau^{x\prime}$ below the static incentive $1/(\theta \pi'_{22})$, as shown in formula (A.31).

3. Optimal domestic innovation.

   FOC over L_1^r considers the marginal cost of innovation, i.e. reducing current production and consumption $\frac{u_{c1}}{P_1} \frac{\partial x_1}{\partial L_1^r}$, and the marginal benefits of (1) more sales to country 2 and increase external revenue in the second period, $\rho \frac{u'_{c1}}{P'_1} \frac{\partial x'_1}{\partial L_1^r}$, (2) lower future home price P'_1, $-\rho \frac{u'_{c1}}{P'_1} \frac{x'_1}{P'_1} \frac{\partial P'_1}{\partial L_{1r}} \geq 0$, (3) higher home technology level thus lower P'_2, which lowers foreign innovation incentives through competition, $\gamma_{r2}(\sigma - 1) \tilde{\rho}_2 \left(\frac{w'_2 L_2}{\theta T_2} \right) \frac{1}{P'_2} \frac{\partial P'_2}{\partial L_1^r}$, if $\sigma > 1$.

   These incentives are captured by the following FOC

$$\frac{u_{c1}}{P_1} \frac{\partial x_1}{\partial L_1^r} + \rho \frac{u'_{c1}}{P'_1} \frac{\partial x'_1}{\partial L_1^r} - \rho \frac{u'_{c1}}{P'_1} \frac{x'_1}{P'_1} \frac{\partial P'_1}{\partial L_1^r} + \gamma_{r2} \tilde{\rho}_2 \left(\frac{w'_2 L_2}{\theta T_2} \right) (\sigma - 1) \frac{1}{P'_2} \frac{\partial P'_2}{\partial L_1^r} = 0.$$

   Plugging in derivatives and using FOC on τ^x, we have

$$\frac{w_1}{\alpha_1} = \tilde{\rho}_1 \left(\frac{w'_1 L_1}{\theta T_1} \right),$$

   which is the same as the one in the private equilibrium. Hence Ramsey will not distort

domestic innovation.

Markov The Markov government cannot commit to future policies. We solve this problem backwardly. In the last period (period 2), the technology is given by (T_1, T_2) achieved through the researchers' choices in period 1. All agents work as workers. In this last period, the government only chooses $\tau^{x\prime}$ to maximize agents' utility,

$$V_2(T_1, T_2) = \max_{\tau^{x\prime}} \frac{(x_1'/P_1')^{1-\sigma}}{1-\sigma}$$

where the expenditures, prices, and trade shares are given by equation (A.22)-(A.27) with $\tau^{m\prime} = 0$.

FOC over $\tau^{x\prime}$ only involves the static effect:

$$\frac{u_{c1}'}{P_1'}\frac{\partial x_1'}{\partial \tau^{x\prime}} = \frac{u_{c1}'}{P_1'}\left[\frac{1}{(1+\tau^{x\prime})^2}\pi_{21}' x_2' + \frac{\tau^{x\prime}}{1+\tau^{x\prime}}\frac{\partial \pi_{21}'}{\partial \tau^{x\prime}} x_2'\right] = 0$$

After simplification, we get the following condition,

$$1 + \tau_M^{x\prime} = \frac{1+\theta\pi_{22}'}{\theta\pi_{22}'}. \tag{A.32}$$

Comparing to the Ramsey policy $\tau_R^{x\prime}$ in the second period, as shown in (A.31), Markov government does not consider the impact of τ_x' on foreign innovation decisions in the first period. Hence, there is no dynamic effect in $\tau_M^{x\prime}$.

We now consider Markov problem in the first period,

$$\max_{\tau^x, L_1^r} \frac{(x_1/P_1)^{1-\sigma}}{1-\sigma} + \rho V_2(\alpha_1 L_1^r, \alpha_2 L_2^r)$$

subject to the private equilibrium (A.16)-(A.21), and foreign worker-researcher choice,

$$\frac{w_2}{\alpha_2} = \rho G\left(\alpha_1 L_1^r, \alpha_2 L_2^r\right) x_2^{\sigma} P_2^{1-\sigma}\left(\frac{1}{\theta}\frac{L_2}{\alpha_2 L_2^r}\right) \qquad (\gamma_{r2})$$

where G function reflects future marginal consumption on extra wages, given by $G(T_1, T_2) = [x_2'(T_1, T_2)]^{-\sigma}[P_2'(T_1, T_2)]^{\sigma-1} w_2'(T_1, T_2)$.

FOC over τ^x

$$\frac{u_{c1}}{P_1}\underbrace{\frac{\partial x_1}{\partial \tau^x}}_{\text{static effect}} \underbrace{-(\sigma-1)\gamma_{r2}\tilde{\rho}_2\left(\frac{w_2' L_2}{\theta T_2}\right)\frac{1}{P_2}\frac{\partial P_2}{\partial \tau^x}}_{\text{dynamic effect}} = 0,$$

which leads to the same formula as in Ramsey,

$$1+\tau_M^x = \frac{1+\theta\pi_{22}}{\theta\pi_{22}+(\sigma-1)\gamma_{r2}\tilde{\rho}_2\left(\frac{w_2' L_2}{\theta T_2}\right)\frac{1}{x_2 u_{c1}}}. \tag{A.33}$$

FOC over L_1^r

$$-u_{c1}\frac{1+\theta}{\theta}w_1 + \rho\frac{\partial V_2}{\partial T_1}\alpha_1 + \gamma_{r2}\rho\frac{\partial G}{\partial T_1}\alpha_1 x_2^{\sigma} P_2^{1-\sigma}\left(\frac{1}{\theta}\frac{L_2}{\alpha_2 L_2^r}\right) = 0.$$

Note

$$\frac{\partial V_2}{\partial T_1} = u_{c1}'\frac{\partial x_1'}{\partial T_1} + \frac{1}{\theta}u_{c1}' x_1' \pi_{11}'\frac{1}{T_1}$$

and so

$$-u_{c1}\frac{1+\theta}{\theta}w_1 L_1^r + \rho u_{c1}'\frac{\tau_x'}{1+\tau_x'}\pi_{21}'\pi_{22}' x_2' + \frac{1}{\theta}\rho u_{c1}' x_1'\pi_{11}' + \gamma_{r2}\rho\frac{\partial G}{\partial T_1}T_1 x_2^{\sigma}P_2^{1-\sigma}\left(\frac{1}{\theta}\frac{L_2}{\alpha_2 L_2^r}\right) = 0.$$

After simplification, we obtain

$$\frac{w_1}{\alpha_1} = \tilde{\rho}_1\left[\frac{1}{\theta}\frac{w_1' L_1}{\theta T_1} + \frac{\theta}{1+\theta}\gamma_{r2}\tilde{\rho}_2\frac{\partial G}{\partial T_1}\frac{T_1}{G}\left(\frac{w_2' L_2}{\theta T_2}\right)\frac{1}{\rho u_{c1}}\right].$$

Hence, there is an extra term in the optimal decision of L_1^r in the Markov, compared to the Ramsey and private equilibrium. Mechanically, this term is considered by FOC of $\tau^{x\prime}$ in Ramsey.

Summary of Ramsey and Markov policies The Ramsey policies are given by the export taxes

$$1+\tau_R^x = \frac{1+\theta\pi_{22}}{\theta\pi_{22}+(\sigma-1)\gamma_{r2}\tilde{\rho}_2\left(\frac{w_2' L_2}{\theta T_2}\right)\frac{1}{x_2 u_{c1}}}, \quad 1+\tau_R^{x\prime} = \frac{1+\theta\pi_{22}'}{\theta\pi_{22}'-(\sigma-1)\gamma_{r2}\tilde{\rho}_2\left(\frac{w_2' L_2}{\theta T_2}\right)\frac{1}{\rho x_2' u_{c1}'}}, \tag{A.34}$$

and the optimal innovation in period 1,

$$\frac{w_1}{\alpha_1} = \tilde{\rho}_1\left(\frac{w_1' L_1}{\theta T_1}\right). \tag{A.35}$$

In contrast, the Markov policies have the following features:

$$1+\tau_M^x=\frac{1+\theta\pi_{22}}{\theta\pi_{22}+(\sigma-1)\gamma_{r2}\tilde{\rho}_2\left(\frac{w_2'L_2}{\theta T_2}\right)\frac{1}{x_2u_{c1}}},\quad 1+\tau_M^{x\prime}=\frac{1+\theta\pi_{22}'}{\theta\pi_{22}'},\tag{A.36}$$

$$\frac{w_1}{\alpha_1}=\tilde{\rho}_1\left[\frac{w_1'L_1}{\theta T_1}+\frac{\theta}{1+\theta}\gamma_{r2}\widetilde{\rho}_2\frac{\partial G}{\partial T_1}\frac{T_1}{G}\left(\frac{w_2'L_2}{\theta T_2}\right)\frac{1}{\rho u_{c1}'}\right].\tag{A.37}$$

To sum up, the simplified example has similar results as our full model. Namely, to affect foreign's forward-looking innovation, Ramsey uses a path of trade policies and doesn't need to distort Home's own innovation. In contrast, Markov has to use both trade policies and innovation policies, which distorts Home's own T to affect foreign future return and thus innovation incentives.

F Characterization of Ramsey Policy in the Baseline

This appendix first lays out the Home government's Ramsey problem. We then characterize the Ramsey policy by taking first-order conditions. In F.3, we prove the steady-state properties.

F.1 Ramsey problem

At period 0, Home planner chooses a sequence of $\left\{L_{1,j,t}^r,\tau_{n,j,t}^x,\tau_{n,j,t}^m\right\}_{t=0}^{\infty}$ to maximize the present value of utility subject to the implementability constraints (private market equilibrium every period):

$$V=\max\sum_{t=0}^{\infty}\rho^t\frac{(x_{1t}/P_{1t})^{1-\sigma}}{1-\sigma},$$

subject to, for each $t \geq 0$

$$T_{n,j,t} = \alpha_{n,j} L^r_{n,j,t} + (1-\delta) T_{n,j,t-1} \qquad (\gamma^T_{n,j,t})$$

$$\frac{w_{nt}}{\alpha_{n,j}} x_{nt}^{-\sigma} P_{nt}^{\sigma-1} = \sum_{s=t}^{\infty} [\rho(1-\delta)]^{s-t} x_{ns}^{-\sigma} P_{ns}^{\sigma-1} \frac{1}{\theta} \frac{w_{ns} L^p_{n,j,s}}{T_{n,j,s}} \qquad (\rho^t \gamma^v_{n,j,t})$$

$$\frac{1+\theta}{\theta} w_{1t} L^p_{1,j,t} = \beta_j \left[\pi_{11,j,t} x_{1t} + \sum_{m \neq 1} \frac{1}{1+\tau^x_{m,j,t}} \pi_{m1,j,t} x_{mt} \right] \qquad (\gamma^L_{1,j,t})$$

$$\frac{1+\theta}{\theta} w_{nt} L^p_{n,j,t} = \beta_j \left[\frac{1}{1+\tau^m_{n,j,t}} \pi_{1n,j,t} x_{1t} + \sum_{m \neq 1} \pi_{mn,j,t} x_{mt} \right] \qquad (\gamma^L_{n,j,t})$$

$$\sum_j \left(L^r_{n,j,t} + L^p_{n,j,t} \right) = L_n \qquad (\mu_{nt})$$

$$x_{1t} = \frac{1+\theta}{\theta} w_{1t} \sum_j L^p_{1,j,t} + \sum_{m \neq 1}^{N} \sum_{j=1}^{J} \beta_j \frac{\tau^x_{m,j,t}}{1+\tau^x_{m,j,t}} \pi_{m1,j,t} x_{mt} + \sum_{i \neq 1}^{N} \sum_{j=1}^{J} \beta_j \frac{\tau^m_{i,j,t}}{1+\tau^m_{i,j,t}} \pi_{1i,j,t} x_{1t} \qquad (\gamma_{xt})$$

where

$$x_{mt} = \frac{1+\theta}{\theta} w_{mt} \sum_j L^p_{m,j,t}$$

$$P_{mt} = \Pi_j \left[T_{1,j,t} (w_{1t}(1+\tau^x_{m,j,t}) d_{m1})^{-\theta} + \sum_{n \neq 1} T_{n,j,t} (w_{nt} d_{mn})^{-\theta} \right]^{-\frac{\beta_j}{\theta}}$$

$$P_{1t} = \Pi_j \left[T_{1,j,t} (w_{1t})^{-\theta} + \sum_{n \neq 1} T_{n,j,t} (w_{nt}(1+\tau^m_{n,j,t}) d_{1n})^{-\theta} \right]^{-\frac{\beta_j}{\theta}} = 1$$

The government decides the entire path of policies that will be honored in the future, as in particular, future policies would affect the foreign individual's expected value of innovation.

The Lagrangian

$$
\begin{aligned}
L = & \sum_{t=0}^{\infty} \rho^t \frac{(x_{1t}/P_{1t})^{1-\sigma}}{1-\sigma} + \sum_{t=0}^{\infty} \sum_{n\neq 1} \sum_{j}^{J} \rho^t \gamma^v_{n,j,t} \left[\sum_{s=t}^{\infty} [\rho(1-\delta)]^{s-t} \frac{1}{\theta} \frac{w_{ns} L^p_{n,j,s}}{T_{n,j,s}} x_{ns}^{-\sigma} P_{ns}^{\sigma-1} - \frac{w_{nt}}{\alpha_{n,j}} x_{nt}^{-\sigma} P_{nt}^{\sigma-1} \right] \\
& + \sum_{t=0}^{\infty} \sum_{n} \sum_{j}^{J} \rho^t \gamma^T_{n,j,t} (\alpha_{n,j} L^r_{n,j,t} + (1-\delta) T_{n,j,t-1} - T_{n,j,t}) \\
& + \sum_{t=0}^{\infty} \sum_{j}^{J} \rho^t \gamma^L_{1,j,t} [\beta_j (\pi_{11,j,t} x_{1t} + \sum_{m\neq 1} \frac{1}{1+\tau^x_{m,j,t}} \pi_{m1,j,t} x_{mt}) - \frac{1+\theta}{\theta} w_{1t} L^p_{1,j,t}] \\
& + \sum_{t=0}^{\infty} \sum_{n\neq 1} \sum_{j}^{J} \rho^t \gamma^L_{n,j,t} [\beta_j (\frac{1}{1+\tau^m_{n,j,t}} \pi_{1n,j,t} x_{1t} + \sum_{m\neq 1} \pi_{mn,j,t} x_{mt}) - \frac{1+\theta}{\theta} w_{nt} L^p_{n,j,t}] \\
& + \sum_{t=0}^{\infty} \sum_{n} \rho^t \mu_{nt} [1 - \sum_{j} \left(L^r_{n,j,t} + L^p_{n,j,t} \right)] \\
& + \sum_{t=0}^{\infty} \rho^t \gamma_{xt} (\frac{1+\theta}{\theta} w_{1t} \sum_{j} L^p_{1,j,t} + \sum_{m\neq 1}^{N} \sum_{j=1}^{J} \beta_j \frac{\tau^x_{m,j,t}}{1+\tau^x_{m,j,t}} \pi_{m1,j,t} x_{mt} + \sum_{i\neq 1}^{N} \sum_{j=1}^{J} \beta_j \frac{\tau^m_{i,j,t}}{1+\tau^m_{i,j,t}} \pi_{1i,j,t} x_{1t} - x_{1t})
\end{aligned}
$$

can be rewritten as

$$
L = \sum_{t=0}^{\infty} \rho^t \left\{ \frac{(x_{1t}/P_{1t})^{1-\sigma}}{1-\sigma} - \sum_{n\neq 1} \sum_{j}^{J} \gamma^v_{n,j,t} \frac{w_{nt}}{\alpha_{n,j}} x_{nt}^{-\sigma} P_{nt}^{\sigma-1} + \sum_{n\neq 1} \sum_{j}^{J} \left(\sum_{s=0}^{t} \gamma^v_{n,j,s} (1-\delta)^{t-s} \right) \frac{1}{\theta} \frac{w_{nt} L^p_{n,j,t}}{T_{n,j,t}} x_{nt}^{-\sigma} P_{nt}^{\sigma-1} + ... \right\}
$$

where we omit those constraints that are the same as in the Markov problem.

Let

$$
\Gamma^v_{n,j,t} = \sum_{s=0}^{t} (1-\delta)^{t-s} \gamma^v_{n,j,s}, \qquad \text{for any } n \neq 1 \text{ and } j
$$

and it can be written recursively as

$$
\Gamma^v_{n,j,t} = (1-\delta) \Gamma^v_{n,j,t-1} + \gamma^v_{n,j,t}
$$

Then,

$$
L = \sum_{t=0}^{\infty} \rho^t \left\{ \frac{(x_{1t}/P_{1t})^{1-\sigma}}{1-\sigma} - \sum_{n\neq 1} \sum_{j}^{J} \gamma^v_{n,j,t} \frac{w_{nt}}{\alpha_{n,j}} x_{nt}^{-\sigma} P_{nt}^{\sigma-1} + \sum_{n\neq 1} \sum_{j}^{J} \left((1-\delta) \Gamma^v_{n,j,t-1} + \gamma^v_{n,j,t} \right) \frac{1}{\theta} \frac{w_{nt} L^p_{n,j,t}}{T_{n,j,t}} x_{nt}^{-\sigma} P_{nt}^{\sigma-1} + ... \right\}
$$

Hence the Ramsey problem can be written as recursively,

$$
\begin{aligned}
L\left(\left\{ \Gamma^v_{n,j,-1}, T_{n,j,-1} \right\} \right) = & \inf_{\gamma^v_{n,j}} \sup_{x_1,..} \frac{(x_{1t}/P_{1t})^{1-\sigma}}{1-\sigma} + \sum_{n\neq 1} \sum_{j}^{J} \left[\frac{1}{\theta} \frac{L^p_{n,j}}{T_{n,j}} \left(\gamma^v_{n,j} + (1-\delta) \Gamma^v_{n,j,-1} \right) - \frac{1}{\alpha_{n,j}} \gamma^v_{n,j} \right] w_n x_n^{-\sigma} P_n^{\sigma-1} \\
& + ... + \rho L \left(\left\{ \Gamma^v_{n,j}, T_{n,j} \right\} \right)
\end{aligned}
$$

where

$$\Gamma^v_{n,j} = (1-\delta)\Gamma^v_{n,j,-1} + \gamma^v_{n,j}.$$

F.2 Optimal conditions for Ramsey policy

FOC over x_1: (with $P_1 = 1$)

$$u_{c_1} + \sum_{n=1}^{N}\sum_{j}\gamma^L_{n,j}\beta_j\frac{1}{1+\tau^m_{n,j}}\pi_{1n,j} - \gamma_x\left(1 - \sum_{n\neq 1}^{N}\sum_{j=1}^{J}\beta_j\frac{\tau^m_{n,j}}{1+\tau^m_{n,j}}\pi_{1n,j}\right) = 0. \tag{A.38}$$

FOC over $L^r_{1,j}$:

$$\mu_1 = \gamma^T_{1,j}\alpha_{1,j}, \tag{A.39}$$

FOC over $L^p_{1,j}$:

$$\mu_1 = \left(\gamma_x - \gamma^L_{1,j}\right)\frac{1+\theta}{\theta}w_1. \tag{A.40}$$

FOC over $L^r_{n,j}$ for $n > 1$:

$$\mu_n = \gamma^T_{n,j}\alpha_{n,j}, \tag{A.41}$$

FOC over $L^p_{n,j}$

$$\begin{aligned}
&\frac{1}{\theta}\frac{w_n}{T_{n,j}}\Gamma^v_{n,j}x_n^{-\sigma}P_n^{\sigma-1} - \gamma^L_{n,j}\frac{1+\theta}{\theta}w_n - \mu_n + \sum_{k=1}^{N_s}\gamma^L_{1,k}\beta_k\frac{1}{1+\tau^x_{n,k}}\pi_{n1,k}\frac{1+\theta}{\theta}w_n \\
&+ \sum_{i\neq 1}^{N}\sum_{k}^{N_s}\gamma^L_{i,k}\beta_k\pi_{ni,k}\frac{1+\theta}{\theta}w_n + \gamma_x\sum_{k}^{N_s}\beta_k\frac{\tau^x_{n,k}}{1+\tau^x_{n,k}}\pi_{n1,k}\frac{1+\theta}{\theta}w_n \\
&+ \sum_{k}^{N_s}\left[\frac{1}{\theta}\frac{L^p_{nk}}{T_{n,k}}\left(\gamma^v_{n,k} + (1-\delta)\Gamma^v_{n,k,-1}\right) - \frac{1}{\alpha_{n,k}}\gamma^v_{n,k}\right]w_n x_n^{-\sigma}P_n^{\sigma-1}(-\sigma)\frac{1+\theta}{\theta}\frac{w_n}{x_n} = 0
\end{aligned} \tag{A.42}$$

FOC over $\tau^m_{n,j}$ for $n > 1$:

$$\begin{aligned}
&-u_{c_1}x_1\beta_j\frac{\pi_{1n,j}}{1+\tau^m_{n,j}} + \gamma^L_{1,j}\beta_j\frac{\partial\pi_{11,j}}{\partial\tau^m_{n,j}}x_1 - \gamma^L_{n,j}\beta_j\frac{\pi_{1n,j}}{\left(1+\tau^m_{n,j}\right)^2}x_1 + \gamma^L_{n,j}\frac{1}{1+\tau^m_{n,j}}\frac{\partial\pi_{1n,j}}{\partial\tau^m_{n,j}}\beta_j x_1 \\
&+ \sum_{i\neq\{1,n\}}\gamma^L_{i,j}\frac{1}{1+\tau^m_{i,j}}\frac{\partial\pi_{1i,j}}{\partial\tau^m_{n,j}}\beta_j x_1 + \gamma_x\beta_j\frac{1}{\left(1+\tau^m_{n,j}\right)^2}\pi_{1n,j}x_1 + \gamma_x\sum_{i\neq 1}^{N}\beta_j\frac{\tau^m_{i,j}}{1+\tau^m_{i,j}}\frac{\partial\pi_{1i,j}}{\partial\tau^m_{n,j}}x_1 = 0
\end{aligned}$$

plug in the derivatives

$$\frac{1}{1+\tau^m_{n,j}} = \frac{u_{c_1} - \sum_{i\neq\{1,n\}}^{N}\left(\gamma_x - \gamma^L_{i,j}\right)\frac{\tau^m_{i,j}}{1+\tau^m_{i,j}}\theta\pi_{1i,j} + \gamma_x\theta(1-\pi_{1n,j}) - \gamma^L_1\theta\pi_{11,j} - \sum_{i\neq\{1,n\}}\gamma^L_{i,j}\theta\pi_{1i,j}}{\left(\gamma_x - \gamma^L_{n,j}\right)\left[1+\theta(1-\pi_{1n,j})\right]}$$

or equivalently,

$$1+\tau^m_{n,j} = \frac{\left(\gamma_x - \gamma^L_{n,j}\right)\left[1+\theta(1-\pi_{1n,j})\right]}{\left(\gamma_x - \gamma^L_1\right)\theta\pi_{11,j} + u_{c_1} + \sum_{i\neq\{1,n\}}^{N}\left(\gamma_x - \gamma^L_{i,j}\right)\frac{1}{1+\tau^m_{i,j}}\theta\pi_{1i,j}}$$

combining with the FOC for x_1 (Eq.A.38), we arrive at $u_{c_1} = \gamma_x - \gamma^L_1$, and for each foreign country $n > 1$ and $j \in [1, J]$

$$1+\tau^m_{n,j} = \frac{\gamma_x - \gamma^L_{n,j}}{u_{c_1}}. \tag{A.43}$$

FOC over $\tau^x_{n,j}$ for $n > 1$

$$\begin{aligned}
&\sum_k^J\left[\gamma^v_{n,k}\left(\frac{1}{\alpha_{n,k}} - \frac{1}{\theta}\frac{L^p_{n,k}}{T_{n,k}}\right) - (1-\delta)\Gamma^v_{n,k,-1}\frac{1}{\theta}\frac{L^p_{n,k}}{T_{n,k}}\right]w_n x_n^{-\sigma}P_n^{\sigma-1}(1-\sigma)\beta_j\frac{\pi_{n1,j}}{1+\tau^x_{n,j}}\\
&+\gamma^L_{1,j}\beta_j\frac{\partial\pi_{11,j}}{\partial\tau^x_{n,j}}x_1 - \gamma^L_{1,j}\beta_j\frac{1}{(1+\tau^x_{n,j})^2}\pi_{n1,j}x_n + \gamma^L_{1,j}\beta_j\frac{1}{1+\tau^x_{n,j}}\frac{\partial\pi_{n1,j}}{\partial\tau^x_{n,j}}x_n + \sum_{i\neq1}^{N}\gamma^L_{i,j}\beta_j\left[\sum_{m\neq1}\frac{\partial\pi_{mi,j}}{\partial\tau^x_{n,j}}x_m\right]\\
&+\gamma_x\beta_j\frac{1}{(1+\tau^x_{n,j})^2}\pi_{n1,j}x_n + \gamma_x\sum_{m\neq1}^{N}\sum_{j=1}^{J}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\frac{\partial\pi_{m1,j}}{\partial\tau^x_{n,j}}x_m = 0
\end{aligned}$$

plugging into derivatives and organizing

$$\begin{aligned}
&\sum_k^J\left[\gamma^v_{n,k}\left(\frac{1}{\alpha_{n,k}} - \frac{1}{\theta}\frac{L^p_{n,k}}{T_{n,k}}\right) - (1-\delta)\Gamma^v_{n,k,-1}\frac{1}{\theta}\frac{L^p_{n,k}}{T_{n,k}}\right]w_n x_n^{-\sigma}P_n^{\sigma-1}(1-\sigma)\beta_j\pi_{n1,j} - \gamma^L_{1,j}\frac{1}{(1+\tau^x_{n,j})}\beta_j\pi_{n1,j}x_n \qquad \text{(A.44)}\\
&-\gamma^L_{1,j}\frac{1}{1+\tau^x_{n,j}}\theta(1-\pi_{n1,j})\beta_j\pi_{n1,j}x_n + \sum_{i\neq1}^{N}\gamma^L_{i,j}\beta_j\theta\pi_{ni,j}\pi_{n1,j}x_n + \gamma_x\beta_j\frac{1}{1+\tau^x_{n,j}}\pi_{n1,j}x_n - \gamma_x\beta_j\frac{\tau^x_{n,j}}{1+\tau^x_{n,j}}\theta\pi_{n1,j}(1-\pi_{n1,j})x_n = 0
\end{aligned}$$

delete $\beta_j\pi_{n1,j}x_n$ on both sides

$$\begin{aligned}
&\sum_k^J\left[\gamma^v_{n,k}\left(\frac{1}{\alpha_{n,k}} - \frac{1}{\theta}\frac{L^p_{n,k}}{T_{n,k}}\right) - (1-\delta)\Gamma^v_{n,k,-1}\frac{1}{\theta}\frac{L^p_{n,k}}{T_{n,k}}\right]w_n x_n^{-\sigma}P_n^{\sigma-1}(1-\sigma)x_n^{-1} - \gamma^L_{1,j}\frac{1}{(1+\tau^x_{n,j})}\\
&-\gamma^L_{1,j}\frac{1}{(1+\tau^x_{n,j})}\theta(1-\pi_{n1,j}) + \sum_{i\neq1}^{N}\gamma^L_{i,j}\theta\pi_{ni,j} + \gamma_x\frac{1}{(1+\tau^x_{n,j})} - \gamma_x\theta(1-\pi^j_{n1}) + \gamma_x\frac{1}{1+\tau^x_{n,j}}\theta(1-\pi^j_{n1}) = 0
\end{aligned}$$

further combine items, we get

$$\frac{\gamma_x - \gamma_{1,j}^L}{1+\tau_{n,j}^x}\left[1+\theta(1-\pi_{n1,j})\right] = \gamma_x\theta(1-\pi_{n1,j}) - \theta\sum_{i\neq 1}^{N}\gamma_{i,j}^L\pi_{ni,j} - \sum_{k}^{J}\left[\Gamma_{n,k}^v\frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}} - \gamma_{n,k}^v\frac{1}{\alpha_{n,k}}\right]w_n x_n^{-\sigma}P_n^{\sigma-1}(\sigma-1)/x_n$$

hence, the optimal export tax

$$1+\tau_{n,j}^x = \frac{(\gamma_x-\gamma_{L1})\left[1+\theta(1-\pi_{n1,j})\right]}{\gamma_x\theta(1-\pi_{n1,j}) - \theta\left(\sum_{i\neq 1}^{N}\gamma_{i,j}^L\pi_{ni,j}\right) + \sum_{k}^{J}\left(\gamma_{n,k}^v\frac{1}{\alpha_{n,k}} - \Gamma_{n,k}^v\frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}}\right)w_n x_n^{-\sigma}P_n^{\sigma-1}(\sigma-1)/x_n}$$

or equivalently,

$$1+\tau_{n,j}^x = \frac{1+\theta(1-\pi_{n1,j})}{\theta\sum_{i\neq 1}(1+\tau_{i,j}^m)\pi_{ni,j} + \sum_{k}^{J}\left(\gamma_{n,k}^v\frac{1}{\alpha_{n,k}} - \Gamma_{n,k}^v\frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}}\right)w_n x_n^{-\sigma}P_n^{\sigma-1}(\sigma-1)/(u_{c_1}x_n)} \tag{A.45}$$

using the result for optimal tariff (Eq.A.43).

FOC over $T_{1,j}$

$$\begin{aligned}
&-\gamma_{1,j}^T + \rho\frac{\partial V}{\partial T_{1,j}} + u_{c_1}x_1\frac{\beta_j}{\theta}\frac{\pi_{11,j}}{T_{1,j}} + \sum_{n\neq 1}^{N}\sum_{k}^{J}\left[\gamma_{n,k}^v\left(\frac{1}{\alpha_{n,k}} - \frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}}\right) - (1-\delta)\Gamma_{n,k,-1}^v\frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}}\right]w_n x_n^{-\sigma}P_n^{\sigma-1}(\sigma-1)\left(\frac{\beta_j}{\theta}\frac{\pi_{n1,j}}{T_{1,j}}\right)\\
&+\gamma_{1,j}^L\beta_j\left[\frac{\partial\pi_{11,j}}{\partial T_{1,j}}x_1 + \sum_{m\neq 1}^{N}\frac{1}{1+\tau_{m,j}^x}\frac{\partial\pi_{m1,j}}{\partial T_{1,j}}x_m\right] + \sum_{n\neq 1}^{N}\gamma_{n,j}^L\beta_j\left[\frac{1}{1+\tau_{n,j}^m}\frac{\partial\pi_{1n,j}}{\partial T_{1,j}}x_1 + \sum_{m\neq 1}\frac{\partial\pi_{mn,j}}{\partial T_{1,j}}x_m\right]\\
&+\gamma_x\sum_{m\neq 1}^{N}\beta_j\frac{\tau_{m,j}^x}{1+\tau_{m,j}^x}\frac{\partial\pi_{m1,j}}{\partial T_{1,j}}x_m + \gamma_x\sum_{i\neq 1}^{N}\beta_j\frac{\tau_{i,j}^m}{1+\tau_{i,j}^m}\frac{\partial\pi_{1i,j}}{\partial T_{1,j}}x_1 = 0
\end{aligned}$$

Putting in envelop $\frac{\partial V}{\partial T_{1,j}} = \gamma_{1,j}'^T(1-\delta)$ and derivatives,

$$\begin{aligned}
\gamma_{1,j}^T &= \rho\gamma_{1,j}'^T(1-\delta) + \frac{1}{\theta}\frac{1}{T_{1,j}}\{u_{c_1}x_1\beta_j\pi_{11,j}\\
&+\sum_{n\neq 1}^{N}\sum_{k}^{J}\left[\gamma_{n,k}^v\left(\frac{1}{\alpha_{n,k}} - \frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}}\right) - (1-\delta)\Gamma_{n,k,-1}^v\frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}}\right]w_n x_n^{-\sigma}P_n^{\sigma-1}(\sigma-1)\left(\beta_j\pi_{n1,j}\right)\\
&+\gamma_{1,j}^L\theta\beta_j\left[\pi_{11,j}(1-\pi_{11,j})x_1 + \sum_{m\neq 1}^{N}\frac{1}{1+\tau_{m,j}^x}\pi_{m1,j}(1-\pi_{m1,j})x_m\right] - \sum_{n\neq 1}^{N}\gamma_{n,j}^L\theta\beta_j\left[\frac{1}{1+\tau_{n,j}^m}\pi_{11,j}\pi_{1n,j}x_1 + \sum_{m\neq 1}^{N}\pi_{m1,j}\pi_{mn,j}x_m\right]\\
&+\theta\gamma_x\sum_{n\neq 1}^{N}\beta_j\frac{\tau_{n,j}^x}{1+\tau_{n,j}^x}\pi_{n1,j}(1-\pi_{n1,j})x_n - \theta\gamma_x\sum_{n\neq 1}^{N}\beta_j\frac{\tau_{n,j}^m}{1+\tau_{n,j}^m}\pi_{11,j}\pi_{1n,j}x_1\}
\end{aligned} \tag{A.46}$$

To get the innovation policy, we combine the FOCs for $L_{1,j}^r$ (A.39), $L_{1,j}^p$ (A.40) and trade policies(A.43) and (A.45), with $T_{1,j}$ (A.46). Specifically, first, summing FOC for $\tau_{n,j}^x$ (Eq.A.44) over $n \neq 1$,

plug back into the FOC for $T_{1,j}$ (Eq.A.46), we get

$$\gamma_{1,j}^{T} = \rho E\left[\gamma_{1,j}^{T\prime}(1-\delta)\right] + \frac{1}{\theta}\frac{1}{T_{1,j}}\{u_{c_1}x_1\beta_j\pi_{11,j} + \gamma_1^L\theta\beta_j\left[\pi_{11,j}(1-\pi_{11,j})x_1\right] + \sum_{n\neq 1}\left(\gamma_x - \gamma_1^L\right)\frac{1}{(1+\tau_{n,j}^x)}\beta_j\pi_{n1,j}x_n$$
$$-\theta\pi_{11,j}x_1\sum_{n\neq 1}^{N}\frac{1}{1+\tau_{n,j}^m}\gamma_{n,j}^L\beta_j\pi_{1n,j} - \theta\gamma_x\sum_{n\neq 1}^{N}\beta_j\frac{\tau_{n,j}^m}{1+\tau_{n,j}^m}\pi_{11,j}\pi_{1m,j}x_1\} \tag{A.47}$$

As can be seen, the term $\sum_{n\neq 1}^{N}\sum_{k}^{J}\left[\gamma_{n,k}^v\left(\frac{1}{\alpha_{n,k}} - \frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}}\right) - (1-\delta)\Gamma_{n,k,-1}^v\frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}}\right]w_n x_n^{-\sigma}P_n^{\sigma-1}(\sigma-1)\left(\frac{\beta_j}{\theta}\frac{\pi_{n1,j}}{T_{1,j}}\right)$ in Home government Euler for $T_{1,j}$ reflects dynamic terms of trade consideration. It can be fully implemented by country-sector specific export tax, as the term

$$\sum_{k}^{J}\left[\gamma_{n,k}^v\left(\frac{1}{\alpha_{n,k}} - \frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}}\right) - (1-\delta)\Gamma_{n,k,-1}^v\frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}}\right]w_n x_n^{-\sigma}P_n^{\sigma-1}(1-\sigma)\beta_j\frac{\pi_{n1,j}}{1+\tau_{n,j}^x}$$

shows up in the FOCs of export taxes (Eq.A.44). So, after substituting FOCs for export taxes, the dynamic term in Euler disappears.

Furthermore, sum over the FOCs for optimal tariffs

$$\begin{aligned}
&u_{c_1}x_1\beta_j\pi_{11,j} + \theta\beta_j\pi_{11,j}x_1\left\{\gamma_1^L\left[(1-\pi_{11,j})\right] - \sum_{n\neq 1}^{N}\frac{1}{1+\tau_{n,j}^m}\gamma_{n,j}^L\pi_{1n,j} - \gamma_x\sum_{n\neq 1}^{N}\frac{\tau_{n,j}^m}{1+\tau_{n,j}^m}\pi_{1m,j}\right\}\\
&+\left(\gamma_x - \gamma_1^L\right)\sum_{n\neq 1}\frac{1}{(1+\tau_{n,j}^x)}\beta_j\pi_{n1,j}x_n\\
&= u_{c_1}x_1\beta_j\pi_{11,j} + \left(\gamma_x - \gamma_1^L\right)\frac{1+\theta}{\theta}w_1L_{1,j}^p - \left(\gamma_x - \gamma_1^L\right)\beta_j\pi_{11,j}x_1\\
&= \left(\gamma_x - \gamma_1^L\right)\frac{1+\theta}{\theta}w_1L_{1,j}^p,
\end{aligned}$$

plug it into Eq.A.47

$$\gamma_{1,j}^{T} = \rho(1-\delta)E\left[\gamma_{1,j}^{T\prime}\right] + \frac{1}{\theta}\frac{1}{T_{1,j}}\left(\gamma_x - \gamma_1^L\right)\frac{1+\theta}{\theta}w_1L_{1,j}^p.$$

Country-sector-specific tariffs implement the intratemporal allocation across countries' sectors. So, after substituting FOCs for tariffs, those terms in Euler disappear.

The FOC for $L_{1,j}^r$ (Eq.A.39) and $L_{1,j}^p$ (Eq.A.40) imply

$$\gamma_{1,j}^{T} = \left(\gamma_x - \gamma_1^L\right)\frac{1+\theta}{\theta}\frac{w_1}{\alpha_{1,j}},$$

plug back into the above equation, we get

$$\left(\gamma_x - \gamma_1^L\right)\frac{1+\theta}{\theta}\frac{w_1}{\alpha_{1,j}} = \rho(1-\delta)\left(\gamma_4' - \gamma_1^{L'}\right)\frac{1+\theta}{\theta}\frac{w_1'}{\alpha_{1,j}} + \frac{1}{\theta}\frac{1}{T_{1,j}}\left(\gamma_x - \gamma_1^L\right)\frac{1+\theta}{\theta}w_1 L_{1,j}^p$$

Since $\gamma_x - \gamma_1^L = u_{c_1}$, it further reduces to:

Home government Euler on home innovation

$$\frac{w_1}{\alpha_{1,j}} = \frac{1}{\theta}\frac{w_1 L_{1,j}^p}{T_{1,j}} + \rho(1-\delta)E\left[\frac{u_{c_1}' w_1'}{u_{c_1}\alpha_{1,j}}\right]. \tag{A.48}$$

It is the same as in the private market equilibrium. The government does not use industry policies.

Optimal Ramsey export tax $\tau_{n,j}^x$

$$1+\tau_{n,j}^x = \frac{1+\theta(1-\pi_{n1,j})}{\theta\sum_{i\neq 1}(1+\tau_{i,j}^m)\pi_{ni,j} + \sum_k^J\left(\gamma_{n,k}^v\frac{1}{\alpha_{n,k}} - \Gamma_{n,k}^v\frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}}\right)w_n x_n^{-\sigma}P_n^{\sigma-1}(\sigma-1)/(u_{c_1}x_n)}$$

Optimal Ramsey tariff $\tau_{n,j}^m$

$$1+\tau_{n,j}^m = \frac{\gamma_x - \gamma_{n,j}^L}{u_{c_1}}.$$

F.3 Proof of $\sum_j^J \frac{\gamma_{n,j}^v}{\alpha_{n,j}} = 0$ at Ramsey steady state

FOC over $L_{n,j}^p$

$$\begin{aligned}&\frac{1}{\theta}\frac{w_n}{T_{n,j}}\Gamma_{n,j}^v x_n^{-\sigma}P_n^{\sigma-1} - \gamma_{n,j}^L\frac{1+\theta}{\theta}w_n - \mu_n + \sum_{k=1}^J\gamma_{1,k}^L\beta_k\frac{1}{1+\tau_{n,k}^x}\pi_{n1,k}\frac{1+\theta}{\theta}w_n + \sum_{i\neq 1}^N\sum_k^J\gamma_{i,k}^L\beta_k\pi_{n,i}\frac{1+\theta}{\theta}w_n\\ &+\gamma_x\sum_k^J\beta_k\frac{\tau_{n,k}^x}{1+\tau_{n,k}^x}\pi_{n1,k}\frac{1+\theta}{\theta}w_n + \sum_k^J\left[\frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}}\Gamma_{n,k}^v - \frac{1}{\alpha_{n,k}}\gamma_{n,k}^v\right]w_n x_n^{-\sigma}P_n^{\sigma-1}(-\sigma)\frac{1+\theta}{\theta}\frac{w_n}{x_n} = 0\end{aligned} \tag{A.49}$$

FOC over w_n

$$
\begin{aligned}
&- u_{c_1} x_1 \sum_j^J \beta_j \pi_{1n,j} + \gamma_1^L \left[\sum_{m\neq 1}^N \sum_j^J \beta_j \frac{1}{1+\tau_{m,j}^x} \theta \pi_{m1,j} \pi_{mn,j} x_m + \sum_j^J \beta_j \frac{1}{1+\tau_{n,j}^x} \pi_{n1,j} x_n \right] \\
&+ \sum_{i\neq\{1\}}^N \sum_j^J \gamma_{i,j}^L \beta_j \left(\sum_{m\neq 1} \theta \pi_{mi,j} \pi_{mn,j} x_m + \pi_{ni,j} x_n \right) - \sum_j^J \gamma_{n,j}^L \left(\theta \beta_j \sum_{m\neq 1} \pi_{mn,j} x_m + \frac{1+\theta}{\theta} w_n L_{n,j}^p \right) \\
&+ \gamma_x \sum_{j=1}^J \beta_j \left(\sum_{m\neq 1}^N \frac{\tau_{m,j}^x}{1+\tau_{m,j}^x} \theta \pi_{m1,j} \pi_{mn,j} x_m + \frac{\tau_{n,j}^x}{1+\tau_{n,j}^x} \pi_{n1,j} x_n \right) + (1-\sigma) \sum_j^J \left[\frac{1}{\theta} \frac{L_{n,j}^p}{T_{n,j}} \Gamma_{n,j}^v - \frac{1}{\alpha_{n,j}} \gamma_{n,j}^v \right] w_n x_n^{-\sigma} P_n^{\sigma-1} \\
&+ \sum_{m\neq 1}^N \sum_j^J \left[\frac{1}{\theta} \frac{L_{m,j}^p}{T_{m,j}} \Gamma_{m,j}^v - \frac{1}{\alpha_{m,j}} \gamma_{m,j}^v \right] w_m x_m^{-\sigma} P_m^{\sigma-1} (\sigma-1) \left(\sum_k \beta_k \pi_{mn,k} \right) = 0
\end{aligned} \tag{A.50}
$$

FOC over $T_{n,j}$

$$
\begin{aligned}
\gamma_{n,j}^T =& \rho E \left[\gamma_{n,j}^{T\prime} (1-\delta) \right] + \frac{1}{T_{n,j}} \{ u_{c_1} x_1 \frac{\beta_j}{\theta} \pi_{1n,j} \\
&- \sum_{i=1}^N \gamma_{i,j}^L \beta_j \sum_{m\neq 1} \pi_{mi,j} \pi_{mn,j} x_m + \gamma_{n,j}^L \beta_j \sum_{m\neq 1} \pi_{mn,j} x_m - u_{c_1} \sum_{m\neq 1}^N \beta_j \frac{\tau_{m,j}^x}{1+\tau_{m,j}^x} \pi_{m1,j} \pi_{mn,j} x_m \\
&- \left[\frac{1}{\theta} \frac{L_{n,j}^p}{T_{n,j}} \Gamma_{n,j}^v \right] w_n x_n^{-\sigma} P_n^{\sigma-1} - \sum_{i\neq 1} \sum_k^J \left[\frac{1}{\theta} \frac{L_{i,k}^p}{T_{i,k}} \Gamma_{i,k}^v - \frac{1}{\alpha_{i,k}} \gamma_{i,k}^v \right] w_i x_i^{-\sigma} P_i^{\sigma-1} (\sigma-1) \frac{\beta_j}{\theta} \pi_{in,j} \}
\end{aligned} \tag{A.51}
$$

We prove the steady state feature using the following steps:

1. At the steady state $\gamma_{n,j}^T = \gamma_{n,j}^{T\prime}$, and sum over j the FOC for $T_{n,j}$:

$$
\begin{aligned}
&- \theta(1-\rho(1-\delta)) \sum_j^J \gamma_{n,j}^T T_{n,j} + u_{c_1} x_1 \sum_j^J \beta_j \pi_{1n,j} - \theta u_{c_1} \sum_j^J \sum_{m\neq 1}^N \beta_j \frac{\tau_{m,j}^x}{1+\tau_{m,j}^x} \pi_{m1,j} \pi_{mn,j} x_m \\
&- \theta \sum_j^J \sum_{m\neq 1} \sum_{i\neq 1}^N \gamma_{i,j}^L \beta_j \pi_{mi,j} \pi_{mn,j} x_m + \theta \sum_j^J \gamma_{n,j}^L \beta_j \sum_{m\neq 1} \pi_{mn,j} x_m \\
&- \sum_j^J \left[\frac{L_{n,j}^p}{T_{n,j}} \Gamma_{n,j}^v \right] w_n x_n^{-\sigma} P_n^{\sigma-1} - \sum_{i\neq 1} \sum_k^J \left[\frac{1}{\theta} \frac{L_{i,k}^p}{T_{i,k}} \Gamma_{i,k}^v - \frac{1}{\alpha_{i,k}} \gamma_{i,k}^v \right] w_i x_i^{-\sigma} P_i^{\sigma-1} (\sigma-1) \sum_j^J \beta_j \pi_{in,j},
\end{aligned}
$$

2. substract FOC for $L_{n,j}^p$ (Eq.A.49) and w_n(Eq.A.50) from the above equation:

$$
\begin{aligned}
&\frac{1}{\theta} \frac{w_n L_{np}}{T_{n,j}} \Gamma_{n,j}^v x_n^{-\sigma} P_n^{\sigma-1} - \gamma_{n,j}^L x_n - \mu_n \sum_j^J L_{n,j}^p + \sum_j^J \gamma_{n,j}^L \frac{1+\theta}{\theta} w_n L_{n,j}^p + \theta(1-\rho(1-\delta)) \sum_j^J \gamma_{n,j}^T T_{n,j} \\
&+ \sum_j^J \left[\frac{L_{n,j}^p}{T_{n,j}} \Gamma_{n,j}^v \right] w_n x_n^{-\sigma} P_n^{\sigma-1} - \sum_k^J \left[\frac{1}{\theta} \frac{L_{n,k}^p}{T_{n,k}} \Gamma_{n,k}^v - \frac{1}{\alpha_{n,k}} \gamma_{n,k}^v \right] w_n x_n^{-\sigma} P_n^{\sigma-1} = 0
\end{aligned}
$$

3. at the steady state

$$T_{n,j} = \frac{\alpha_{n,j} L^r_{n,j}}{\delta}, \quad \mu_n = \gamma^T_{n,j} \alpha_{n,j},$$

hence

$$\frac{1}{\theta}\frac{w_n L_{np}}{T_{n,j}}\Gamma^v_{n,j} x_n^{-\sigma} P_n^{\sigma-1} - \gamma^L_{n,j} x_n - \mu_n \sum_j^J L^p_{n,j} + \sum_j^J \gamma^L_{n,j} \frac{1+\theta}{\theta} w_n L^p_{n,j} + \theta(1-\rho(1-\delta)) \sum_j \frac{1}{\delta} \mu_n L^r_{n,j}$$
$$+ \sum_j^J \left[\frac{L^p_{n,j}}{T_{n,j}} \Gamma^v_{n,j} \right] w_n x_n^{-\sigma} P_n^{\sigma-1} - \sum_k^J \left[\frac{1}{\theta} \frac{L^p_{n,k}}{T_{n,k}} \Gamma^v_{n,k} - \frac{1}{\alpha_{n,k}} \gamma^v_{n,k} \right] w_n x_n^{-\sigma} P_n^{\sigma-1} = 0 \tag{A.52}$$

At the steady state, the research intensity is given by

$$\bar{r}_j = \frac{\delta}{\theta(1-\rho(1-\delta)) + \delta}.$$

Using this research intensity, $L^p_{n,j} = (1-\bar{r}) L_{n,j}$, $L^p_n = (1-\bar{r}) L_n$, and $L^r_{n,j} = \bar{r} L_{n,j}$, we can cancel out the two terms involving μ_n in equation (A.52). Then we have

$$\frac{1}{\theta}\frac{w_n L^p_n}{T_{n,j}}\Gamma^v_{n,j} x_n^{-\sigma} P_n^{\sigma-1} - \gamma^L_{n,j} x_n + \sum_j^J \gamma^L_{n,j} \frac{1+\theta}{\theta} w_n L^p_{n,j}$$
$$+ \sum_j^J \left[\frac{L^p_{n,j}}{T_{n,j}} \Gamma^v_{n,j} \right] w_n x_n^{-\sigma} P_n^{\sigma-1} - \sum_k^J \left[\frac{1}{\theta} \frac{L^p_{n,k}}{T_{n,k}} \Gamma^v_{n,k} - \frac{1}{\alpha_{n,k}} \gamma^v_{n,k} \right] w_n x_n^{-\sigma} P_n^{\sigma-1} = 0.$$

4. Multiply the above equation $\frac{\frac{1+\theta}{\theta} w_n L^p_{n,j}}{x_n}$ on both sides and sum over j

$$\sum_j^J \left[\frac{L^p_{n,j}}{T_{n,j}} \Gamma^v_{n,j} \right] w_n x_n^{-\sigma} P_n^{\sigma-1} + \sum_k^J \frac{1}{\alpha_{n,k}} \gamma^v_{n,k} w_n x_n^{-\sigma} P_n^{\sigma-1} = 0.$$

Finally,

$$\sum_k^J \frac{\gamma^v_{n,k}}{\alpha_{n,k}} (1 + \frac{\theta(1-\rho(1-\delta))}{\delta}) w_n x_n^{-\sigma} P_n^{\sigma-1} = 0.$$

The equation only holds when for every $n \neq 1$, $\sum_k^J \frac{\gamma^v_{n,k}}{\alpha_{n,k}} = 0$.

G Characterization of Markov Policy in the Baseline

In this appendix, we first lay out the Markov problem, and then characterize the Markov policy by taking first-order conditions. In G.3, we prove the steady-state properties.

G.1 Markov problem

Home government (country 1) chooses optimal unilateral trade policies and domestic R&D, and solves

$$V\left(\left\{T_{n,j,-1}\right\}\right) = \max_{\left\{L^r_{n,j},L^p_{n,j}T_{n,j},w_n,x_n,\tau^x_{n,j},\tau^m_{n,j}\right\}} \frac{\left(x_1/P_1\right)^{1-\sigma}}{1-\sigma} + \rho\left[V\left(\left\{T_{n,j}\right\}\right)\right] \tag{A.53}$$

Subject to

$$\frac{w_n}{\alpha_{n,j}} = \frac{1}{\theta}\frac{w_n L^p_{n,j}}{T_{n,j}} + \tilde{\rho}_n v'_{n,j} \qquad (\gamma^r_{n,j}, \quad J\times(N-1)) \tag{A.54}$$

$$T_{n,j} = \alpha_{n,j}L^r_{n,j} + (1-\delta)T_{n,j,-1} \quad (\gamma^T_{n,j}, \quad J\times N) \tag{A.55}$$

$$\frac{1+\theta}{\theta}w_1 L^p_{1,j} = \beta_j\left[\pi_{11,j}x_1 + \sum_{m\neq 1}\frac{1}{1+\tau^x_{m,j}}\pi_{m1,j}x_m\right] \quad (\gamma^L_{1,j}, \quad J) \tag{A.56}$$

$$\frac{1+\theta}{\theta}w_n L^p_{n,j} = \beta_j\left[\frac{1}{1+\tau^m_{n,j}}\pi_{1n,j}x_1 + \sum_{m\neq 1}\pi_{mn,j}x_m\right] \quad (\gamma^L_{n,j}, \quad J\times(N-1)) \tag{A.57}$$

$$\sum_j\left(L^r_{n,j} + L^p_{n,j}\right) = L_n \quad (\mu_n, \quad N) \tag{A.58}$$

$$x_1 = \frac{1+\theta}{\theta}w_1\sum_j L^p_{1,j} + \sum_{m\neq 1}^{N}\sum_{j=1}^{J}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\pi_{m1,j}x_m + \sum_{i\neq 1}^{N}\sum_{j=1}^{J}\beta_j\frac{\tau^m_{i,j}}{1+\tau^m_{i,j}}\pi_{1i,j}x_1 \quad (\gamma_x) \tag{A.59}$$

where

$$x_m = \frac{1+\theta}{\theta}w_m\sum_j L^p_{m,j} \tag{A.60}$$

$$P_m = \Pi_j\left[T_{1,j}(w_1(1+\tau^x_{m,j})d_{m1})^{-\theta} + \sum_{n\neq 1}T_{n,j}(w_n d_{mn})^{-\theta}\right]^{-\beta_j/\theta} \tag{A.61}$$

$$P_1 = \Pi_j\left[T_{1,j}w_1^{-\theta} + \sum_{n\neq 1}T_{n,j}(w_n(1+\tau^m_{n,j})d_{1n})^{-\theta}\right]^{-\beta_j/\theta} = 1, \tag{A.62}$$

$$\pi_{11,j} = \frac{T_{1,j}(w_1)^{-\theta}}{T_{1,j}(w_1)^{-\theta} + \sum_{n\neq 1}T_{n,j}(w_n(1+\tau^m_{n,j})d_{1n})^{-\theta}} \tag{A.63}$$

$$\pi_{m1,j} = \frac{T_{1,j}(w_1(1+\tau^x_{m,j})d_{m1})^{-\theta}}{T_{1,j}(w_1(1+\tau^x_{m,j})d_{m1})^{-\theta} + \sum_{n\neq 1}T_{n,j}(w_n d_{mn})^{-\theta}} \tag{A.64}$$

$$\pi_{mn,j} = \frac{T_{n,j}(w_n d_{mn})^{-\theta}}{T_{1,j}(w_1(1+\tau^x_{m,j})d_{m1})^{-\theta} + \sum_{i\neq\{1\}}T_{i,j}(w_i d_{mi})^{-\theta}} \tag{A.65}$$

$$\pi_{1n,j} = \frac{T_{n,j}(w_n(1+\tau^m_{n,j})d_{1n})^{-\theta}}{T_{1,j}(w_1)^{-\theta} + \sum_{i\neq\{1\}}T_{i,j}(w_i(1+\tau^m_{i,j})d_{1i})^{-\theta}} \tag{A.66}$$

Note that $\sum_n \pi_{mn,j} = 1$ for any m and P_1 normalize to 1.

Note that sum up (A.56) we get the balanced trade condition for country 1

$$\sum_{n=1}^{N}\sum_{j=1}^{J}\beta_j\frac{1}{1+\tau_{n,j}^m}\pi_{1n,j}x_1 = \sum_{n=1}^{N}\sum_{j=1}^{J}\beta_j\pi_{n1,j}x_n,$$

and sum up (A.57) we get the balanced trade condition for country $n \neq 1$,

$$\sum_{m=1}^{N}\sum_{j=1}^{J}\beta_j\pi_{nm,j}x_n = \sum_{m\neq 1}^{N}\sum_{j=1}^{J}\beta_j\pi_{mn,j}x_m + \sum_{j=1}^{J}\beta_j\frac{1}{1+\tau_{n,j}^m}\pi_{1n,j}x_1.$$

As one of them is redundant, we can drop one of the (A.57), so we end up with number $(N-1)J-1$ for $\gamma_{n,j}^L$ with $n \neq 1$.

G.2 Optimal conditions for Markov policy

Here, we first take the first-order conditions over the Markov problem (A.53), assuming the interior solutions. We then derive the optimal conditions for Markov policy from these first-order conditions.

First order conditions For ease of notation, we make the following definitions:

$$G_{n,j} = \left\{\left(x_n'\right)^{-\sigma}\left(P_n'\right)^{\sigma-1}\frac{w_n'}{\alpha_{n,j}}\right\}, \qquad M_{n,j} = G_{n,j}x_n^{\sigma}P_n^{1-\sigma}, \tag{A.67}$$

where $G_{n,j}$ captures the future value of a researcher in future prices, and $M_{n,j}$ is the future value of the researcher in current prices.

FOC over x_1: (with $P_1 = 1$)

$$u_{c_1} + \sum_{n=1}^{N}\sum_{j}\gamma_{n,j}^L\beta_j\frac{1}{1+\tau_{n,j}^m}\pi_{1n,j} - \gamma_x\left(1 - \sum_{n\neq 1}^{N}\sum_{j=1}^{J}\beta_j\frac{\tau_{n,j}^m}{1+\tau_{n,j}^m}\pi_{1n,j}\right) = 0. \tag{A.68}$$

FOC over $L_{1,j}^r$:

$$\mu_1 = \gamma_{1,j}^T\alpha_{1,j}, \tag{A.69}$$

FOC over $L_{1,j}^p$:

$$\mu_1 = \left(\gamma_x - \gamma_{1,j}^L\right)\frac{1+\theta}{\theta}w_1. \tag{A.70}$$

FOC over $L^r_{n,j}$ for $n > 1$:

$$\mu_n = \gamma^T_{n,j} \alpha_{n,j}, \tag{A.71}$$

FOC over $L^p_{n,j}$ for $n > 1$:

$$\begin{aligned}\mu_n = {} & \gamma^r_{n,j} \frac{1}{\theta} \frac{w_n}{T_{n,j}} - \gamma^L_{n,j} \frac{1+\theta}{\theta} w_n + \sum_{k=1}^{J} \gamma^L_{1,k} \beta_k \frac{1}{1+\tau^x_{n,k}} \pi_{n1,k} \frac{1+\theta}{\theta} w_n \\ & + \sum_{i\neq 1}^{N} \sum_{k}^{J} \gamma^L_{i,k} \beta_k \pi_{ni,k} \frac{1+\theta}{\theta} w_n + \gamma_x \sum_{k}^{J} \beta_k \frac{\tau^x_{n,k}}{1+\tau^x_{n,k}} \pi_{n1,k} \frac{1+\theta}{\theta} w_n + \sum_{k}^{J} \gamma^r_{n,k} \rho (1-\delta) \sigma M_{n,k} \frac{1+\theta}{\theta} \frac{w_n}{x_n}\end{aligned} \tag{A.72}$$

FOC over $\tau^x_{n,j}$ for $n > 1$:

$$\begin{aligned}& \left(\sum_k \gamma^r_{n,k} \rho (1-\delta)(1-\sigma) M_{n,k} \right) \beta_j \pi_{n1,j} - \gamma^L_{1,j} \frac{1}{(1+\tau^x_{n,j})} \beta_j \pi_{n1,j} x_n - \gamma^L_{1,j} \frac{1}{(1+\tau^x_{n,j})} \theta (1-\pi_{n1,j}) \beta_j \pi_{n1,j} x_n \\ & + \sum_{i\neq 1}^{N} \gamma^L_{i,j} \beta_j \left[\theta \pi_{ni,j} \pi_{n1,j} x_n \right] + \gamma_x \beta_j \frac{1}{(1+\tau^x_{n,j})} \pi_{n1,j} x_n - \gamma_x \beta_j \frac{\tau^x_{n,j}}{1+\tau^x_{n,j}} \theta \pi_{n1,j} (1-\pi_{n1,j}) x_n = 0\end{aligned} \tag{A.73}$$

FOC over $\tau^m_{n,j}$ for $n > 1$:

$$\begin{aligned}& - u_{c_1} x_1 \beta_j \frac{\pi_{1n,j}}{1+\tau^m_{n,j}} + \gamma^L_{1,j} \beta_j \frac{\partial \pi_{11,j}}{\partial \tau^m_{n,j}} x_1 - \gamma^L_{n,j} \beta_j \frac{\pi_{1n,j}}{\left(1+\tau^m_{n,j}\right)^2} x_1 + \gamma^L_{n,j} \frac{1}{1+\tau^m_{n,j}} \frac{\partial \pi_{1n,j}}{\partial \tau^m_{n,j}} \beta_j x_1 \\ & + \sum_{i\neq\{1,n\}} \gamma^L_{i,j} \frac{1}{1+\tau^m_{i,j}} \frac{\partial \pi_{1i,j}}{\partial \tau^m_{n,j}} \beta_j x_1 + \gamma_x \beta_j \frac{1}{\left(1+\tau^m_{n,j}\right)^2} \pi_{1n,j} x_1 + \gamma_x \sum_{i\neq 1}^{N} \beta_j \frac{\tau^m_{i,j}}{1+\tau^m_{i,j}} \frac{\partial \pi_{1i,j}}{\partial \tau^m_{n,j}} x_1 = 0\end{aligned} \tag{A.74}$$

FOC over $T_{1,j}$:

$$\begin{aligned}& - \gamma^T_{1,j} + \rho \frac{\partial V}{\partial T_{1,j}} + u_{c_1} x_1 \frac{\beta_j}{\theta} \frac{\pi_{11,j}}{T_{1,j}} - \left(\sum_{i\neq 1}^{N} \sum_{k}^{J} \gamma^r_{i,k} \rho (1-\delta)(1-\sigma) M_{i,k} \right) \left(\frac{\beta_j}{\theta} \frac{\pi_{i1,j}}{T_{1,j}} \right) \\ & + \gamma^L_{1,j} \beta_j \left[\frac{\partial \pi_{11,j}}{\partial T_{1,j}} x_1 + \sum_{m\neq 1}^{N} \frac{1}{1+\tau^x_{m,j}} \frac{\partial \pi_{m1,j}}{\partial T_{1,j}} x_2 \right] + \sum_{n\neq 1}^{N} \gamma^L_{n,j} \beta_j \left[\frac{1}{1+\tau^m_{n,j}} \frac{\partial \pi_{1n,j}}{\partial T_{1,j}} x_1 + \sum_{m\neq 1} \frac{\partial \pi_{mn,j}}{\partial T_{1,j}} x_m \right] \\ & + \gamma_x \sum_{m\neq 1}^{N} \beta_j \frac{\tau^x_{m,j}}{1+\tau^x_{m,j}} \frac{\partial \pi_{m1,j}}{\partial T_{1,j}} x_m + \gamma_x \sum_{i\neq 1}^{N} \beta_j \frac{\tau^m_{i,j}}{1+\tau^m_{i,j}} \frac{\partial \pi_{1i,j}}{\partial T_{1,j}} x_1 + \rho (1-\delta) \sum_{n\neq 1}^{J} \sum_{k} \gamma^r_{n,k} \frac{\partial G_{n,k}}{\partial T_{1,j}} x_n^{\sigma} P_n^{1-\sigma} = 0\end{aligned} \tag{A.75}$$

where $\frac{\partial V}{\partial T_{1,j,-1}} = \gamma^T_{1,j} (1-\delta)$ by the envelop theorem.

FOC over $T_{n,j}$ for $n > 1$:

$$\begin{aligned}
\gamma^T_{n,j} = {} & \rho\left[\gamma^{T\prime}_{n,j}(1-\delta)\right] + u_{c_1}x_1\frac{\beta_j}{\theta}\frac{\pi_{1n,j}}{T_{n,j}} \\
& + \gamma^L_{1,j}\beta_j\left[\frac{\partial\pi_{11,j}}{\partial T_{n,j}}x_1 + \sum_{m\neq 1}\frac{1}{1+\tau^x_{m,j}}\frac{\partial\pi_{m1,j}}{\partial T_{n,j}}x_m\right] + \sum_{i\neq 1}^{N}\gamma^L_{i,j}\beta_j\left[\frac{1}{1+\tau^m_{i,j}}\frac{\partial\pi_{1i,j}}{\partial T_{n,j}}x_1 + \sum_{m\neq 1}\frac{\partial\pi_{mi,j}}{\partial T_{n,j}}x_m\right] \\
& + \gamma_x\sum_{m\neq 1}^{N}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\frac{\partial\pi_{m1,j}}{\partial T_{n,j}}x_m + \gamma_x\sum_{i\neq 1}^{N}\beta_j\frac{\tau^m_{i,j}}{1+\tau^m_{i,j}}\frac{\partial\pi_{1i,j}}{\partial T_{n,j}}x_1 - \gamma^r_{n,j}\frac{1}{\theta}\frac{w_n L^p_{n,j}}{\left(T_{n,j}\right)^2} \\
& - \sum_{i\neq 1}^{N}\left(\sum_k^J\gamma^r_{i,k}\rho(1-\delta)(1-\sigma)M_{i,k}\right)\frac{\beta_j}{\theta}\frac{\pi_{in,j}}{T_{n,j}} + \sum_{i\neq 1}^{N}\sum_k^J\gamma^r_{i,k}\rho(1-\delta)\frac{\partial G_{i,k}}{\partial T_{n,j}}x_i^{\sigma}P_i^{1-\sigma}
\end{aligned} \tag{A.76}$$

FOC over w_1:

$$\begin{aligned}
& - u_{c_1}x_1\sum_j^J\beta_j\frac{\pi_{11,j}}{w_1} - \sum_j^J\gamma^L_{1,j}\frac{1+\theta}{\theta}L^p_{1,j} + \sum_j^J\gamma^L_{1,j}\beta_j\frac{\partial\pi_{11,j}}{\partial w_1}x_1 + \sum_j^J\gamma^L_{1,j}\beta_j\sum_{m\neq 1}^{N}\frac{1}{1+\tau^x_{m,j}}\frac{\partial\pi_{m1,j}}{\partial w_1}x_m \\
& + \sum_{n\neq 1}^{N}\sum_j^J\gamma^L_{n,j}\beta_j\frac{1}{1+\tau^m_{n,j}}\frac{\partial\pi_{1n,j}}{\partial w_1}x_1 + \sum_{n\neq 1}^{N}\sum_j^J\gamma^L_{n,j}\beta_j\sum_{m\neq 1}^{N}\frac{\partial\pi_{mn,j}}{\partial w_1}x_m \\
& + \gamma_x\frac{1+\theta}{\theta}\sum_j L^p_{1,j} + \gamma_x\sum_{m\neq 1}^{N}\sum_{j=1}^{J}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\frac{\partial\pi_{m1,j}}{\partial w_1}x_m + \gamma_x\sum_{i\neq 1}^{N}\sum_{j=1}^{J}\beta_j\frac{\tau^m_{i,j}}{1+\tau^m_{i,j}}\frac{\partial\pi_{1i,j}}{\partial w_1}x_1 \\
& + \sum_{n\neq 1}^{N}\sum_j^J\gamma^r_{n,j}\rho(1-\delta)(1-\sigma)M_{n,j}\left(\sum_k^J\beta_k\pi_{n1,k}\right)w_1^{-1} = 0
\end{aligned} \tag{A.77}$$

FOC over w_n for $n > 1$:

$$\begin{aligned}
& - u_{c_1}x_1\sum_j^J\beta_j\frac{\pi_{1n,j}}{w_n} + \sum_j^J\gamma^L_{1,j}\beta_j\frac{\partial\pi_{11,j}}{\partial w_n}x_1 + \sum_j^J\gamma^L_{1,j}\beta_j\sum_{m\neq 1}^{N}\frac{1}{1+\tau^x_{m,j}}\frac{\partial\pi_{m1,j}}{\partial w_n}x_m + \sum_j^J\gamma^L_{1,j}\beta_j\frac{1}{1+\tau^x_{n,j}}\pi_{n1,j}\frac{x_n}{w_n} \\
& - \sum_j^J\gamma^L_{n,j}\frac{1+\theta}{\theta}L^p_{n,j} + \sum_{i\neq 1}^{N}\sum_j^J\gamma^L_{i,j}\beta_j\frac{1}{1+\tau^m_{i,j}}\frac{\partial\pi_{1i,j}}{\partial w_n}x_1 + \sum_{i\neq 1}^{N}\sum_j^J\gamma^L_{i,j}\beta_j\sum_{m\neq 1}\frac{\partial\pi_{mi,j}}{\partial w_n}x_m + \sum_{i\neq 1}^{N}\sum_j^J\gamma^L_{i,j}\beta_j\pi_{ni,j}\frac{x_n}{w_n} \\
& + \gamma_x\sum_{m\neq 1}^{N}\sum_{j=1}^{J}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\frac{\partial\pi_{m1,j}}{\partial w_n}x_m + \gamma_x\sum_{j=1}^{J}\beta_j\frac{\tau^x_{n,j}}{1+\tau^x_{n,j}}\pi_{n1,j}\frac{x_n}{w_n} + \gamma_x\sum_{i\neq 1}^{N}\sum_{j=1}^{J}\beta_j\frac{\tau^m_{i,j}}{1+\tau^m_{i,j}}\frac{\partial\pi_{1i,j}}{\partial w_n}x_1 \\
& + \sum_j^J\gamma^r_{n,j}\left[\frac{1}{\theta}\frac{L^p_{n,j}}{T_{n,j}} - \frac{1}{\alpha_{n,j}} + \rho(1-\delta)\sigma M_{n,j}\frac{1}{w_n}\right] + \sum_{m\neq 1}^{N}\sum_j^J\gamma^r_{m,j}\left[\rho(1-\delta)(1-\sigma)M_{m,j}\left(\sum_k\beta_k\frac{\pi_{mn,k}}{w_n}\right)\right] = 0
\end{aligned} \tag{A.78}$$

Simplifying conditions First, according to FOC for $L^p_{1,j}$ (A.70), it has to be the case that $\gamma^L_{1,j} = \gamma^L_1$ for any sector j at Home country. This implies that the marginal return of production labor has to be equalized across sectors. We can further prove that $u_{c_1} = \gamma_x - \gamma^L_1$ later.

With the condition $\gamma_{1,j}^L = \gamma_1^L$, we can further simplify the optimal export tax equation (A.73) as, for $n > 1$ and $j \in [1, J]$

$$1 + \tau_{n,j}^x = \frac{\left(\gamma_x - \gamma_1^L\right)\left[1 + \theta(1 - \pi_{n1,j})\right]}{\gamma_x \theta(1 - \pi_{n1,j}) - \theta\left(\sum_{i\neq 1}^N \gamma_{i,j}^L \pi_{ni,j}\right) + \left(\sum_k \gamma_{n,k}^r \rho(1-\delta)(\sigma - 1) M_{n,k}\right) x_n^{-1}}. \tag{A.79}$$

For the import tariff, we can plug in the derivatives and further simplify the equation (A.74) to

$$\frac{1}{1 + \tau_{n,j}^m} = \frac{u_{c_1} - \sum_{i\neq\{1,n\}}^N \left(\gamma_x - \gamma_{i,j}^L\right) \frac{\tau_{i,j}^m}{1+\tau_{i,j}^m} \theta \pi_{1i,j} + \gamma_x \theta(1 - \pi_{1n,j}) - \gamma_1^L \theta \pi_{11,j} - \sum_{i\neq\{1,n\}} \gamma_{i,j}^L \theta \pi_{1i,j}}{\left(\gamma_x - \gamma_{n,j}^L\right)\left[1 + \theta(1 - \pi_{1n,j})\right]}$$

or equivalently,

$$1 + \tau_{n,j}^m = \frac{\left(\gamma_x - \gamma_{n,j}^L\right)\left[1 + \theta(1 - \pi_{1n,j})\right]}{\left(\gamma_x - \gamma_1^L\right)\theta \pi_{11,j} + u_{c_1} + \sum_{i\neq\{1,n\}}^N \left(\gamma_x - \gamma_{i,j}^L\right) \frac{1}{1+\tau_{i,j}^m} \theta \pi_{1i,j}}$$

combining with the FOC for x_1 (Eq.A.68), we arrive at $u_{c_1} = \gamma_x - \gamma_1^L$, and for each foreign country $n > 1$ and $j \in [1, J]$

$$1 + \tau_{n,j}^m = \frac{\gamma_x - \gamma_{n,j}^L}{u_{c_1}}. \tag{A.80}$$

To get the innovation policy, we combine the FOCs for $L_{1,j}^r$ (A.69), $L_{1,j}^p$ (A.70) and trade policies (A.73), (A.74) with $T_{1,j}$ (A.75). Specifically, putting in envelope and derivatives

$$\begin{aligned}
\gamma_{1,j}^T = {} & \rho\left[\gamma_{1,j}^{T\prime}(1-\delta)\right] + \frac{1}{\theta}\frac{1}{T_{1,j}}\Bigg[u_{c_1} x_1 \beta_j \pi_{11,j} - \left(\sum_{n\neq 1}^N \sum_k^J \gamma_{n,k}^r \rho(1-\delta)(1-\sigma) M_{n,k}\right)(\beta_j \pi_{n1,j}) \\
& + \gamma_1^L \theta \beta_j \left[\pi_{11,j}(1 - \pi_{11,j}) x_1 + \sum_{m\neq 1}^N \frac{1}{1+\tau_{m,j}^x} \pi_{m1,j}(1 - \pi_{m1,j}) x_m\right] \\
& - \sum_{n\neq 1}^N \gamma_{n,j}^L \theta \beta_j \left[\frac{1}{1+\tau_{n,j}^m} \pi_{11,j} \pi_{1n,j} x_1 + \sum_{m\neq 1}^N \pi_{m1,j} \pi_{mn,j} x_m\right] + \theta \gamma_x \sum_{n\neq 1}^N \beta_j \frac{\tau_{n,j}^x}{1+\tau_{n,j}^x} \pi_{n1,j}(1 - \pi_{n1,j}) x_n \\
& - \theta \gamma_x \sum_{n\neq 1}^N \beta_j \frac{\tau_{n,j}^m}{1+\tau_{n,j}^m} \pi_{11,j} \pi_{1m,j} x_1 + \theta \sum_{n\neq 1}^N \sum_k^J \gamma_{n,k}^r \rho(1-\delta) \frac{\partial G_{n,k}}{\partial T_{1,j}} x_n^\sigma P_n^{1-\sigma} T_{1,j}\Bigg] = 0,
\end{aligned}$$

sum over the FOC for $\tau_{n,j}^x$, over $n \neq 1$ (Eq.A.73), and substitute it into the above equation, we can

further write it as:

$$\gamma_{1,j}^{T} = \rho\left[\gamma_{1,j}^{T\prime}(1-\delta)\right] + \frac{1}{\theta}\frac{1}{T_{1,j}}\{u_{c_1}x_1\beta_j\pi_{11,j} + \gamma_1^L\theta\beta_j\left[\pi_{11,j}(1-\pi_{11,j})x_1\right] + \sum_{n\neq 1}\left(\gamma_x - \gamma_1^L\right)\frac{1}{(1+\tau_{n,j}^x)}\beta_j\pi_{n1,j}x_n$$
$$-\theta\pi_{11,j}x_1\sum_{n\neq 1}^{N}\frac{1}{1+\tau_{n,j}^m}\gamma_{n,j}^L\beta_j\pi_{1n,j} - \theta\gamma_x\sum_{n\neq 1}^{N}\beta_j\frac{\tau_{n,j}^m}{1+\tau_{n,j}^m}\pi_{11,j}\pi_{1m,j}x_1\} + \sum_{n\neq 1}^{N}\sum_{k}^{J}\gamma_{n,k}^r\rho(1-\delta)\frac{\partial G_{n,k}}{\partial T_{1,j}}x_n^{\sigma}P_n^{1-\sigma}.$$

furthermore, because optimal tariffs satisfy

$$u_{c_1}x_1\beta_j\pi_{11,j} + \theta\beta_j\pi_{11,j}x_1\left\{\gamma_1^L(1-\pi_{11,j}) - \sum_{n\neq 1}^{N}\frac{1}{1+\tau_{n,j}^m}\gamma_{n,j}^L\pi_{1n,j} - \gamma_x\sum_{n\neq 1}^{N}\frac{\tau_{n,j}^m}{1+\tau_{n,j}^m}\pi_{1m,j}\right\}$$
$$+\left(\gamma_x - \gamma_1^L\right)\sum_{n\neq 1}\frac{1}{(1+\tau_{n,j}^x)}\beta_j\pi_{n1,j}x_n$$
$$= u_{c_1}x_1\beta_j\pi_{11,j} + \left(\gamma_x - \gamma_1^L\right)\frac{1+\theta}{\theta}w_1L_{1,j}^p - \left(\gamma_x - \gamma_1^L\right)\beta_j\pi_{11,j}x_1 = \left(\gamma_x - \gamma_1^L\right)\frac{1+\theta}{\theta}w_1L_{1,j}^p,$$

hence FOC for T_{1j} can be further simplified

$$\gamma_{1,j}^{T} = \rho\left[\gamma_{1,j}^{T\prime}(1-\delta)\right] + \frac{1}{\theta}\frac{1}{T_{1,j}}\left(\gamma_x - \gamma_1^L\right)\frac{1+\theta}{\theta}w_1L_{1,j}^p + \sum_{n\neq 1}^{N}\sum_{k}^{J}\gamma_{n,k}^r\rho(1-\delta)\frac{\partial G_{n,k}}{\partial T_{1,j}}x_n^{\sigma}P_n^{1-\sigma}.$$

The FOC for $L_{1,j}^r$ implies

$$\gamma_{1,j}^{T} = \left(\gamma_x - \gamma_1^L\right)\frac{1+\theta}{\theta}\frac{w_1}{\alpha_{1,j}}$$

plug back into the above equation, we have the optimal conditions for Home innovation for each sector j:

$$\frac{w_1}{\alpha_{1,j}} = \frac{1}{\theta}\frac{w_1L_{1,j}^p}{T_{1,j}} + \rho\left[\frac{u_{c_1}'}{u_{c_1}}\frac{w_1'}{\alpha_{1,j}}(1-\delta)\right] + \frac{\theta}{1+\theta}\frac{1}{u_{c_1}}\sum_{n\neq 1}^{N}\sum_{k}^{J}\gamma_{n,k}^r\rho(1-\delta)\frac{\partial G_{n,k}}{\partial T_{1,j}}x_n^{\sigma}P_n^{1-\sigma}, \tag{A.81}$$

this gives us the optimal innovation policy. Note that the Markov government optimal innovation is different from the private market equilibrium condition.

The rest of the FOCs help to pin down multipliers. Similarly, plugging in the derivatives and FOCs for tariffs $\tau_{n,j}^m$ (A.74) to FOC for $T_{n,j}$ (A.76) , we have a simplified FOC for $T_{n,j}$ for each country

$n > 1$ and $j \in [1, J]$.

$$\begin{aligned}\gamma^T_{n,j} =& \rho\left[\gamma^{T\prime}_{n,j}(1-\delta)\right] + \frac{1}{T_{n,j}}\{u_{c_1}x_1\frac{\beta_j}{\theta}\pi_{1n,j} \\ &- \sum_{i=1}^{N}\gamma^L_{i,j}\beta_j\sum_{m\neq 1}\pi_{mi,j}\pi_{mn,j}x_m + \gamma^L_{n,j}\beta_j\sum_{m\neq 1}\pi_{mn,j}x_m - u_{c_1}\sum_{m\neq 1}^{N}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\pi_{m1,j}\pi_{mn,j}x_m \\ &- \gamma^r_{n,j}\frac{1}{\theta}\frac{w_n L^p_{n,j}}{(T_{n,j})} - \sum_{i\neq 1}^{N}\left(\sum_k^J \gamma^r_{i,k}\rho(1-\delta)(1-\sigma)M_{i,k}\right)\frac{\beta_j}{\theta}\pi_{in,j}\} + \sum_{i\neq 1}^{N}\sum_k^J\gamma^r_{i,k}\rho(1-\delta)\frac{\partial G_{i,k}}{\partial T_{n,j}}x_i^\sigma P_i^{1-\sigma}.\end{aligned} \tag{A.82}$$

Using the derivatives and the optimal tariff conditions (A.74), we can write the FOC for w_1 (A.77) as

$$\begin{aligned}&-u_{c_1}x_1\sum_j^J\beta_j\pi_{11,j} - \gamma^L_1\sum_j^J\left[\frac{1+\theta}{\theta}w_1L^p_{1,j} + \beta_j\sum_{m\neq 1}^{N}\frac{1}{1+\tau^x_{m,j}}\theta\pi_{m1,j}(1-\pi_{m1,j})x_m + \beta_j\theta\sum_{m\neq 1}^{N}\pi_{m1,j}\pi_{m1,j}x_m\right] \\ &+\sum_n^N\sum_j^J\gamma^L_{n,j}\beta_j\sum_{m\neq 1}^{N}\theta\pi_{mn,j}\pi_{m1,j}x_m + \gamma_x\sum_j^J\left[\frac{1+\theta}{\theta}w_1L^p_{1,j} - \sum_{m\neq 1}^{N}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\theta\pi_{m1,j}(1-\pi_{m1,j})x_m\right] \\ &+\sum_{n\neq 1}^{N}\sum_j^J\gamma^r_{n,j}\rho(1-\delta)(1-\sigma)M_{n,j}\left(\sum_k^J\beta_k\pi_{n1,k}\right) = 0,\end{aligned}$$

and FOC for w_n for $n > 1$ (A.78) as

$$\begin{aligned}&-u_{c_1}x_1\sum_j^J\beta_j\pi_{1n,j} + \gamma^L_1\left[\sum_{m\neq 1}^{N}\sum_j^J\beta_j\frac{1}{1+\tau^x_{m,j}}\theta\pi_{m1,j}\pi_{mn,j}x_m + \sum_j^J\beta_j\frac{1}{1+\tau^x_{n,j}}\pi_{n1,j}x_n\right] \\ &+\sum_{i\neq\{1\}}^{N}\sum_j^J\gamma^L_{i,j}\beta_j\left(\sum_{m\neq 1}\theta\pi_{mi,j}\pi_{mn,j}x_m + \pi_{ni,j}x_n\right) - \sum_j^J\gamma^L_{n,j}\left(\theta\beta_j\sum_{m\neq 1}\pi_{mn,j}x_m + \frac{1+\theta}{\theta}w_nL^p_{n,j}\right) \\ &+\gamma_x\sum_{j=1}^{J}\beta_j\left(\sum_{m\neq 1}^{N}\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\theta\pi_{m1,j}\pi_{mn,j}x_m + \frac{\tau^x_{n,j}}{1+\tau^x_{n,j}}\pi_{n1,j}x_n\right) \\ &+\sum_j^J\gamma^r_{n,j}\left[\rho(1-\delta)(\sigma-1)M_{n,j}\right] + \sum_{m\neq 1}^{N}\sum_j^J\gamma^r_{m,j}\left[\rho(1-\delta)(1-\sigma)M_{m,j}\left(\sum_k^J\beta_k\pi_{mn,k}\right)\right] = 0.\end{aligned} \tag{A.83}$$

Summary of optimal Markov conditions The Markov equilibrium includes the variables

$$\Big\{\underbrace{L^r_{n,j}}_{J\times N}, \underbrace{L^p_{n,j}}_{J\times N}, \underbrace{T_{n,j}}_{J\times N}, \underbrace{\tau^x_{n,j}}_{J\times(N-1)}, \underbrace{\tau^m_{n,j}}_{J\times(N-1)}, \underbrace{\gamma^L_{n,j}}_{J\times(N-1)-1}, \underbrace{\gamma^r_{n,j}}_{J\times(N-1)}, \underbrace{\gamma^T_{n,j}}_{J\times(N-1)}, \underbrace{\mu_n}_{N-1}, \underbrace{w_n}_{N}, \gamma^L_1, \gamma_x\Big\}$$

that satisfies the private equilibrium (A.54)-(A.66), the optimal Home innovation (A.81), export tax formula (A.79), import tariff formula (A.80), FOC for foreign research effort $L^r_{n,j}$ (A.71), FOC for foreign production labor $L^p_{n,j}$ (A.72), FOC for foreign $T_{n,j}$ (A.82), FOC for foreign wages (A.83), and

the condition $u_{c_1} = \gamma_x - \gamma_1^L$.

G.3 Proof of $\sum_{j=1}^{J} \frac{\gamma_{n,j}^r}{\alpha_{n,j}} = 0$ at the steady state

In this appendix, we prove the following equation holds at the steady state

$$\sum_{j}^{J} \frac{\gamma_{n,j}^r}{\alpha_{n,j}} = 0. \tag{A.84}$$

1. Because of tax neutrality, we can normalize $\gamma_1^L = 0$, then $\gamma_x = u_{c_1}$.

2. Use the FOC for $L_{n,j}^p$ (A.72), we get

$$\begin{aligned}
&\gamma_{n,j}^r \frac{1}{\theta} \frac{w_n}{T_{n,j}} - \gamma_{n,j}^L \frac{1+\theta}{\theta} w_n - \mu_n + \sum_{i\neq 1}^{N} \sum_{k}^{J} \gamma_{i,k}^L \beta_k \pi_{ni} \frac{1+\theta}{\theta} w_n \\
&+ u_{c_1} \sum_{k}^{J} \beta_k \frac{\tau_{n,k}^x}{1+\tau_{n,k}^x} \pi_{n1,k} \frac{1+\theta}{\theta} w_n + \sum_{k}^{J} \gamma_{n,k}^r \rho(1-\delta)\sigma M_{n,k} \frac{1+\theta}{\theta} \frac{w_n}{x_n} = 0
\end{aligned}$$

multiply by total production labor $\sum_j L_{n,j}^p$ and use the relation $x_n = \frac{1+\theta}{\theta} w_n \sum_j L_{n,j}^p$,

$$\begin{aligned}
&\gamma_{n,j}^r \frac{1}{1+\theta} \frac{x_n}{T_{n,j}} - \gamma_{n,j}^L x_n - \mu_n \sum_j L_{n,j}^p + \sum_{i\neq 1}^{N} \sum_{k}^{J} \gamma_{i,k}^L \beta_k \pi_{ni,k} x_n \\
&+ u_{c_1} \sum_{k}^{J} \beta_k \frac{\tau_{n,k}^x}{1+\tau_{n,k}^x} \pi_{n1,k} x_n + \sum_{k}^{J} \gamma_{n,k}^r \rho(1-\delta)\sigma M_{n,k} = 0
\end{aligned} \tag{A.85}$$

3. Use FOC for w_n (A.83), we get

$$\begin{aligned}
&- u_{c_1} x_1 \sum_{j}^{J} \beta_j \pi_{1n,j} + \sum_{i\neq 1}^{N} \sum_{j}^{J} \gamma_{i,j}^L \beta_j \left(\sum_{m\neq 1} \theta \pi_{mi,j} \pi_{mn,j} x_m \right) \\
&+ \sum_{i\neq 1}^{N} \sum_{j}^{J} \gamma_{i,j}^L \beta_j \left(\pi_{ni,j} x_n \right) - \sum_{j}^{J} \gamma_{n,j}^L \left(\theta \beta_j \sum_{m\neq 1} \pi_{mn,j} x_m + \frac{1+\theta}{\theta} w_n L_{n,j}^p \right) \\
&+ u_{c_1} \sum_{j=1}^{J} \beta_j \left(\sum_{m\neq 1}^{N} \frac{\tau_{m,j}^x}{1+\tau_{m,j}^x} \theta \pi_{m1,j} \pi_{mn,j} x_m + \frac{\tau_{n,j}^x}{1+\tau_{n,j}^x} \pi_{n1,j} x_n \right) + \sum_{j}^{J} \gamma_{n,j}^r \left[\rho(1-\delta)(\sigma-1) M_{n,j} \right] \\
&+ \sum_{m\neq 1}^{N} \sum_{j}^{J} \gamma_{m,j}^r \left[\rho(1-\delta)(1-\sigma) M_{m,j} \left(\sum_{k}^{J} \beta_k \pi_{mn,k} \right) \right] = 0,
\end{aligned} \tag{A.86}$$

4. Substract equation (A.86) from equation (A.85):

$$\gamma^r_{n,j}\frac{1}{1+\theta}\frac{x_n}{T_{n,j}} - \gamma^L_{n,j}x_n - \mu_n\sum_j^J(L^p_{n,j}) + u_{c_1}x_1\sum_j^J\beta_j\pi_{1n,j} - \sum_{i\neq1}^N\sum_j^J\gamma^L_{i,j}\beta_j\left(\sum_{m\neq1}\theta\pi_{mi,j}\pi_{mn,j}x_m\right)$$
$$+\sum_j^J\gamma^L_{n,j}\theta\beta_j\sum_{m\neq1}\pi_{mn,j}x_m + \sum_j^J\gamma^L_{n,j}\frac{1+\theta}{\theta}w_nL^p_{n,j} - u_{c_1}\sum_{j=1}^J\beta_j\left(\sum_{m\neq1}^N\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\theta\pi_{m1,j}\pi_{mn,j}x_m\right)$$
$$+\sum_j^J\gamma^r_{n,j}\left[\rho(1-\delta)M_{n,j}\right] + \sum_{m\neq1}^N\sum_j^J\gamma^r_{m,j}\left[\rho(1-\delta)(\sigma-1)M_{m,j}\left(\sum_k^J\beta_k\pi_{mn,k}\right)\right] = 0, \tag{A.87}$$

5. At the steady state $\gamma^T_{n,j} = \gamma^{T\prime}_{n,j}$, the FOC for $T_{n,j}$ for $n > 1$ (A.82) becomes,

$$\theta(1-\rho(1-\delta))\gamma^T_{n,j}T_{n,j} = -\gamma^r_{n,j}\frac{w_nL^p_{n,j}}{T_{n,j}} + u_{c_1}x_1\beta_j\pi_{1n,j} - \theta u_{c_1}\sum_{m\neq1}^N\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\pi_{m1,j}\pi_{mn,j}x_m$$
$$-\theta\sum_{m\neq1}\sum_{i\neq1}^N\gamma^L_{i,j}\beta_j\pi_{mi,j}\pi_{mn,j}x_m + \theta\gamma^L_{n,j}\beta_j\sum_{m\neq1}\pi_{mn,j}x_m$$
$$-\sum_{i\neq1}^N\left(\sum_k^J\gamma^r_{i,k}\rho(1-\delta)(1-\sigma)M_{i,k}\right)\beta_j\pi_{in,j} + \theta T_{n,j}\sum_{i\neq1}^N\sum_k^J\gamma^r_{i,k}\rho(1-\delta)\frac{\partial G_{i,k}}{\partial T_{n,j}}x_i^\sigma P_i^{1-\sigma} \tag{A.88}$$

summing over j, we get

$$-\theta(1-\rho(1-\delta))\sum_j^J\gamma^T_{n,j}T_{n,j} - \sum_j^J\gamma^r_{n,j}\frac{w_nL^p_{n,j}}{T_{n,j}} + u_{c_1}x_1\sum_j^J\beta_j\pi_{1n,j} - \theta u_{c_1}\sum_j^J\sum_{m\neq1}^N\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\pi_{m1,j}\pi_{mn,j}x_m$$
$$-\theta\sum_j^J\sum_{m\neq1}\sum_{i\neq1}^N\gamma^L_{i,j}\beta_j\pi_{mi,j}\pi_{mn,j}x_m + \theta\sum_j^J\gamma^L_{n,j}\beta_j\sum_{m\neq1}\pi_{mn,j}x_m$$
$$-\sum_{i\neq1}^N\left(\sum_k^J\gamma^r_{i,k}\rho(1-\delta)(1-\sigma)M_{i,k}\right)\sum_j\beta_j\pi_{in,j} + \sum_j\theta T_{n,j}\sum_{i\neq1}^N\sum_k^J\gamma^r_{i,k}\rho(1-\delta)\frac{\partial G_{i,k}}{\partial T_{n,j}}x_i^\sigma P_i^{1-\sigma}, \tag{A.89}$$

6. Combine equation (A.87) and (A.89)

$$\gamma^r_{n,j}\frac{1}{1+\theta}\frac{x_n}{T_{n,j}} - \gamma^L_{n,j}x_n - \mu_n\sum_j^JL^p_{n,j} + \sum_j^J\gamma^L_{n,j}\frac{1+\theta}{\theta}w_nL^p_{n,j} + \theta(1-\rho(1-\delta))\sum_j^J\gamma^T_{n,j}T_{n,j} + \sum_j^J\gamma^r_{n,j}\frac{w_nL^p_{n,j}}{T_{n,j}}$$
$$+\sum_j^J\gamma^r_{n,j}\left[\rho(1-\delta)M_{n,j}\right] - \sum_j\theta T_{n,j}\sum_{i\neq1}^N\sum_k^J\gamma^r_{i,k}\rho(1-\delta)\frac{\partial G_{i,k}}{\partial T_{n,j}}x_i^\sigma P_i^{1-\sigma} = 0. \tag{A.90}$$

According to FOC for $L^r_{n,j}$ for $n > 1$ (A.71), we have $\gamma^T_{n,j} = \mu_n/\alpha_{n,j}$. In addition, at the steady

state $T_{n,j} = \frac{\alpha_{n,j} L^r_{n,j}}{\delta} = \frac{\alpha_{n,j}}{\delta} \bar{r} L_{n,j}$. Hence, equation (A.90) becomes

$$\gamma^r_{n,j} \frac{1}{1+\theta} \frac{x_n}{T_{n,j}} - \gamma^L_{n,j} x_n - \mu_n (1-\bar{r}) L_n + \sum_j^J \gamma^L_{n,j} \frac{1+\theta}{\theta} w_n L^p_{n,j} + \theta(1-\rho(1-\delta)) \frac{1}{\delta} \bar{r} \mu_n L_n$$
$$+ \sum_j^J \gamma^r_{n,j} \frac{w_n L^p_{n,j}}{T_{n,j}} + \sum_j^J \gamma^r_{n,j} \left[\rho(1-\delta) M_{n,j}\right] - \sum_j \theta T_{n,j} \sum_{i \neq 1}^N \sum_k^J \gamma^r_{i,k} \rho(1-\delta) \frac{\partial G_{i,k}}{\partial T_{n,j}} x_i^\sigma P_i^{1-\sigma} = 0.$$

Using $\bar{r} = \frac{\delta}{\theta(1-\rho(1-\delta))+\delta}$ proved in Proposition 2 and definition of $G_{i,k}$ from (A.67), we have

$$\gamma^r_{n,j} \frac{1}{1+\theta} \frac{x_n}{T_{n,j}} - \gamma^L_{n,j} x_n + \sum_j^J \gamma^L_{n,j} \frac{1+\theta}{\theta} w_n L^p_{n,j} + \sum_j^J \gamma^r_{n,j} \frac{w_n \delta (1-\bar{r})}{\alpha_{n,j} \bar{r}} + \sum_j^J \gamma^r_{n,j} \left[\rho(1-\delta) \frac{w_n}{\alpha_{n,j}}\right]$$
$$- \sum_j \theta T_{n,j} \sum_{i \neq 1}^N \sum_k^J \gamma^r_{i,k} \rho(1-\delta) \frac{\partial (x_i^{'-\sigma} P_i^{'\sigma-1} w_i')}{\partial T_{n,j}} \frac{1}{\alpha_{i,k}} x_i^\sigma P_i^{1-\sigma} = 0. \tag{A.91}$$

7. Multiply equation (A.91) with $\frac{\frac{1+\theta}{\theta} w_n L^p_{n,j}}{x_n}$ on both sides and sum over j

$$\sum_j^J \gamma^r_{n,j} \frac{1}{\theta} \frac{w_n L^p_{n,j}}{T_{n,j}} + \sum_j^J \gamma^r_{n,j} \frac{w_n \delta (1-\bar{r})}{\alpha_{n,j} \bar{r}} + \sum_j^J \gamma^r_{n,j} \rho(1-\delta) \frac{w_n}{\alpha_{n,j}}$$
$$- \sum_j \theta T_{n,j} \sum_{i \neq 1}^N \sum_k^J \frac{\gamma^r_{i,k}}{\alpha_{i,k}} \rho(1-\delta) \frac{\partial (x_i^{'-\sigma} P_i^{'\sigma-1} w_i')}{\partial T_{n,j}} x_i^\sigma P_i^{1-\sigma} = 0$$

which implies

$$\sum_j^J \gamma^r_{n,j} \frac{w_n (1-\bar{r}) L_{n,j}}{\theta \alpha_{n,j} \bar{r} L_{n,j} / \delta} + \sum_j^J \gamma^r_{n,j} \frac{w_n \delta (1-\bar{r})}{\alpha_{n,j} \bar{r}} + \sum_j^J \gamma^r_{n,j} \rho(1-\delta) \frac{w_n}{\alpha_{n,j}}$$
$$- \sum_j \theta T_{n,j} \sum_{i \neq 1}^N \sum_k^J \frac{\gamma^r_{i,k}}{\alpha_{i,k}} \rho(1-\delta) \frac{\partial (x_i^{'-\sigma} P_i^{'\sigma-1} w_i')}{\partial T_{n,j}} x_i^\sigma P_i^{1-\sigma} = 0.$$

combining terms, we have

$$\left(\sum_k^J \frac{\gamma^r_{n,k}}{\alpha_{n,k}}\right) (1 + \theta(1-\rho(1-\delta))) w_n - \rho(1-\delta) \sum_j \theta T_{n,j} \sum_{i \neq 1}^N \left(\sum_k^J \frac{\gamma^r_{i,k}}{\alpha_{i,k}}\right) \frac{\partial (x_i^{'-\sigma} P_i^{'\sigma-1} w_i')}{\partial T_{n,j}} x_i^\sigma P_i^{1-\sigma} = 0.$$

Zero vector is a solution of the homogenous system of linear equations. For every $n \neq 1$, $\sum_k^J \frac{\gamma^r_{n,k}}{\alpha_{n,k}} = 0$.

H Proof for Proposition 5 and 6: optimal dynamic policies one-sector two-country

Proof for Proposition 5 Ramsey optimal policy in one-sector two-country case.

The Euler equation, A.48 in the one sector case becomes:

$$\frac{w_1}{\alpha_1} = \frac{1}{\theta}\frac{w_1 L_1^p}{T_1} + \rho(1-\delta)E\left[\frac{u'_{c_1} w'_1}{u_{c_1}\alpha_1}\right].$$

This is the same as the private innovation equation.

Lerner symmetry holds; we normalize the tariff to zero. The optimal Ramsey export tax, Eq.A.45 becomes:

$$1+\tau^{x,R} = \frac{1+\theta\pi_{22}}{\theta\pi_{22} + (\sigma-1)\left(\gamma_2^v \frac{w_2}{\alpha_2} - \Gamma_2^v \frac{w_2}{\theta}\frac{L_2^p}{T_2}\right)\frac{u_{c_2}}{P_2}/(u_{c_1}x_2)}.$$

Furthermore, According to Appendix F.3, at the SS, $\sum_j^J \frac{\gamma_{n,j}^v}{\alpha_{n,j}} = 0$. This implies under one sector $\gamma_2^v = 0$ and $\Gamma_2^v = 0$ at the SS. □

Proof for Proposition 6 Markov optimal policy under one-sector two-country case.

Home's optimal condition for innovation (A.81) becomes

$$\frac{w_1}{\alpha_1} = \frac{1}{\theta}\frac{w_1 L_1^p}{T_1} + \tilde{\rho}_1 v_1 + \underbrace{\frac{\gamma_2^r}{\alpha_2}\frac{\theta}{1+\theta}\frac{1}{u_{c_1}}\frac{\rho(1-\delta)}{u_{c_2}/P_2}\frac{\partial\left(u'_{c_2}\frac{w'_2}{P'_2}\right)}{\partial T_1}}_{\text{innovation wedge}},$$

where we have substituted $G_2 = (x'_2)^{-\sigma}(P'_2)^{\sigma-1}w'_2/\alpha_2$ with $u'_{c_2}\frac{w'_2}{\alpha_2 P'_2}$. We also used the definition of $\tilde{\rho}_1$ and $v_1 = w_1/\alpha_1$. This proved equation (32) in Proposition 6.

The Lerner symmetry holds; we normalize the home tariff to be zero. Using the tariff formula (A.80) and $u_{c_1} = \gamma_x - \gamma_1^L$, we can back out the multiplier $\gamma_2^L = \gamma_x - u_{c_1}(1+\tau_{n,j}^m) = 0$. Using these conditions and the assumption of one sector, equation (A.79) becomes:

$$1+\tau^{x,M} = \frac{1+\theta\pi_{22}}{\theta\pi_{22} + \gamma_2^r(\sigma-1)\rho(1-\delta)\frac{u'_{c_2}/P'_2}{u_{c_2}/P_2}\frac{w'_2}{\alpha_2}\frac{1}{u_{c_1}x_2}},$$

where we used the condition $\pi_{22} = 1-\pi_{21}$ and definition of M_n from (A.67) under one sector. This proved equation (31) in Proposition 6.

According to Appendix G.3, $\sum_j^J \frac{\gamma_{n,j}^r}{\alpha_{n,j}} = 0$. This implies under one sector $\gamma_2^r = 0$ at the SS. □

I Proof for Proposition 7

According to Appendix F.3, in Ramsey problem at the steady state, $\sum_j^J \frac{\gamma_{n,j}^v}{\alpha_{n,j}} = 0$, then

$$\sum_k \Gamma_{n,k}^v \frac{L_{n,k}^p}{\theta T_{n,k}} = \sum_k \Gamma_{n,k}^v \frac{L_{n,k}^p}{\theta \frac{\alpha_{n,k} L_{n,k}^r}{\delta}} = \sum_k \frac{\Gamma_{n,k}^v}{\alpha_{n,k}} \frac{L_{n,k}^p}{\theta \frac{L_{n,k}^r}{\delta}} = \sum_k \frac{\Gamma_{n,k}^v}{\alpha_{n,k}} \frac{(1-r)L_{n,k}}{\theta \frac{rL_{n,k}}{\delta}} = 0$$

Hence at steady state, these items disappear in export tax Eq.A.45. Furthermore, according to Appendix G.3, in Markov problem, $\sum_j^J \frac{\gamma_{n,j}^r}{\alpha_{n,j}} = 0$ at the steady state. Then Eq.A.45 shows Ramsey export taxes have the same formula as in the Markov A.79 at the SS.

Ramsey tariff Eq.A.43 also has the same formula as in the Markov (A.80). As at the steady state, $\sum_j^J \frac{\gamma_{n,j}^v}{\alpha_{n,j}} = 0$, the Ramsey FOCs of $L_{n,j}^p$, w_n, and $T_{n,j}$ (A.49, A.50, A.51) are also the same as Markov, hence $\gamma_{n,j}^L$ are the same and optimal tariffs are the same.

Eq.A.48 in Section F.2 shows Ramsey doesn't use any industry policy. For Markov innovation policy, we have proved that at the steady state $\sum_j \gamma_{n,j}^r / \alpha_{n,j} = 0$ in Appendix G.3. As a result, $(\sum_k \gamma_{n,k}^r \rho(1-\delta)(\sigma-1)M_{n,k})x_n^{-1} = 0$ and $\sum_{n\neq 1}^N \sum_k^J \gamma_{n,k}^r \rho(1-\delta)\frac{\partial G_{n,k}}{\partial T_{1,j}} x_n^\sigma P_n^{1-\sigma} = 0$. The optimal innovation condition A.81 can be simplified as

$$\frac{w_1}{\alpha_{1,j}} = \frac{1}{\theta}\frac{w_1 L_{1,j}^p}{T_{1,j}} + \rho\,(1-\delta)\left[\frac{u'_{c_1}}{u_{c_1}}\frac{w'_1}{\alpha_{1,j}}\right],$$

which is the same as private free entry condition for Home researcher. Hence, there are no distortions on Home innovation.

With no domestic externalities, in the steady state, Markov and Ramsey governments choose the same policies and have the same allocations.

J Proof for Proposition 8

Section I already proved optimal policies do not distort domestic innovation (zero innovation taxes) at the steady state with multiple countries.

In this section, we prove the following hold

$$\sum_{n\neq 1}(\tau_{n,j}^{m}+\xi_n)\frac{1+\theta}{\theta}w_n L_{n,j}^{p}=\frac{1+\theta}{\theta}w_1 L_{1,j}^{p}-\beta_j x_1 \tag{A.92}$$

$$\frac{\tau_{n,j}^{m}}{1+\tau_{n,j}^{m}}\frac{\beta_j\pi_{1n,j}x_1}{\frac{1+\theta}{\theta}w_n L_{n,j}^{p}}=\delta_n+\frac{\sum_{m\neq 1}\frac{\tau_{m,j}^{x}}{1+\tau_{m,j}^{x}}\beta_j\pi_{mn,j}x_m}{\frac{1+\theta}{\theta}w_n L_{n,j}^{p}} \tag{A.93}$$

Proof:

1. Proof of Eq.A.92. We use the FOCs for $L_{n,j}^{p}$ and $T_{n,j}$ to derive the optimal policy for tariffs. We also need to use other optimal conditions, e.g., export tax $\tau_{n,j}^{x}$, to simplify the formula.

   Sum over $m\neq 1$ of the FOCs on $\tau_{n,j}^{x}$

$$\sum_{m\neq 1}\sum_{i\neq 1}^{N}\gamma_{i,j}^{L}\beta_j\left[\theta\pi_{mi,j}\pi_{m1,j}x_m\right]+u_{c_1}\sum_{m\neq 1}\beta_j\frac{1}{(1+\tau_{m,j}^{x})}\pi_{m1,j}x_m-u_{c_1}\theta\sum_{m\neq 1}\sum_{n\neq 1}\beta_j\frac{\tau_{m,j}^{x}}{1+\tau_{m,j}^{x}}\pi_{m1,j}\pi_{mn,j}x_m=0$$

   Sum over $n\neq 1$ of the FOCs on $T_{n,j}$

$$\sum_{n\neq 1}\theta(1-\rho(1-\delta))\gamma_{n,j}^{T}T_{n,j}=-\sum_{n\neq 1}\gamma_{n,j}^{r}\frac{w_n L_{n,j}^{p}}{T_{n,j}}+\sum_{n\neq 1}u_{c_1}x_1\beta_j\pi_{1n,j}-\theta u_{c_1}\sum_{n\neq 1}\sum_{m\neq 1}^{N}\beta_j\frac{\tau_{m,j}^{x}}{1+\tau_{m,j}^{x}}\pi_{m1,j}\pi_{mn,j}x_m$$
$$-\theta\sum_{n\neq 1}\sum_{m\neq 1}\sum_{i\neq 1}^{N}\gamma_{i,j}^{L}\beta_j\pi_{mi,j}\pi_{mn,j}x_m+\theta\sum_{n\neq 1}\gamma_{n,j}^{L}\beta_j\sum_{m\neq 1}\pi_{mn,j}x_m$$

   Combine these two:

$$\sum_{n\neq 1}\theta(1-\rho(1-\delta))\gamma_{n,j}^{T}T_{n,j}=-\sum_{n\neq 1}\gamma_{n,j}^{r}\frac{w_n L_{n,j}^{p}}{T_{n,j}}+\sum_{n\neq 1}u_{c_1}x_1\beta_j\pi_{1n,j}-u_{c_1}\sum_{m\neq 1}\beta_j\frac{1}{(1+\tau_{m,j}^{x})}\pi_{m1,j}x_m$$

   which is

$$\begin{aligned}\sum_{n\neq 1}\theta(1-\rho(1-\delta))\gamma_{n,j}^{T}T_{n,j}&=-\sum_{n\neq 1}\frac{\gamma_{n,j}^{r}}{T_{n,j}}w_n L_{n,j}^{p}+u_{c_1}\left[\sum_{n\neq 1}\beta_j\pi_{1n,j}x_1-\sum_{m\neq 1}\beta_j\frac{1}{(1+\tau_{m,j}^{x})}\pi_{m1,j}x_m\right]\\&=-\sum_{n\neq 1}\frac{\gamma_{n,j}^{r}}{T_{n,j}}w_n L_{n,j}^{p}-u_{c_1}(\frac{1+\theta}{\theta}w_1 L_{1,j}^{p}-\beta_j x_1)\end{aligned}$$

At the steady state $\gamma_{n,j}^T \alpha_{n,j} - \mu_n = 0$,

$$\frac{\theta(1-\rho(1-\delta))}{\delta} \sum_{n\neq 1} \mu_n L_{n,j}^r = -\sum_{n\neq 1} \frac{\gamma_{n,j}^r}{T_{n,j}} w_n L_{n,j}^p - u_{c_1} \underbrace{(\frac{1+\theta}{\theta} w_1 L_{1,j}^p - \beta_j x_1)}_{\text{Home net export of sector } j} \tag{A.94}$$

Recall the FOC for $L_{n,j}^p$, sum over country n

$$-\gamma_{n,j}^r \frac{1}{\theta} \frac{w_n}{T_{n,j}} = -\gamma_{n,j}^L \frac{1+\theta}{\theta} w_n + \mu_n - L_{1,j}^p \sum_{i\neq 1}^{N} \sum_{k}^{J} \gamma_{i,k}^L \beta_k \pi_{ni} \frac{1+\theta}{\theta} w_n - u_{c_1} \sum_{k}^{J} \beta_k \frac{\tau_{n,k}^x}{1+\tau_{n,k}^x} \pi_{n1,k} \frac{1+\theta}{\theta} w_n$$

and substitute back to Eq.A.94, and use $\tau_{n,j}^m = -\frac{\gamma_{n,j}^L}{u_{c_1}}$,

$$\frac{\theta(1-\rho(1-\delta))}{\delta} \sum_{n\neq 1} \mu_n L_{n,j}^r - \sum_{n\neq 1} \mu_n L_{n,j}^p + \sum_{n\neq 1} \sum_{i\neq 1}^{N} \left(\sum_{k}^{J} \gamma_{i,k}^L \beta_k \pi_{ni,k} \right) \frac{1+\theta}{\theta} w_n L_{n,j}^p +$$
$$u_{c_1} \sum_{n\neq 1} \left(\sum_{k}^{J} \beta_k \frac{\tau_{n,k}^x}{1+\tau_{n,k}^x} \pi_{n1,k} \right) \frac{1+\theta}{\theta} w_n L_{n,j}^p = u_{c_1} \sum_{n\neq 1} \tau_{n,j}^m \frac{1+\theta}{\theta} w_n L_{n,j}^p - u_{c_1} (\frac{1+\theta}{\theta} w_1 L_{1,j}^p - \beta_j x_1)$$

reorganize

$$\sum_{n\neq 1} \tau_{n,j}^m \frac{1+\theta}{\theta} w_n L_{n,j}^p + \sum_{n\neq 1} \sum_{i\neq 1}^{N} \left(\sum_{k}^{J} \tau_{i,k}^m \beta_k \pi_{ni,k} \right) \frac{1+\theta}{\theta} w_n L_{n,j}^p - \sum_{n\neq 1} \left(\sum_{k}^{J} \beta_k \frac{\tau_{n,k}^x}{1+\tau_{n,k}^x} \pi_{n1,k} \right) \frac{1+\theta}{\theta} w_n L_{n,j}^p$$
$$= \frac{1+\theta}{\theta} w_1 L_{1,j}^p - \beta_j x_1$$

It can be further simplified to

$$\sum_{n\neq 1} (\tau_{n,j}^m + \sum_{i\neq 1}^{N} \sum_{k}^{J} \tau_{i,k}^m \beta_k \pi_{ni,k} - \sum_{k}^{J} \beta_k \frac{\tau_{n,k}^x}{1+\tau_{n,k}^x} \pi_{n1,k}) \frac{1+\theta}{\theta} w_n L_{n,j}^p = \frac{1+\theta}{\theta} w_1 L_{1,j}^p - \beta_j x_1 \tag{A.95}$$

With multiple countries, Home's tariffs are on average higher for sectors with relatively higher net exports. For that sector, the weighted average of Home's tariff across country is higher. Different from the two-country case, now tariffs on one foreign country would affect other foreign countries' sector profit and incentive to do innovation, hence technology, so the levels of optimal tariffs are jointly determined across country and sector.

In the special cases of two countries or multiple countries but only Home and one foreign country (country 2) produce sector j and j' goods, we don't need to consider the supply of other foreign countries that have been affected by the tariff imposed on country 2. The left-

hand side of the equation (A.95) only has country 2. Then again we get

$$\tau^m_{2,j} - \tau^m_{2,j'} = \frac{1}{\theta}\left(\frac{\frac{1+\theta}{\theta}w_1 L^p_{1,j} - \beta_j x_1}{\frac{1+\theta}{\theta}w_2 L^p_{2,j}} - \frac{\frac{1+\theta}{\theta}w_1 L^p_{1,j'} - \beta_{j'} x_1}{\frac{1+\theta}{\theta}w_2 L^p_{2,j'}}\right)$$

2. Now we prove Eq.A.93

Go back to the FOC for $T_{n,j}$

$$\begin{aligned}\theta(1-\rho(1-\delta))\gamma^T_{n,j}T_{n,j} = &-\gamma^r_{n,j}\frac{w_n L^p_{n,j}}{T_{n,j}} + u_{c_1}x_1\beta_j\pi_{1n,j} - \theta u_{c_1}\sum_{m\neq 1}^{N}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\pi_{m1,j}\pi_{mn,j}x_m \\ &- \theta\sum_{m\neq 1}\sum_{i\neq 1}^{N}\gamma^L_{i,j}\beta_j\pi_{mi,j}\pi_{mn,j}x_m + \theta\gamma^L_{n,j}\beta_j\sum_{m\neq 1}\pi_{mn,j}x_m\end{aligned}$$

using the FOC for $\tau^x_{n,j}$

$$\theta\sum_{i\neq 1}^{N}\gamma^L_{i,j}\beta_j\pi_{ni,j} = -u_{c_1}\beta_j\frac{1}{(1+\tau^x_{n,j})} + u_{c_1}\beta_j\frac{\tau^x_{n,j}}{1+\tau^x_{n,j}}\theta(1-\pi_{n1,j})$$

the FOC for $T_{n,j}$ becomes

$$\begin{aligned}&\theta(1-\rho(1-\delta))\gamma^T_{n,j}T_{n,j} = -\gamma^r_{n,j}\frac{w_n L^p_{n,j}}{T_{n,j}} + u_{c_1}x_1\beta_j\pi_{1n,j} - \theta u_{c_1}\sum_{m\neq 1}^{N}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\pi_{m1,j}\pi_{mn,j}x_m \\ &+\sum_{m\neq 1}\pi_{mn,j}x_m\left(u_{c_1}\beta_j\frac{1}{(1+\tau^x_{m,j})}\right) - \sum_{m\neq 1}\pi_{mn,j}x_m\left(u_{c_1}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\theta(1-\pi_{m1,j})\right) + \theta\gamma^L_{n,j}\beta_j\sum_{m\neq 1}\pi_{mn,j}x_m\end{aligned}$$

Combine items,

$$\begin{aligned}&\theta(1-\rho(1-\delta))\frac{1}{\delta}\mu_n L^r_{n,j} = -\gamma^r_{n,j}\frac{w_n L^p_{n,j}}{T_{n,j}} + u_{c_1}x_1\beta_j\pi_{1n,j} \\ &+u_{c_1}\sum_{m\neq 1}\beta_j\frac{1}{(1+\tau^x_{m,j})}\pi_{mn,j}x_m - \theta u_{c_1}\sum_{m\neq 1}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\pi_{mn,j}x_m + \theta\gamma^L_{n,j}\beta_j\sum_{m\neq 1}\pi_{mn,j}x_m\end{aligned}$$

replace the last term with $\left(\frac{1+\theta}{\theta}w_n L^p_{n,j} - \beta_j\frac{1}{1+\tau^m_{n,j}}\pi_{1n,j}x_1\right)$

$$\begin{aligned}&\theta(1-\rho(1-\delta))\frac{1}{\delta}\mu_n L^r_{n,j} = -\gamma^r_{n,j}\frac{w_n L^p_{n,j}}{T_{n,j}} + u_{c_1}x_1\beta_j\pi_{1n,j} \\ &+u_{c_1}\sum_{m\neq 1}\beta_j\frac{1}{(1+\tau^x_{m,j})}\pi_{mn,j}x_m - \theta u_{c_1}\sum_{m\neq 1}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\pi_{mn,j}x_m + \gamma^L_{n,j}\left((1+\theta)w_n L^p_{n,j} - \theta\beta_j\frac{1}{1+\tau^m_{n,j}}\pi_{1n,j}x_1\right)\end{aligned}$$

Using the following equations

$$\gamma^L_{n,j} = \gamma^r_{n,j}\frac{1}{1+\theta}\frac{1}{T_{n,j}} + \sum_j^J \gamma^L_{n,j}\frac{\frac{1+\theta}{\theta}w_n L^p_{n,j}}{x_n}, \quad \gamma^L_{n,j} = -\tau^m_{n,j}u_{c_1},$$

we get

$$\theta(1-\rho(1-\delta))\frac{1}{\delta}\mu_n L^r_{n,j} = (1+\theta)w_n L^p_{n,j}\left[\sum_j^J \gamma^L_{n,j}\frac{\frac{1+\theta}{\theta}w_n L^p_{n,j}}{x_n}\right] + u_{c_1}x_1\beta_j\pi_{1n,j}$$
$$+ u_{c_1}\sum_{m\neq 1}\beta_j\frac{1}{(1+\tau^x_{m,j})}\pi_{mn,j}x_m - \theta u_{c_1}\sum_{m\neq 1}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\pi_{mn,j}x_m + \theta\tau^m_{n,j}u_{c_1}\beta_j\frac{1}{1+\tau^m_{n,j}}\pi_{1n,j}x_1$$

Using the FOC for τ_x, we get

$$\theta(1-\rho(1-\delta))\frac{1}{\delta}\mu_n L^r_{n,j} = (1+\theta)w_n L^p_{n,j}\left[\sum_j^J \gamma^L_{n,j}\frac{\frac{1+\theta}{\theta}w_n L^p_{n,j}}{x_n}\right]$$
$$+ (1+\theta)u_{c_1}\beta_j\frac{\tau^m_{n,j}}{1+\tau^m_{n,j}}\pi_{1n,j}x_1 + u_{c_1}\frac{1+\theta}{\theta}w_n L^p_{n,j} - (1+\theta)u_{c_1}\sum_{m\neq 1}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\pi_{mn,j}x_m$$

or

$$\frac{\tau^m_{n,j}}{1+\tau^m_{n,j}}\frac{\beta_j\pi_{1n,j}x_1}{\frac{1+\theta}{\theta}w_n L^p_{n,j}} = \delta_n + \frac{\sum_{m\neq 1}\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\pi_{mn,j}x_m}{\frac{1+\theta}{\theta}w_n L^p_{n,j}}$$

where

$$\delta_n = \frac{1}{1+\theta}\{-\theta\sum_j^J \gamma^L_{n,j}\frac{\frac{1+\theta}{\theta}w_n L^p_{n,j}}{u_{c_1}x_n} + \mu_n\frac{\theta(1-\rho(1-\delta))}{\delta}\frac{\bar{r}}{1-\bar{r}}\frac{1}{u_{c_1}w_n}\frac{\theta}{1+\theta} - 1\}$$
$$= \frac{1}{1+\theta}\{-\theta\sum_j^J \gamma^L_{n,j}\frac{\frac{1+\theta}{\theta}w_n L^p_{n,j}}{u_{c_1}x_n} + \frac{\mu_n L^p_n}{u_{c_1}x_n} - 1\} = \frac{1}{1+\theta}\{\theta\sum_j^J \tau^m_{n,j}\frac{\frac{1+\theta}{\theta}w_n L^p_{n,j}}{x_n} + \frac{\mu_n L^p_n}{u_{c_1}x_n} - 1\}.$$

K Proof for Corollary 1

According to our characterization of Markov problem in equation (A.79) in section G of this appendix, the optimal export tax $\tau^x_{n,j}$ satisifes

$$1+\tau^x_{n,j} = \frac{\left(\gamma_x - \gamma^L_1\right)\left[1+\theta(1-\pi_{n1,j})\right]}{\gamma_x\theta(1-\pi_{n1,j}) - \theta\left(\sum_{i\neq 1}^N \gamma^L_{i,j}\pi_{ni,j}\right) + \left(\sum_k \gamma^r_{n,k}\rho(1-\delta)(\sigma-1)M_{n,k}\right)x_n^{-1}},$$

and the optimal $\tau^m_{n,j}$ tariff satisfies

$$1+\tau^m_{n,j}=\frac{\gamma_x-\gamma^L_{n,j}}{\gamma_x-\gamma^L_1}.$$

Given $\sum_j^J \frac{\gamma^r_{n,j}}{\alpha_{n,j}}=0$ at SS,

$$1+\tau^x_{n,j}=\frac{\left(\gamma_x-\gamma^L_1\right)\left[1+\theta(1-\pi_{n1,j})\right]}{\gamma_x\theta(1-\pi_{n1,j})-\theta\left(\sum_{i\neq 1}^N\gamma^L_{i,j}\pi_{ni,j}\right)}=\frac{1+\theta(1-\pi_{n1,j})}{\theta\sum_{i\neq 1}(1+\tau^m_{i,j})\pi_{ni,j}}.$$

Now we prove the optimal Markov tariffs across sectors at the SS in the baseline model with two countries.

Using the FOC for $L^p_{n,j}$ (Eq.A.72), in a two-country case, for any two sectors,

$$\frac{\gamma^r_{2,j}}{T_{2,j}(1+\theta)}-\gamma^L_{2,j}=\frac{\gamma^r_{2,j'}}{T_{2,j'}(1+\theta)}-\gamma^L_{2,j'} \tag{A.96}$$

Plug in $\sum_j \frac{\gamma^r_{n,j}}{\alpha_{n,j}}=0$ to the FOC for $T_{n,j}$ at SS (Eq.A.88)

$$\begin{aligned}\theta(1-\rho(1-\delta))\gamma^T_{n,j}T_{n,j}=&-\gamma^r_{n,j}\frac{w_nL^p_{n,j}}{T_{n,j}}+u_{c_1}x_1\beta_j\pi_{1n,j}-\theta u_{c_1}\sum_{m\neq 1}^N\beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\pi_{m1,j}\pi_{mn,j}x_m\\&-\theta\sum_{m\neq 1}\sum_{i\neq 1}^N\gamma^L_{i,j}\beta_j\pi_{mi,j}\pi_{mn,j}x_m+\theta\gamma^L_{n,j}\beta_j\sum_{m\neq 1}\pi_{mn,j}x_m\end{aligned}$$

In a two-country case, this equation becomes

$$\begin{aligned}\theta(1-\rho(1-\delta))\gamma^T_{2,j}T_{2,j}=&\,u_{c_1}x_1\beta_j\pi_{12,j}-\theta u_{c_1}\beta_j\frac{\tau^x_{2,j}}{1+\tau^x_{2,j}}\pi_{21,j}\pi^j_{22}x_2\\&-\theta\gamma^L_{2,j}\beta_j\pi_{22,j}\pi_{22,j}x_2+\theta\gamma^L_{2,j}\beta_j\pi_{22,j}x_2-\gamma^r_{2,j}\frac{w_2L^p_{2,j}}{T_{2,j}}\end{aligned} \tag{A.97}$$

From the FOC for $L^r_{n,j}$

$$\mu_n=\gamma^T_{n,j}\alpha_{n,j} \tag{A.98}$$

Combine A.97 and A.98, for any two sector j and j',

$$\frac{\alpha_{2,j}}{T_{2,j}}\left\{u_{c_1}x_1\beta_j\pi_{12,j}-\theta u_{c_1}\beta_j\frac{\tau^x_{2,j}}{1+\tau^x_{2,j}}\pi_{21,j}\pi_{22,j}x_2-\theta\gamma^L_{2,j}\beta_j\pi_{22,j}\pi_{22,j}x_2+\theta\gamma^L_{2,j}\beta_j\pi_{22,j}x_2-\gamma^r_{2,j}\frac{w_2L^p_{2,j}}{T_{2,j}}\right\} \tag{A.99}$$

$$=\frac{\alpha_{2,j'}}{T_{2,j'}}\left\{u_{c_1}x_1\beta_{j'}\pi_{12,j'}-\theta u_{c_1}\beta_{j'}\frac{\tau^x_{2,j'}}{1+\tau^x_{2,j'}}\pi_{21,j'}\pi_{22,j'}x_2-\theta\gamma^L_{2,j'}\beta_{j'}\pi_{22,j'}\pi_{22,j'}x_2+\theta\gamma^L_{2,j'}\beta_{j'}\pi_{22,j'}x_2-\gamma_{2,j'}\frac{w_2L^p_{2,j'}}{T_{2,j'}}\right\}$$

In the two-country case,

$$1+\tau^m_{2,j}=\frac{\gamma_x-\gamma^L_{2,j}}{u_{c_1}}, \tag{A.100}$$

$$1+\tau^x_{2,j}=\frac{\left(\gamma_x-\gamma^L_1\right)\left[1+\theta(1-\pi_{21,j})\right]}{\gamma_x\theta(1-\pi_{21,j})-\theta\gamma^L_{2,j}\pi_{22,j}}=\frac{1+\theta\pi_{22,j}}{\theta\pi_{22,j}(1+\tau^m_{2,j})}$$

Equation (A.99) can be further simplified to

$$\frac{\alpha_{2,j}}{T_{2,j}}\left\{u_{c_1}x_1\beta_j\pi_{12,j}-u_{c_1}\frac{1}{1+\tau^x_{2,j}}\beta_j\pi_{21,j}x_2-\gamma^r_{2,j}\frac{w_2L^p_{2,j}}{T_{2,j}}\right\}=\frac{\alpha_{2,j'}}{T_{2,j'}}\left\{u_{c_1}x_1\beta_{j'}\pi_{12,j'}-u_{c_1}\frac{1}{1+\tau^x_{2,j'}}\beta'_j\pi_{21,j'}x_2-\gamma^r_{2,j'}\frac{w_2L^p_{2,j'}}{T_{2,j'}}\right.$$

Combine A.96 and A.100, and goods market clearing conditions with the above equation, we get

$$\tau^m_{2,j}-\tau^m_{2,j'}=\frac{1}{\theta}\left(\frac{\frac{1+\theta}{\theta}w_1L^p_{1,j}-\beta_jx_1}{\frac{1+\theta}{\theta}w_2L^p_{2,j}}-\frac{\frac{1+\theta}{\theta}w_1L^p_{1,j'}-\beta_{j'}x_1}{\frac{1+\theta}{\theta}w_2L^p_{2,j'}}\right).$$

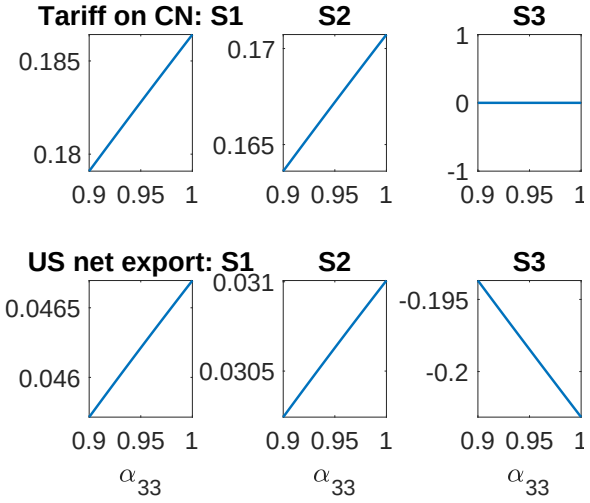
L Numerical example

The interlinking relationships and spillover of tariffs across countries and sectors may appear to be complex but follow the basic principle outlined in the proposition. That is— the within-country tariff ranking has to do with the country's comparative advantage. That strict ranking disappears when we deviate from the special case, but still follows some basic trade patterns—further explored in our quantitative analysis. Still, to gain some intuition, and to investigate predictions of theory, we consider a few numerical examples.

Consider the world with three sectors (S1, S2, and S3) and three economies: the United States (US), China (CN), and the rest of the world (ROW). The innovation efficiencies in the three sectors accord with US $(1, 0.9, 0.9)$, CN $(0.9, 0.9, 1)$, ROW $(0, 0, \alpha_{33})$ with α_{33} varying from 0.9 to 1. Since ROW's innovation efficiency are zero and cannot produce goods from S1 and S2, it imports them from the United States and China. Furthermore, S2 is China's comparative advantage (CA) sector,

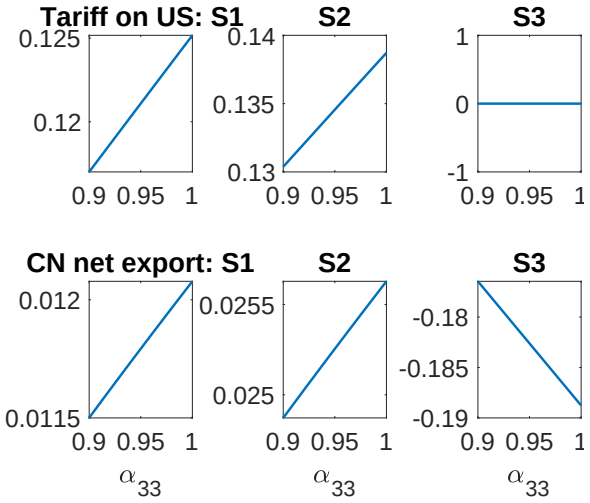
and S1 is U.S.'s comparative advantage sector.

Figure A-1: Optimal Policies for the United States when Global Demand Rises



Note: The upper panel plots U.S. optimal tariff on imports from China in sector 1 ($S1$), sector 2 ($S2$), and sector 3 ($S3$) when sector 3 in ROW becomes more efficient, i.e. higher α_{33}. The lower panel plots the corresponding US's net export in sector 1, 2, and 3.

Figure A-2: Optimal Policies for China when Global Demand Rises



Note: The upper panel plots China's optimal tariff on imports from the U.S. in sectors S1, S2, and S3 when S3 in ROW becomes more efficient, i.e. higher α_{33}. The lower panel plots China's net export in the three sectors.

Consider the case in which ROW becomes larger, with a higher α_{33}. This leads to a surge in the global demand for goods in S1 and S2. Note that the U.S. always imposes tariffs on S1, which is its CA sector. Larger global demand for S1 and S2 boosts exports in these sectors, and thus the U.S. increases tariffs on both sectors. Moreover, the tariff on S2 increases by more as the net export in this sector rises faster—the same reasoning for China.

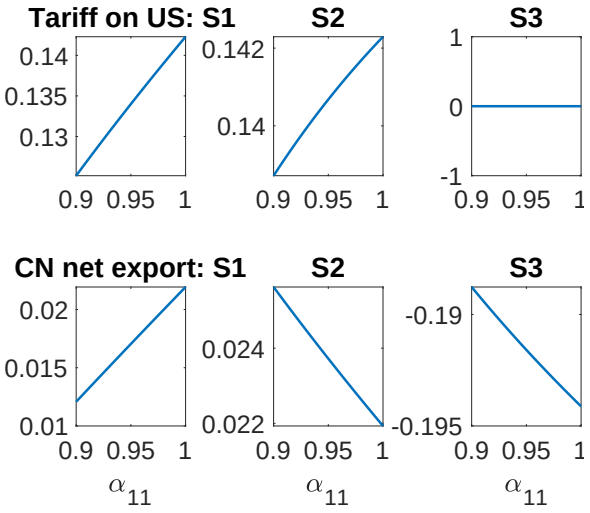
Figure A-1 shows that the U.S. raises tariffs on China's sector 2 goods more than tariffs on sector 1 goods. In addition, the U.S. imposes higher tariffs on ROW's sector 3 goods, aiming to induce Chinese labor to shift from sector 1 and 2 to sector 3. However, the impact of this policy on the rise of Chinese labor in sector 3 is not significant, although it dampens China's labor flowing into sectors 1 and 2. The same optimal policy applies symmetrically to China, which levies a higher tariff on the US goods in sector 1 and 2, but higher in sector 2, and on ROW in sector 3, see Figure A-2. The larger the rise in global demand, the larger the tariffs levied by both China and the U.S. to endogenously improve each country's own technology in that sector.

Next, consider a case where there is an increase in China S1's innovation efficiency α_{11}, which takes on a value between 0.9 and 1. Figure A-3 shows that China levies a relatively higher tariff on the U.S. in S1 to discourage innovation in the U.S. It would also impose a higher tariff on ROW's sector 3 so as to induce U.S. labor to flow towards sector 3. Overall, this results in U.S. labor (and hence innovation) falling in sector 1.

However, from the perspective of the U.S., it will lower tariffs imposed on China's S1 goods. The reason is that China has become more productive in producing these goods, and thus it is more efficient for the U.S. to increase its imports from China in sector 1. Note, however, that this result may change if there are other forces at play—for instance, an event that causes there to be more profits to be reaped in that sector. This can occur if, along with rising Chinese efficiency, there is also a rise in the US efficiency in this sector, or in the importance for sector 1 goods (see Figure A-4), or if overall trade costs have fallen. In this case, the U.S. would want to raise its tariffs on Chinese sector 1 goods to stave off competition. These results show that in our framework, competitive policies arise not from a country becoming better at producing another country's comparative advantage goods (in this case more trade is better), but from the incrementally bigger market that is 'up for grabs' and where competition is made more intense when the other country is a closer rival.

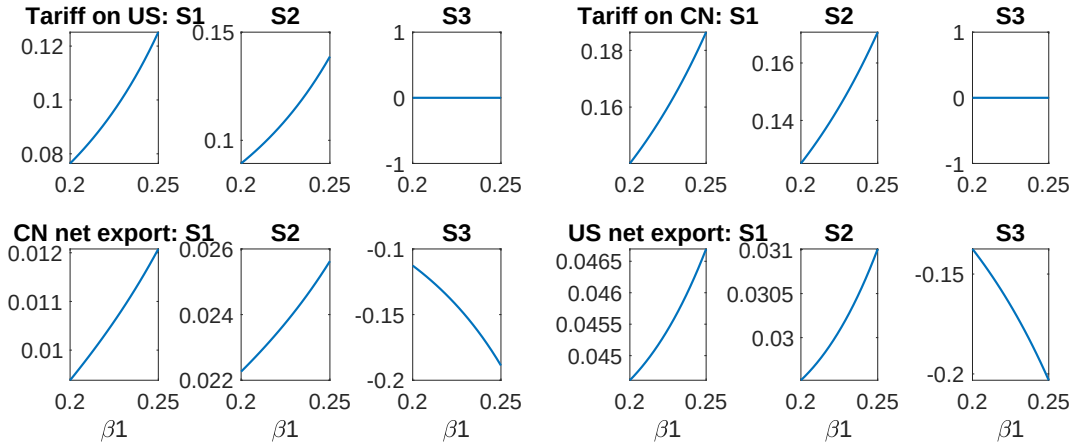
We have shown examples whereby a rise in net exports induces an increase in import tariffs—in the case where T is endogenous and optimal policies across sectors and countries are jointly determined. There are other points to be made here. First, the larger the comparative advantage, the larger the motive to impose tariffs on one another. For instance, as the overall trade costs fall, there is more trade across all countries, and hence there is a larger incentive to undertake policies. Second, for each foreign country, there is no simple relationship between optimal tariffs and Home sector's degree of comparative advantage compare to the Foreign country. As is shown from the example above, China will raise its tariffs on the U.S. in sector 1 more if the global demand for

Figure A-3: Optimal Policies for China when S1 Efficiency Rises



Note: The upper panel plots China's optimal tariff on imports from the U.S. in sector 1 ($S1$), sector 2 ($S2$), and sector 3 ($S3$) when sector 1 in China becomes more efficient, i.e., higher α_{11}. The lower panel plots the corresponding China's net export in sector 1, 2, and 3.

Figure A-4: Optimal Policies for China and US when S1 becomes more important



(a) China's Tariffs (b) US's Tariffs

Note: The upper panel plots the optimal tariff in sector 1 ($S1$), sector 2 ($S2$), and sector 3 ($S3$) when sector 1 becomes more important, i.e., higher β_1. The lower panel plots the corresponding net export in sector 1, 2, and 3.

good 1 and 2 rises—even though sector 1 is not China's comparative advantage sector. By contrast, if overall trade costs fall, then China will increase its tariff on sector 2. The reason is that China will increase its imports from the U.S. in sector 1 and thus lower its tariffs on U.S.'s sector 1 goods. The key barometer here is the change in *net exports* in any given sector: a country will raise tariffs imposed on the other country if its net exports in a sector rises; conversely, it will lower tariffs in sectors where its net exports falls.

M Proof for Proposition 9: optimal Ramsey policies

The way we solve the Ramsey with externality is very similar to the baseline in Appendix F, which lists optimal conditions for Ramsey policy, and Appendix F.3 proved $\sum_j^J \frac{\gamma_{n,j}^r}{\alpha_{n,j}} = 0$ at baseline SS. The prove of $\sum_j^J \frac{\gamma_{n,j}^r}{\alpha_{n,j}(L_{n,j,t}^r)^{\varepsilon-1}} = 0$ with externality is similar. To avoid repetition, we just highlight the differences here.

We use the following FOCs of the Ramsey problem.

FOC over $L_{n,j}^p$

$$\varepsilon\frac{1}{\theta}\frac{w_n}{T_{n,j}}\Gamma_{n,j}^v x_n^{-\sigma}P_n^{\sigma-1} - \gamma_{n,j}^L\frac{1+\theta}{\theta}w_n - \mu_n + \sum_{k=1}^J \gamma_{1,k}^L\beta_k\frac{1}{1+\tau_{n,k}^x}\pi_{n1,k}\frac{1+\theta}{\theta}w_n + \sum_{i\neq 1}^N\sum_k^J \gamma_{i,k}^L\beta_k\pi_{n,i}\frac{1+\theta}{\theta}w_n$$
$$+\gamma_x\sum_k^J\beta_k\frac{\tau_{n,k}^x}{1+\tau_{n,k}^x}\pi_{n1,k}\frac{1+\theta}{\theta}w_n + \sum_k^J\left[\varepsilon\frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}}\Gamma_{n,k}^v - \frac{1}{\alpha_{n,k}}\gamma_{n,k}^v\right]w_n x_n^{-\sigma}P_n^{\sigma-1}(-\sigma)\frac{1+\theta}{\theta}\frac{w_n}{x_n} = 0$$

FOC over w_n

$$-u_{c_1}x_1\sum_j^J\beta_j\pi_{1n,j} + \gamma_1^L\left[\sum_{m\neq 1}^N\sum_j^J\beta_j\frac{1}{1+\tau_{m,j}^x}\theta\pi_{m1,j}\pi_{mn,j}x_m + \sum_j^J\beta_j\frac{1}{1+\tau_{n,j}^x}\pi_{n1,j}x_n\right]$$
$$+\sum_{i\neq\{1\}}^N\sum_j^J\gamma_{i,j}^L\beta_j\left(\sum_{m\neq 1}\theta\pi_{mi,j}\pi_{mn,j}x_m + \pi_{ni,j}x_n\right) - \sum_j^J\gamma_{n,j}^L\left(\theta\beta_j\sum_{m\neq 1}\pi_{mn,j}x_m + \frac{1+\theta}{\theta}w_n L_{n,j}^p\right)$$
$$+\gamma_x\sum_{j=1}^J\beta_j\left(\sum_{m\neq 1}^N\frac{\tau_{m,j}^x}{1+\tau_{m,j}^x}\theta\pi_{m1,j}\pi_{mn,j}x_m + \frac{\tau_{n,j}^x}{1+\tau_{n,j}^x}\pi_{n1,j}x_n\right) + (1-\sigma)\sum_j^J\left[\varepsilon\frac{1}{\theta}\frac{L_{n,j}^p}{T_{n,j}}\Gamma_{n,j}^v - \frac{1}{\alpha_{n,j}}\gamma_{n,j}^v\right]w_n x_n^{-\sigma}P_n^{\sigma-1}$$
$$+\sum_{m\neq 1}^N\sum_j^J\left[\varepsilon\frac{1}{\theta}\frac{L_{m,j}^p}{T_{m,j}}\Gamma_{m,j}^v - \frac{1}{\alpha_m^j}\gamma_{m,j}^v\right]w_m x_m^{-\sigma}P_m^{\sigma-1}(\sigma-1)\left(\sum_k\beta_k\pi_{mn,k}\right) = 0$$

FOC over $T_{n,j}$

$$\gamma_{n,j}^T = \rho E\left[\gamma_{n,j}^{T\prime}(1-\delta)\right] + \frac{1}{T_{n,j}}\{u_{c_1}x_1\frac{\beta_j}{\theta}\pi_{1n,j}$$
$$-\sum_{i=1}^N\gamma_{i,j}^L\beta_j\sum_{m\neq 1}\pi_{mi,j}\pi_{mn,j}x_m + \gamma_{n,j}^L\beta_j\sum_{m\neq 1}\pi_{mn,j}x_m - u_{c_1}\sum_{m\neq 1}^N\beta_j\frac{\tau_{m,j}^x}{1+\tau_{m,j}^x}\pi_{m1,j}\pi_{mn,j}x_m$$
$$-\left[\varepsilon\frac{1}{\theta}\frac{L_{n,j}^p}{T_{n,j}}\Gamma_{n,j}^v\right]w_n x_n^{-\sigma}P_n^{\sigma-1} - \sum_{i\neq 1}\sum_k^J\left[\varepsilon\frac{1}{\theta}\frac{L_{i,k}^p}{T_{i,k}}\Gamma_{i,k}^v - \frac{1}{\alpha_{i,k}}\gamma_{i,k}^v\right]w_i x_i^{-\sigma}P_i^{\sigma-1}(\sigma-1)\frac{\beta_j}{\theta}\pi_{in,j}\}$$

1. At the steady state $\gamma^T_{n,j} = \gamma^T_{n,j'}$, and sum over j the FOC for $T_{n,j}$:

$$
\begin{aligned}
&-\theta(1-\rho(1-\delta))\sum_j^J \gamma^T_{n,j}T_{n,j} + u_{c_1}x_1\sum_j^J \beta_j\pi_{1n,j} - \theta u_{c_1}\sum_j^J\sum_{m\neq 1}^N \beta_j\frac{\tau^x_{m,j}}{1+\tau^x_{m,j}}\pi_{m1,j}\pi_{mn,j}x_m \\
&-\theta\sum_j^J\sum_{m\neq 1}\sum_{i\neq 1}^N \gamma^L_{i,j}\beta_j\pi_{mi,j}\pi_{mn,j}x_m + \theta\sum_j^J\gamma^L_{n,j}\beta_j\sum_{m\neq 1}\pi_{mn,j}x_m \\
&-\sum_j^J\left[\varepsilon\frac{L^p_{n,j}}{T_{n,j}}\Gamma^v_{n,j}\right]w_n x_n^{-\sigma}P_n^{\sigma-1} - \sum_{i\neq 1}\sum_k^J\left[\varepsilon\frac{1}{\theta}\frac{L^p_{i,k}}{T_{i,k}}\Gamma^v_{i,k} - \frac{1}{\alpha_{i,k}}\gamma^v_{i,k}\right]w_i x_i^{-\sigma}P_i^{\sigma-1}(\sigma-1)\sum_j^J\beta_j\pi_{in,j},
\end{aligned}
\tag{A.101}
$$

2. Substract the FOCs for $L^p_{n,j}$ and w_n from the above equation:

$$
\begin{aligned}
&\varepsilon\frac{1}{\theta}\frac{w_nL_{np}}{T_{n,j}}\Gamma^v_{n,j}x_n^{-\sigma}P_n^{\sigma-1} - \gamma^L_{n,j}x_n - \mu_n\sum_j^J L^p_{n,j} + \sum_j^J\gamma^L_{n,j}\frac{1+\theta}{\theta}w_nL^p_{n,j} + \theta(1-\rho(1-\delta))\sum_j^J\gamma^T_{n,j}T_{n,j} \\
&+\sum_j^J\left[\varepsilon\frac{L^p_{n,j}}{T_{n,j}}\Gamma^v_{n,j}\right]w_nx_n^{-\sigma}P_n^{\sigma-1} - \sum_k^J\left[\varepsilon\frac{1}{\theta}\frac{L^p_{n,k}}{T_{n,k}}\Gamma^v_{n,k} - \frac{1}{\alpha_{n,k}}\gamma^v_{n,k}\right]w_nx_n^{-\sigma}P_n^{\sigma-1} = 0
\end{aligned}
\tag{A.102}
$$

3. At the steady state

$$
T_{n,j} = \frac{\alpha_{n,j}(L^r_{n,j})^{\varepsilon-1}L^r_{n,j}}{\delta} = \frac{\alpha_{n,j}}{\delta}(L^r_{n,j})^{\varepsilon}, \quad \mu_n = \varepsilon\gamma^T_{n,j}\alpha_{n,j}(L^r_{n,j})^{\varepsilon-1} + \gamma^v_{n,j}(\varepsilon-1)\frac{w_n}{\alpha_{n,j}(L^r_{n,j})^{\varepsilon}}x_n^{-\sigma}P_n^{\sigma-1},
$$

hence

$$
\begin{aligned}
&\varepsilon\frac{1}{\theta}\frac{w_nL_{np}}{T_{n,j}}\Gamma^v_{n,j}x_n^{-\sigma}P_n^{\sigma-1} - \gamma^L_{n,j}x_n - \mu_n\sum_j^J L^p_{n,j} + \sum_j^J\gamma^L_{n,j}\frac{1+\theta}{\theta}w_nL^p_{n,j} \\
&+\theta(1-\rho(1-\delta))\left(\sum_j\frac{1}{\delta\varepsilon}\mu_nL^r_{n,j} - \sum_j^J\gamma^v_{n,j}\frac{(\varepsilon-1)w_n}{\delta\varepsilon\alpha_{n,j}(L^r_{n,j})^{\varepsilon-1}}x_n^{-\sigma}P_n^{\sigma-1}\right) \\
&+\sum_j^J\left[\varepsilon\frac{L^p_{n,j}}{T_{n,j}}\Gamma^v_{n,j}\right]w_nx_n^{-\sigma}P_n^{\sigma-1} - \sum_k^J\left[\varepsilon\frac{1}{\theta}\frac{L^p_{n,k}}{T_{n,k}}\Gamma^v_{n,k} - \frac{1}{\alpha_{n,k}}\gamma^v_{n,k}\right]w_nx_n^{-\sigma}P_n^{\sigma-1} = 0
\end{aligned}
\tag{A.103}
$$

At the steady state, the research intensity is given by

$$
\bar{r}_j = \frac{\delta\varepsilon}{\theta(1-\rho(1-\delta))+\delta\varepsilon}.
$$

Using this research intensity, $L^p_{n,j} = (1-\bar{r})L_{n,j}$, $L^p_n = (1-\bar{r})L_n$, and $L^r_{n,j} = \bar{r}L_{n,j}$, we can cancel

out the two terms involving μ_n in equation (A.103). Then we have

$$\varepsilon\frac{1}{\theta}\frac{w_n L_n^p}{T_{n,j}}\Gamma_{n,j}^v x_n^{-\sigma}P_n^{\sigma-1}-\gamma_{n,j}^L x_n+\sum_j^J\gamma_{n,j}^L\frac{1+\theta}{\theta}w_n L_{n,j}^p-\theta(1-\rho(1-\delta))(\sum_j^J\gamma_{n,j}^v\frac{(\varepsilon-1)w_n}{\delta\varepsilon\alpha_{n,j}(rL_{n,j})^{\varepsilon-1}})x_n^{-\sigma}P_n^{\sigma-1}$$
$$+\sum_j^J\left[\varepsilon\frac{L_{n,j}^p}{T_{n,j}}\Gamma_{n,j}^v\right]w_n x_n^{-\sigma}P_n^{\sigma-1}-\sum_k^J\left[\varepsilon\frac{1}{\theta}\frac{L_{n,k}^p}{T_{n,k}}\Gamma_{n,k}^v-\frac{1}{\alpha_{n,k}}\gamma_{n,k}^v\right]w_n x_n^{-\sigma}P_n^{\sigma-1}=0.$$

4. Multiply the above equation $\frac{\frac{1+\theta}{\theta}w_n L_{n,j}^p}{x_n}$ on both sides and sum over j

$$-\theta(1-\rho(1-\delta))(\sum_j^J\gamma_{n,j}^v\frac{(\varepsilon-1)w_n}{\delta\varepsilon\alpha_{n,j}(rL_{n,j})^{\varepsilon-1}})x_n^{-\sigma}P_n^{\sigma-1}+\sum_j^J\left[\varepsilon\frac{L_{n,j}^p}{T_{n,j}}\Gamma_{n,j}^v\right]w_n x_n^{-\sigma}P_n^{\sigma-1}+\sum_k^J\frac{1}{\alpha_{n,k}}\gamma_{n,k}^v w_n x_n^{-\sigma}P_n^{\sigma-1}=0.$$

Finally,

$$\sum_k^J\frac{\gamma_{n,k}^v}{\alpha_{n,k}(rL_{n,k})^{\varepsilon-1}}(1+\frac{\theta(1-\rho(1-\delta))}{\delta\varepsilon})w_n x_n^{-\sigma}P_n^{\sigma-1}=0,$$

and the equation only holds when for every $n\neq 1$, $\sum_k^J\frac{\gamma_{n,k}^v}{\alpha_{n,k}(rL_{n,k})^{\varepsilon-1}}=0$,

Furthermore,

$$\sum_k\varepsilon\Gamma_{n,k}^v\frac{w_n L_{n,k}^p}{\theta T_{n,k}}=\sum_k\varepsilon\Gamma_{n,k}^v\frac{w_n L_{n,k}^p}{\theta\frac{\alpha_{n,k}(L_{n,k}^r)^{\varepsilon-1}L_{n,k}^r}{\delta}}=\sum_k\varepsilon\frac{\Gamma_{n,k}^v}{\alpha_{n,k}(L_{n,k}^r)^{\varepsilon-1}}\frac{w_n L_{n,k}^p}{\theta\frac{L_{n,k}^r}{\delta}}$$
$$=\sum_k\varepsilon\frac{\Gamma_{n,k}^v}{\alpha_{n,k}(L_{n,k}^r)^{\varepsilon-1}}\frac{w_n(1-r)L_{n,k}}{\theta\frac{rL_{n,k}}{\delta}}=0$$

Hence at SS even with externality $\varepsilon\neq 1$, $\sum_k^J\frac{\gamma_{n,k}^v}{\alpha_{n,k}(rL_{n,k})^{\varepsilon-1}}=0$, and $\sum_k\varepsilon\Gamma_{n,k}^v\frac{w_n L_{n,k}^p}{\theta T_{n,k}}=0$, for any foreign country n.

N Proof for Proposition 10: optimal Markov policies

Appendix G lists optimal conditions for Markov policy in the baseline. Appendix G.3 proved $\sum_j^J\frac{\gamma_{n,j}^r}{\alpha_{n,j}}=0$ at baseline SS. The prove of $\sum_j^J\frac{\gamma_{n,j}^r}{\alpha_{n,j}(L_{n,j,t}^r)^{\varepsilon-1}}=0$ with externality is similar, and to avoid repetition, we just highlight the differences here.

All the optimal conditions are the same except that we consider ε in front of the current profit

reflecting Foreign country's Pigouvian innovation tax/subsidy

$$\frac{w_n}{\alpha_{n,j}(L^r_{n,j,t})^{\varepsilon-1}} = \varepsilon\frac{w_n L^p_{n,j}}{\theta T_{n,j}} + \tilde{\rho}_n v'_{n,j}, \qquad (\gamma^r_{n,j}),$$

Follow the same steps in Appendix G.3, Eq.A.90 now becomes

$$\gamma^r_{n,j}\frac{\varepsilon}{1+\theta}\frac{x_n}{T_{n,j}} - \gamma^L_{n,j}x_n - \mu_n\sum_j^J L^p_{n,j} + \sum_j^J \gamma^L_{n,j}\frac{1+\theta}{\theta}w_n L^p_{n,j} + \theta(1-\rho(1-\delta))\sum_j^J \gamma^T_{n,j}T_{n,j} + \sum_j^J \gamma^r_{n,j}\frac{\varepsilon w_n L^p_{n,j}}{T_{n,j}}$$
$$+\sum_j^J \gamma^r_{n,j}\left[\rho(1-\delta)M_{n,j}\right] - \sum_j \theta T_{n,j}\sum_{i\neq 1}^N\sum_k^J \gamma^r_{i,k}\rho(1-\delta)\frac{\partial G_{i,k}}{\partial T_{n,j}}x_i^\sigma P_i^{1-\sigma} = 0$$

At the steady state

$$T_{n,j} = \frac{\alpha_{n,j}(L^r_{n,j})^{\varepsilon-1}L^r_{n,j}}{\delta} = \frac{\alpha_{n,j}}{\delta}(L^r_{n,j})^{\varepsilon}, \quad \mu_n = \varepsilon\gamma^T_{n,j}\alpha_{n,j}(L^r_{n,j})^{\varepsilon-1} + \gamma^r_{n,j}(\varepsilon-1)\frac{w_n}{\alpha_{n,j}(L^r_{n,j})^{\varepsilon}},$$

hence

$$\gamma^r_{n,j}\frac{\varepsilon}{1+\theta}\frac{x_n}{T_{n,j}} - \gamma^L_{n,j}x_n - \mu_n\sum_j^J L^p_{n,j} + \sum_j^J \gamma^L_{n,j}\frac{1+\theta}{\theta}w_n L^p_{n,j} + \theta(1-\rho(1-\delta))\left(\sum_j \frac{1}{\delta\varepsilon}\mu_n L^r_{n,j} - \sum_j^J \gamma^r_{n,j}\frac{(\varepsilon-1)w_n}{\delta\varepsilon\alpha_{n,j}(L^r_{n,j})^{\varepsilon-1}}\right)$$
$$+\sum_j^J \gamma^r_{n,j}\frac{\varepsilon w_n L^p_{n,j}}{T_{n,j}} + \sum_j^J \gamma^r_{n,j}\left[\rho(1-\delta)M_{n,j}\right] - \sum_j \theta T_{n,j}\sum_{i\neq 1}^N\sum_k^J \gamma^r_{i,k}\rho(1-\delta)\frac{\partial G_{i,k}}{\partial T_{n,j}}x_i^\sigma P_i^{1-\sigma} = 0. \qquad (A.104)$$

At the steady state, the research intensity is given by

$$\bar{r}_j = \frac{\delta\varepsilon}{\theta(1-\rho(1-\delta)) + \delta\varepsilon}.$$

Using this research intensity, $L^p_{n,j} = (1-\bar{r})L_{n,j}$, and $L^r_{n,j} = \bar{r}L_{n,j}$, we can cancel out the two terms involving μ_n in Eq.A.104. Finally, we have

$$\gamma^r_{n,j}\frac{\varepsilon}{1+\theta}\frac{x_n}{T_{n,j}} - \gamma^L_{n,j}x_n + \sum_j^J \gamma^L_{n,j}\frac{1+\theta}{\theta}w_n L^p_{n,j} - \theta(1-\rho(1-\delta))(\sum_j^J \gamma^r_{n,j}\frac{(\varepsilon-1)w_n}{\delta\varepsilon\alpha_{n,j}(rL_{n,j})^{\varepsilon-1}})$$
$$+\sum_j^J \gamma^r_{n,j}\frac{\varepsilon w_n L^p_{n,j}}{T_{n,j}} + \sum_j^J \gamma^r_{n,j}\left[\rho(1-\delta)M_{n,j}\right] - \sum_j \theta T_{n,j}\sum_{i\neq 1}^N\sum_k^J \gamma^r_{i,k}\rho(1-\delta)\frac{\partial G_{i,k}}{\partial T_{n,j}}x_i^\sigma P_i^{1-\sigma} = 0.$$

Multiply the above equation $\frac{\frac{1+\theta}{\theta} w_n L_{n,j}^p}{x_n}$ on both sides and sum over j

$$\sum_j^J \gamma_{n,j}^r \frac{1}{\theta} \frac{\varepsilon w_n L_{n,j}^p}{T_{n,j}} - \theta(1-\rho(1-\delta))(\sum_j^J \gamma_{n,j}^r \frac{(\varepsilon-1)w_n}{\delta\varepsilon\alpha_{n,j}(rL_{n,j})^{\varepsilon-1}})$$
$$+ \sum_j^J \gamma_{n,j}^r \frac{\varepsilon w_n L_{n,j}^p}{T_{n,j}} + \sum_j^J \gamma_{n,j}^r \rho(1-\delta) \frac{w_n}{\alpha_{n,j}(rL_{n,j})^{\varepsilon-1}} - \sum_j \theta T_{n,j} \sum_{i\neq 1}^N \sum_k^J \gamma_{i,k}^r \rho(1-\delta) \frac{\partial x_i'^{-\sigma} P_i'^{\sigma-1} \frac{w_i'}{\alpha_{i,k}(rL_{i,k}')^{\varepsilon-1}}}{\partial T_{n,j}} x_i^\sigma P_i^{1-\sigma} = 0,$$

$$\sum_k^J \frac{\gamma_{n,k}^r}{\alpha_{n,k}(rL_{n,k})^{\varepsilon-1}} (1 + \frac{\theta(1-\rho(1-\delta))(1+(\delta-1)\varepsilon)}{\delta\varepsilon}) w_n$$
$$- \sum_j \theta T_{n,j} \sum_{i\neq 1}^N \sum_k^J \gamma_{i,k}^r \rho(1-\delta) \frac{\partial x_i'^{-\sigma} P_i'^{\sigma-1} \frac{w_i'}{\alpha_{i,k}(rL_{i,k}')^{\varepsilon-1}}}{\partial T_{n,j}} x_i^\sigma P_i^{1-\sigma} = 0$$

The equation holds when for every $n \neq 1$, $\sum_k^J \frac{\gamma_{n,k}^r}{\alpha_{n,k}(rL_{n,k})^{\varepsilon-1}} = 0$, and

$$\sum_j \theta T_{n,j} \sum_{i\neq 1}^N \sum_k^J \gamma_{i,k}^r \rho(1-\delta) \frac{\partial x_i'^{-\sigma} P_i'^{\sigma-1} \frac{w_i'}{\alpha_{i,k}(rL_{i,k}')^{\varepsilon-1}}}{\partial T_{n,j}} x_i^\sigma P_i^{1-\sigma} = 0. \tag{A.105}$$

With $\sum_k^J \frac{\gamma_{n,k}^r}{\alpha_{n,k}(rL_{n,k})^{\varepsilon-1}} = 0$, export taxes at the SS have the same formula as without externality. For innovation, Eq.A.105 implies the average wedge goes to zero at the SS, however, the wedges in each industry at SS will not go to zero.