International Trade and Innovation Dynamics with Endogenous Markups

by Laurent Cavenaile, Pau Roldan-Blanco, Tom Schmitz

Online Appendix

18th September 2020

Contents

A Derivations and Proofs **2**
A.1 Solution of the static Bertrand game 2
A.2 Proof of Lemma 1 2

B Numerical Appendix **4**
B.1 The numerical solution of the model 4
B.1.1 The static solution 4
B.1.2 Dynamic innovation choices 5
B.1.3 The invariant distribution of technology gaps 5
B.2 Internal calibration 5
B.3 Transition Dynamics Algorithm 7

C Robustness Results **9**

D Social Planner Problem **13**
D.1 Optimal labor allocation 13
D.2 Optimal innovation policies 18

A Derivations and Proofs

A.1 Solution of the static Bertrand game

In this section, we derive the solution of the Bertrand game between the Home and the Foreign leader. The static problem of the Home leader is

$$\max_{\{p^H_{jH,t},p^F_{jH,t}\}} \left\{ \underbrace{\left(p^H_{jH,t} - \frac{w_t}{q_{jH,t}}\right) y^H_{jH,t}}_{\text{Domestic profits}} + \underbrace{\left(p^F_{jH,t} - \tau\frac{w_t}{q_{jH,t}}\right) y^F_{jH,t}}_{\text{Foreign profits}} \right\}, \tag{A.1}$$

where the quantities sold at Home and in Foreign, $y^H_{jH,t}$ and $y^F_{jH,t}$, are given by the demand function (10). As specified in the main text, the Home leader takes the price of the Foreign leader and the fringes as given. The first-order optimality conditions of Problem (A.1) are

$$\text{Domestic price } (p^H_{jH,t}): \qquad y^H_{jH,t} + p^H_{jH,t}\frac{\partial y^H_{jH,t}}{\partial p^H_{jH,t}} = \frac{w_t}{q_{jH,t}}\frac{\partial y^H_{jH,t}}{\partial p^H_{jH,t}} \tag{A.2}$$

$$\text{Export price } (p^F_{jH,t}): \qquad y^F_{jH,t} + p^F_{jH,t}\frac{\partial y^F_{jH,t}}{\partial p^F_{jH,t}} = \tau\frac{w_t}{q_{jH,t}}\frac{\partial y^F_{jH,t}}{\partial p^F_{jH,t}} \tag{A.3}$$

These optimality conditions show that decisions on the Home and on the Foreign market are independent from each other. In each country, the leader equates its marginal cost to the marginal benefit of increasing prices by one unit. Combining the first-order conditions with the demand function (Equation (10)) and the definition of market shares (Equation (12)), we obtain Equation (11) and an expression for the export price of the Home leader, which by symmetry yields Equation (13).

Finally, replacing Equations (10), (11) and (13) into the leader's profit function, defined as the value of Problem (A.1), we get the expression for profits stated in Equation (15).

A.2 Proof of Lemma 1

In this proof, we focus on the Home country throughout. Using the definitions in Equations (3) and (4), we can express total output as

$$\begin{aligned}\ln \boldsymbol{Y}^H_t &= \int_0^1 \ln\left(Y^H_{jt}\right)\mathrm{d}j = \frac{\eta}{\eta-1}\int_0^1 \ln\left(\sum_{c=H,C_H,F} (\omega_c)^{\frac{1}{\eta}}\left(y^H_{jc,t}\right)^{\frac{\eta-1}{\eta}}\right)\mathrm{d}j \\ &= \frac{\eta}{\eta-1}\int_0^1 \left[\ln\left(\left(y^H_{jH,t}\right)^{\frac{\eta-1}{\eta}}\right) + \ln\left(\sum_{c=H,C_H,F}(\omega_c)^{\frac{1}{\eta}}\left(\frac{y^H_{jc,t}}{y^H_{jH,t}}\right)^{\frac{\eta-1}{\eta}}\right)\right]\mathrm{d}j\end{aligned} \tag{A.4}$$

$$= \int_0^1 \ln(y^H_{jH,t})\mathrm{d}j + \underbrace{\frac{\eta}{\eta-1}\int_0^1 \ln\left(\sum_{c=H,C_H,F} (\omega_c)^{\frac{1}{\eta}} \left(\frac{y^H_{jc,t}}{y^H_{jH,t}}\right)^{\frac{\eta-1}{\eta}}\right)\mathrm{d}j}_{\equiv \Psi^H_t}.$$

Using the fact that relative outputs only depend on technology gaps, we can express Ψ^H_t as

$$\Psi^H_t = \frac{\eta}{\eta-1}\sum_{n=-\infty}^{+\infty}\sum_{n_C=0}^{+\infty}\left(\varphi(n,n_C)\ln\left(\sum_{c=H,C_H,F}(\omega_c)^{\frac{1}{\eta}}\left(\frac{y^H_c(n,n_C)}{y^H_H(n,n_C)}\right)^{\frac{\eta-1}{\eta}}\right)\right). \tag{A.5}$$

Conditional on the technology gap, output ratios are constant over time. The technology gap distribution is constant over time as well, and therefore, Ψ^H_t is a constant.

Furthermore, we can rewrite

$$\begin{aligned}\int_0^1 \ln(y^H_{jH,t})\mathrm{d}j &= \int_0^1 \ln(q_{jH,t})\mathrm{d}j + \int_0^1 \ln(\ell^H_{jH,t})\mathrm{d}j \\ &= \int_0^1 \ln(q_{jH,t})\mathrm{d}j + \sum_{n=-\infty}^{+\infty}\sum_{n_C=0}^{+\infty}\left(\varphi(n,n_C)\ln\left(\ell^H_H(n,n_C)\right)\right).\end{aligned}$$

Again, the second term of the above equation is a constant. These intermediate results imply that output growth in Home is given by

$$\frac{\dot{\mathbf{Y}}^H_t}{\mathbf{Y}^H_t} = \dot{\Theta}^H_t, \tag{A.6}$$

where

$$\Theta^H_t = \int_0^1 \ln(q_{jH,t})\mathrm{d}j. \tag{A.7}$$

The productivity of Home leaders increases because of Home innovations (done by incumbents and potential entrants). The law of large numbers implies that in a short time interval of length $\Delta > 0$, the mass of Home innovations realized in an industry with technology gap $\underline{n}$ is $i_H(\underline{n})\Delta$, and the total mass of Home innovations realized in all industries is $\Delta\cdot\left(\sum_{n=-\infty}^{+\infty}\sum_{n_C=0}^{+\infty}\varphi(n,n_C)i_H(n,n_C)\right)$. Each innovation increases productivity by a factor $1+\lambda$. Therefore, we have

$$\Theta^H_{t+\Delta} = \Delta\cdot\left(\sum_{n=-\infty}^{+\infty}\sum_{n_C=0}^{+\infty}\varphi(n,n_C)i_H(n,n_C)\right)\cdot\ln(1+\lambda) + \Theta^H_t. \tag{A.8}$$

To arrive at this expression, we have used the log-linearity of Θ^H_t,which implies that the effect of an innovation does not depend on the productivity level of the industry to which it applies. Substracting Θ^H_t from both sides of the previous expression, dividing by Δ and taking the limit for

$\Delta \to 0$, we get

$$\dot{\Theta}_t^H = \left(\sum_{n=-\infty}^{+\infty} \sum_{n_C=0}^{+\infty} \varphi(n, n_C) i_H(n, n_C) \right) \cdot \ln(1+\lambda),$$

which gives the expression for aggregate growth stated in Lemma 1. It is easy to verify that aggregate symmetry implies that Foreign must grow at the same rate. ■

B Numerical Appendix

B.1 The numerical solution of the model

This section provides further details on the numerical solution of our model. This solution is greatly simplified by the fact that static decisions are independent of innovation choices, and innovation choices are themselves independent of the technology gap distribution.

As technology gaps between firms can a priori become infinitely large, we impose an upper bound $n_{\max} = 25$ on both n (the technology gap between Home leader and Foreign leader) and on $\max(n_C, n_C - n)$ (the technology gap between the fringes and the most productive leader). Imposing these bounds a priori changes firm behavior, as leaders recognize that they can never acquire an advantage exceeding $n_{\max}$. However, we make sure that bounds are irrelevant in practice, by verifying that the mass of firms in states in which the technology gap is maximal is always smaller than 0.001.

B.1.1 The static solution

To solve for relative prices, markups and market shares in an industry with technology gap $\underline{n} = (n, n_C)$, we reduce Equations (11) to (14) to a system of two equations in two unknowns. Simple algebra shows

$$\frac{p_F^H(\underline{n})}{p_H^H(\underline{n})} = (1+\lambda)^n \cdot \tau \cdot \frac{\left(\frac{\eta}{\eta-1} - \sigma_F^H(\underline{n})\right)\left(1 - \sigma_H^H(\underline{n})\right)}{\left(\frac{\eta}{\eta-1} - \sigma_H^H(\underline{n})\right)\left(1 - \sigma_F^H(\underline{n})\right)} \tag{B.1}$$

where we have used that $q_H/q_F = (1+\lambda)^n$. Likewise, the relative price of the Home competitive fringe with respect to the Home leader holds

$$\frac{p_{C_H}^H(\underline{n})}{p_H^H(\underline{n})} = (1+\lambda)^{n_C} \cdot \frac{1 - \sigma_H^H(\underline{n})}{\frac{\eta}{\eta-1} - \sigma_H^H(\underline{n})} \tag{B.2}$$

As shown in Equation (12), market shares are themselves a function of relative prices. For instance, the market share of the Home leader on the Home market is equal to

$$\sigma_H^H(\underline{n}) = \frac{1}{1 + \frac{\omega_F}{\omega_H}\left(\frac{p_F^H(\underline{n})}{p_H^H(\underline{n})}\right)^{1-\eta} + \frac{\omega_C}{\omega_H}\left(\frac{p_{C_H}^H(\underline{n})}{p_H^H(\underline{n})}\right)^{1-\eta}}, \tag{B.3}$$

and the market share of the Foreign leader on the Home market is equal to

$$\sigma_F^H(\underline{n}) = \frac{\frac{\omega_F}{\omega_H}\left(\frac{p_F^H(\underline{n})}{p_H^H(\underline{n})}\right)^{1-\eta}}{1 + \frac{\omega_F}{\omega_H}\left(\frac{p_F^H(\underline{n})}{p_H^H(\underline{n})}\right)^{1-\eta} + \frac{\omega_C}{\omega_H}\left(\frac{p_{C_H}^H(\underline{n})}{p_H^H(\underline{n})}\right)^{1-\eta}}. \tag{B.4}$$

Thus, we can use Equations (B.3) and (B.4) to substitute out market shares from Equations (B.1) and (B.2). This yields a system of two equations with two unknowns, the relative prices $p_F^H(\underline{n})/p_H^H(\underline{n})$ and $p_{C_H}^H(\underline{n})/p_H^H(\underline{n})$. This system can be solved numerically. Furthermore, once we know relative prices, we can immediately deduce market shares and markups for all firms.

B.1.2 Dynamic innovation choices

To solve for the value functions of Home leaders in both sectors, we use a simple Value Function Iteration algorithm, described below.

1. Guess an initial value function $\left(v_H^{(k=0)}(\underline{n})\right)$, where the superscript k stands for the current iteration.
2. For any given iteration $k \in \mathbb{N}$, deduce the optimal R&D choices of entrants and incumbents in both countries, using Equations (24) and (25), and the symmetry implied by Equation (22).
3. Deduce the new implied values of the value function $v_H^{new}(\underline{n})$, given by Equation (23).
4. Update the guess for the value function according to

$$v_H^{(k+1)}(\underline{n}) = \iota v_H^{new}(\underline{n}) + (1-\iota)\, v_H^{(k)}(\underline{n}),$$

   where $\iota \in (0,1)$ is a dampening parameter.
5. Iterate on 2-4 until the difference between $v_H^{new}(\underline{n})$ and $v_H^{(k)}(\underline{n})$ is sufficiently small.

B.1.3 The invariant distribution of technology gaps

Once all innovation rates are known, Equations (26), (27) and the condition that the distribution sums up to 1 form a linear system of equations, which can be easily solved numerically.

B.2 Internal calibration

In order to estimate the vector of internal parameters $\boldsymbol{\theta}$, we define a model-data distance function

$$\sum_{m=1}^{M}\left|\frac{\text{Moment}_m\,(\text{Data}) - \text{Moment}_m\,(\text{Model}, \boldsymbol{\theta})}{\frac{1}{2}\Big(\text{Moment}_m\,(\text{Data}) + \text{Moment}_m\,(\text{Model}, \boldsymbol{\theta})\Big)}\right|.$$

where M is the number of moments. We find the vector $\boldsymbol{\theta}$ that minimizes this function using a Differential Evolution algorithm, developed by Markus Buehren and available for download at https://it.mathworks.com/matlabcentral/fileexchange/18593-differential-evolution. This method requires setting bounds on each parameter. Table B.1 lists the bounds we use, and shows that none of them are binding in our baseline calibration.

	λ	χ_i	χ_e	η	τ	ω_H	ξ	ζ
Lower bound	0.07	1.2	22	10	1.2	0.49	0.2	3.5
Upper bound	0.09	1.5	27	13	1.4	0.5	0.3	6

Table B.1: Parameter bounds for the Differential Evolution algorithm.

Next, we describe how we compute the different targeted moments in the model.

Aggregate and standard deviation of import shares

$$\frac{M_t}{Y_t} = \sum_{n=-\infty}^{+\infty} \sum_{n_C=0}^{+\infty} \varphi(n, n_C) \sigma_F^H(n, n_C)$$

$$StDevImpSh = \sqrt{\sum_{n=-\infty}^{+\infty} \sum_{n_C=0}^{+\infty} \varphi(n, n_C) \left(\sigma_F^H(n, n_C) - \frac{M_t}{Y_t} \right)^2}$$

Average markup We compute the sales-weighted average markup as

$$\overline{\boldsymbol{\mu}} = \sum_{n=-\infty}^{+\infty} \sum_{n_C=0}^{+\infty} \overline{\mu}_H(n, n_C) \overline{\sigma}_H(n, n_C)$$

where

$$\overline{\mu}_H(n, n_C) \equiv \frac{\sigma_H^H(n, n_C)\mu_H^H(n, n_C) + \sigma_{C_H}^H(n, n_C)\mu_{C_H}^H(n, n_C) + \sigma_H^F(n, n_C)\mu_H^F(n, n_C)}{\sigma_H^H(n, n_C) + \sigma_{C_H}^H(n, n_C) + \sigma_H^F(n, n_C)}$$

$$\overline{\sigma}_H(n, n_C) \equiv \frac{\varphi(n, n_C) \left(\sigma_H^H(n, n_C) + \sigma_{C_H}^H(n, n_C) + \sigma_H^F(n, n_C) \right)}{\sum_{\tilde{n}=-\infty}^{+\infty} \sum_{\tilde{n}_C=0}^{+\infty} \varphi(\tilde{n}, \tilde{n}_C) \left(\sigma_H^H(\tilde{n}, \tilde{n}_C) + \sigma_{C_H}^H(\tilde{n}, \tilde{n}_C) + \sigma_H^F(\tilde{n}, \tilde{n}_C) \right)}$$

are the sales-weighted average markup of domestically-produced goods within an industry, and the sales weight of this industry, respectively. This definition for the average markup is motivated by the fact that in the Compustat data we observe sales of domestic firms, but not of foreign firms selling in the domestic market.

Standard deviation of markups: We compute a sales-weighted standard deviation of markups as

$$\widehat{\mu}_H = \sqrt{\sum_{n=-\infty}^{+\infty} \sum_{n_C=0}^{+\infty} \overline{\sigma}_H(n, n_C)\widehat{\mu}_H(n, n_C)}$$

where

$$\widehat{\mu}_H(n, n_C) \equiv \frac{\sigma_H^H(n, n_C)\Big(\mu_H^H(n, n_C) - \overline{\boldsymbol{\mu}}\Big)^2 + \sigma_{C_H}^H(n, n_C)\Big(\mu_{C_H}^H(n, n_C) - \overline{\boldsymbol{\mu}}\Big)^2 + \sigma_H^F(n, n_C)\Big(\mu_H^F(n, n_C) - \overline{\boldsymbol{\mu}}\Big)^2}{\sigma_H^H(n, n_C) + \sigma_{C_H}^H(n, n_C) + \sigma_H^F(n, n_C)}$$

is a sales-weighted average standard deviation of markups within an industry.

Entry rate

$$EntryRate = \sum_{n=-\infty}^{+\infty} \sum_{n_C=0}^{+\infty} \varphi(n, n_C) x_H(n, n_C)$$

Contribution of entrants to growth We define the contribution of entrants to growth as the percentage of innovations realized by entrants:

$$ContEntGrowth = \frac{\sum_{n=-\infty}^{+\infty} \sum_{n_C=0}^{+\infty} \varphi(n, n_C) x_H(n, n_C)}{\sum_{n=-\infty}^{+\infty} \sum_{n_C=0}^{+\infty} \varphi(n, n_C) i_H(n, n_C)}$$

Employment share of the fringe We use the labor demand of the fringe (Equation (16)) and the aggregate markup $\boldsymbol{\mu}$ to compute

$$EmpShareFringe = \boldsymbol{\mu} \sum_{n=-\infty}^{+\infty} \sum_{n_C=0}^{+\infty} \varphi(n, n_C) \sigma_{C_H}^H(n, n_C)$$

B.3 Transition Dynamics Algorithm

To compute transition dynamics, we assume that the economy is initially on a given BGP. At some time t_0, it is hit by a permanent and unexpected shock which lowers trade costs τ. Eventually, the economy will then converge to a new low trade cost BGP. This section describes how we compute outcomes during the transition period.

We assume that the transition is completed after some long period of time T (in practice, we set $T = 1,000$). Furthermore, we discretize the model, considering short time interals of length $\Delta = 0.01$. Thus, we solve the model on a time grid $\mathcal{T} = \{\Delta, 2\Delta, 3\Delta, \ldots, T - \Delta, T\}$.

1. Set the iteration counter to $k = 0$. Guess paths for the growth rate of output and the interest rate, $\underline{g}_Y^{(k=0)} = \Big\{g_{Y,t}^{(k=0)} : t \in \mathcal{T}\Big\}$ and $\underline{r}^{(k=0)} = \Big\{r_t^{(k=0)} : t \in \mathcal{T}\Big\}$.

2. For any given iteration $k \in \mathbb{N}$, and going backwards from $t = T - \Delta$ to $t = \Delta$:

   (a) Solve for the innovation rates $\left\{ z_{H,t}^{(k)}(\underline{n}), x_{H,t}^{(k)}(\underline{n}) \right\}$ using:

$$
\begin{aligned}
z_{H,t}^{(k)}(\underline{n}) &= \left(e^{-\Delta r_{t+\Delta}^{(k)}} \frac{v_{H,t+\Delta}^{(k)}(n+1, n_C+1) - v_{H,t+\Delta}^{(k)}(\underline{n})}{\psi_i \chi_i} \right)^{\frac{1}{\psi_i - 1}} \\
x_{H,t}^{(k)}(\underline{n}) &= \left(e^{-\Delta r_{t+\Delta}^{(k)}} \frac{v_{H,t+\Delta}^{(k)}(n+1, n_C+1)}{\psi_e \chi_e} \right)^{\frac{1}{\psi_e - 1}}
\end{aligned}
$$

   with initial condition (i.e. at $t = T - \Delta$) given by $v_{H,T}^{(k)}(\underline{n}) = v_H^{final}(\underline{n})$, the value function from the final BGP.

   (b) Use the results from (a) to get $v_{H,t}^{(k)}(\underline{n})$ from the discrete-time Bellman equation:

$$
\begin{aligned}
v_{H,t}^{(k)}(\underline{n}) = \; & \pi_H^{H,final}(\underline{n})\Delta + \pi_H^{F,final}(\underline{n})\Delta - \chi_i \left(z_{H,t}^{(k)}(\underline{n}) \right)^{\psi_i} \Delta \\
& + \Delta e^{\Delta (g_{Y,t+\Delta}^{(k)} - r_{t+\Delta}^{(k)})} \Bigg[- x_{H,t}^{(k)}(\underline{n}) v_{H,t+\Delta}^{(k)}(\underline{n}) \\
& \quad + \left(z_{H,t}^{(k)}(\underline{n}) + \mathbb{1}_{n<0}\xi \right) \left(v_{H,t+\Delta}^{(k)}(n+1, n_C+1) - v_{H,t+\Delta}^{(k)}(\underline{n}) \right) \\
& \quad + \left(x_{F,t}^{(k)}(\underline{n}) + z_{F,t}^{(k)}(\underline{n}) + \mathbb{1}_{n>0}\xi \right) \left(v_{H,t+\Delta}^{(k)}(n-1, n_C) - v_{H,t+\Delta}^{(k)}(\underline{n}) \right) \\
& \quad + \zeta \left(v_{H,t+\Delta}^{(k)}(n, \max(n, 0, n_C - 1)) - v_{H,t+\Delta}^{(k)}(\underline{n}) \right) + \frac{v_{H,t+\Delta}^{(k)}(\underline{n})}{\Delta} \Bigg]
\end{aligned}
$$

   where $\pi_H^{H,final}(\underline{n})$ and $\pi_H^{F,final}(\underline{n})$ are the domestic and foreign profits from the final BGP.

3. Using $\left\{ z_{H,t}^{(k)}(\underline{n}), x_{H,t}^{(k)}(\underline{n}) : t \in \mathcal{T} \right\}$ and $\left\{ \varphi^{initial}(\underline{n}) \right\}$, the stationary technology gap distribution from the initial BGP, compute $\left\{ \varphi_t^{(k)}(\underline{n}) : t \in \mathcal{T} \right\}$ using the flow equations.

4. Compute the implied aggregate output and aggregate consumption levels, for each $t \in \mathcal{T}$.

   (a) For aggregate output, use Equation (A.4):

$$
\boldsymbol{Y}_t^{(k)} = \exp \left(\Theta_t^{(k)} + \sum_{n=-\infty}^{+\infty} \sum_{n_C=0}^{+\infty} \varphi_t^{(k)}(\underline{n}) \ln \left(\ell_H^{H,final}(\underline{n}) \right) + \Psi_t^{H,(k)} \right)
$$

   where $\Psi_t^{H,(k)}$ is defined as in Equation (A.5), $\Theta_t^{(k)}$ is defined as in Equation (A.7), and $\ell_H^{H,final}(\underline{n})$ is the labor allocation from the final BGP. To compute $\left\{ \Theta_t^{(k)} : t \in \mathcal{T} \right\}$, we normalize $\Theta_0^{(k)} = 0$ and use Equation (A.8) to update.

   (b) For aggregate consumption, use $\boldsymbol{C}_t^{(k)} = \boldsymbol{Y}_t^{(k)} \left(1 - \frac{\boldsymbol{R}_t^{(k)}}{\boldsymbol{Y}_t^{(k)}} \right)$ by the resource constraint, where

$$\frac{\boldsymbol{R}_t^{(k)}}{\boldsymbol{Y}_t^{(k)}} = \sum_{n=-\infty}^{+\infty} \sum_{n_C=0}^{+\infty} \varphi_t^{(k)}(\underline{n}) \left(\chi_i \left(z_{H,t}^{(k)}(\underline{n}) \right)^{\psi_i} + \chi_e \left(x_{H,t}^{(k)}(\underline{n}) \right)^{\psi_e} \right)$$

5. For each $t \in \mathcal{T}$, compute the implied growth rates of output and consumption, and get a new interest rate from the Euler equation:

$$g_{Y,t}^{new} = \frac{1}{\Delta} \left(\frac{\boldsymbol{Y}_t^{(k)}}{\boldsymbol{Y}_{t-\Delta}^{(k)}} - 1 \right) \qquad \text{and} \qquad r_t^{new} = \frac{1}{\Delta} \left(\frac{\boldsymbol{C}_t^{(k)}}{\boldsymbol{C}_{t-\Delta}^{(k)}} - 1 \right) + \rho$$

6. Collect these results in vectors $\underline{g}_Y^{new} = \left\{ g_{Y,t}^{new} : t \in \mathcal{T} \right\}$ and $\underline{r}^{new} = \left\{ r_t^{new} : t \in \mathcal{T} \right\}$. Stop if $\|\underline{g}_Y^{new} - \underline{g}_Y^{(k)}\| < \varepsilon$ and $\|\underline{r}^{new} - \underline{r}^{(k)}\| < \varepsilon$ for some small tolerance $\varepsilon > 0$, where $\|\cdot\|$ denotes the sup-norm. Otherwise, define:

$$\underline{g}_Y^{(k+1)} = \iota \underline{g}_Y^{new} + (1-\iota) \underline{g}_Y^{(k)} \qquad \text{and} \qquad \underline{r}^{(k+1)} = \iota \underline{r}^{new} + (1-\iota) \underline{r}^{(k)}$$

where $\iota \in (0,1)$ is a dampening parameter, and go back to 2. with $[k] \leftarrow [k+1]$.

C Robustness Results

Table C.1 lists the internally calibrated parameter values for the different robustness checks described in Section 3.4, and Table C.2 shows how the different calibrations fit the targeted moments. External parameter values are set to their baseline values described in the main text, and targets are the same as the baseline ones in Table 2 (with the exception of the low markup calibration).

In the model with fixed costs of exporting, we assume that leaders need to pay a flow cost $\kappa \boldsymbol{Y}_t$, with $\kappa > 0$, in order to export their product. In equilibrium, the Home leader exports if and only if its export profits are sufficiently high, that is, if and only if $\pi_H^F(\underline{n}) \geq \kappa$. Using Equation (15), this yields a threshold market share: the Home leader exports if its market share $\sigma_H^F(\underline{n})$ exceeds

$$\widehat{\sigma} = \frac{\kappa \eta}{1 + \kappa(\eta - 1)}.$$

All other equilibrium conditions are unchanged in this extended model. To discipline the new parameter κ, we target the share of leaders that export. Bernard *et al.* (2018) show that in 2007, 35% of US manufacturing firms exported. Harris and Moffat (2011) show that over the period 2004-2008, the prevalence of exporting was about 41% higher among R&D performing firms (the equivalent of leaders in our model) than in the general population of manufacturing firms in the United Kingdom. Assuming that the same relationship also holds for the United States, we obtain that 49% of R&D-performing firms export.

Table C.3 summarizes the results of the different robustness checks. It presents, for each robustness check, the analogue of Panel 1 of Table 3. That is, we consider for each case a transition from a high trade cost BGP (in which the trade-to-GDP ratio is half as high as in the baseline) to the baseline low trade cost BGP.

			(1)	(2)	(3)	(4)
	Parameter	**Baseline**	**Lower** μ	$\eta = 7$	$\eta = 16$	κ **cost**
λ	Innovation step	0.0869	0.0623	0.091	0.084	0.0747
χ_i	R&D scale, incumbents	1.37	0.607	1.416	1.332	1.576
χ_e	R&D scale, entrants	23.179	11.846	30	20.248	24.21
η	Within-industry elasticity	11.161	15.6	7	16	8.505
τ	Variable trade cost	1.2746	1.1874	1.324	1.242	1.2752
ω_H	Quality of Home leader	0.4928	0.491	0.4878	0.4953	0.4944
ξ	Catch-up rate leaders	0.2348	0.332	0.2031	0.2544	0.204
ζ	Catch-up rate fringes	5.446	1.000	5.000	4.989	7.706
κ	Fixed cost of exporting	.	.	.	.	0.038
	Period: 1 year					

Table C.1: Internally calibrated parameters for the different calibrations. Notes: Internally calibrated parameters are obtained by indirect inference, targeting the data moments listed in Table C.2. Additionally, for the model with fixed export costs, we target the share of exporting firms.

Moment	Baseline	(1) **Lower μ**	(2) $\eta = 7$	(3) $\eta = 16$	(4) **κ cost**	**Data**	**Data Source**
A. From aggregate data							
Productivity growth	1.61%	1.59%	1.61%	1.61%	1.59%	1.58%	EU KLEMS
R&D share of VA	8.3%	6.3%	9.7%	7.6%	10.8%	9.8%	OECD
Import share	24.0%	23.4%	27.0%	21.5%	23.3%	23.5%	US Census, NBER-CES
St. dev. of import shares	18.1%	17.2%	14.6%	21.2%	21.2%	21.3%	US Census, NBER-CES
Share of exporting firms	.	.	.	.	63.4%	49.0%	Bernard *et al.* (2018)
B. From firm-level data							
Average markup	33.0%	20.4%	39.7%	30.5%	36.6%	35.4%*	See main text
St. dev. of markups	49.5%	26.7%	47.0%	51.3%	43.0%	48.1%*	See main text
Entry rate	5.2%	6.4%	5.2%	5.2%	5.9%	6.5%	US Census
Contr. entrants to growth	27.0%	24.2%	28.0%	26.0%	26.6%	25.7%	Akcigit and Kerr (2018)
Emp. share of fringe	17.5%	17.6%	20.5%	16.2%	17.8%	18.2%	US Census, NSF

Table C.2: Targeted moments: model versus data, for the different calibrations. Notes: All data moments refer to the US manufacturing sector. For the lower markup calibration (Column (1)), the target for the average markup is 17.7%, and the target for the standard deviation of markups is 24.1%.

	(1)	(2)	(3)	(4)	(5)
Variable	$\text{BGP}_{\text{initial}}$	$\text{BGP}_{\text{final}}$	**Total change**	**Impact**	**Transition**
Panel 1. *Lower markup target*					
Productivity growth	1.50%	1.59%	+0.09	+0.06	+0.03
Aggregate markup	16.18%	17.15%	+0.97	−0.82	+1.79
Trade share	11.64%	23.36%	+11.72	+10.43	+1.29
Emp. share of the fringe	25.59%	17.58%	−8.01	−7.27	−0.74
Trade cost (τ)	1.3026	1.1874	.	.	.
Panel 2. *Fixing* $\eta = 7$					
Productivity growth	1.48%	1.61%	+0.13	+0.09	+0.04
Aggregate markup	30.70%	32.20%	+1.5	−1.08	+2.58
Trade share	13.50%	27.05%	+13.55	+12.58	+0.97
Emp. share of the fringe	30.31%	20.51%	−9.80	−9.17	−0.63
Trade cost (τ)	1.6404	1.3240	.	.	.
Panel 3. *Fixing* $\eta = 16$					
Productivity growth	1.49%	1.61%	+0.12	+0.08	+0.04
Aggregate markup	20.02%	22.21%	+2.19	−0.87	+3.06
Trade share	10.73%	21.48%	+10.75	+8.84	+1.91
Emp. share of the fringe	23.51%	16.23%	−7.28	−6.50	−0.78
Trade cost (τ)	1.3637	1.2420	.	.	.
Panel 4. *Fixed export cost* κ					
Productivity growth	1.45%	1.59%	+0.15	+0.04	+0.11
Aggregate markup	33.02%	33.98%	+0.96	−2.00	+2.96
Trade share	11.62%	23.26%	+11.64	+11.04	+0.60
Emp. share of the fringe	26.69%	17.79%	−8.90	−8.65	−0.25
Trade cost (τ)	1.3492	1.2752	.	.	.

Table C.3: The quantitative importance of the innovation feedback effect for different robustness checks. Notes: In all cases, the final BGP corresponds to the baseline calibration, and the initial BGP has a higher value of τ, yielding a trade-to-GDP ratio that is half as high as in the baseline. All differences in Columns (3) to (5) are stated in percentage points. For the results in the baseline calibration, see Table 3. For the algorithm that computes the transition dynamics between BGPs, see Appendix B.3.

D Social Planner Problem

This section analyzes the problem of a global social planner who cares about both Home and Foreign, with equal weights. As in the decentralized solution, we first solve for the optimal allocation of labor across firms, and then for the optimal innovation policies.

D.1 Optimal labor allocation

First, we determine the optimal allocation of labor across firms at any given point in time, conditional on the technology gap distribution. Lemma 1 characterizes the optimal allocation for the Home market. Allocations for the Foreign market are exactly analogous.

Lemma 1 *The planner allocates labor $\ell_c^{H*}(\underline{n}) = \sigma_c^{H*}(\underline{n})$ to firm $c = H, C_H, F$ of industry $\underline{n} \equiv (n, n_C)$, where*

$$\sigma_H^{H*}(\underline{n}) = \left(1 + \frac{\omega_C}{\omega_H}(1+\lambda)^{-n_C(\eta-1)} + \frac{\omega_F}{\omega_H}\left(\frac{(1+\lambda)^{-n}}{\tau}\right)^{\eta-1}\right)^{-1} \tag{D.1}$$

$$\sigma_{C_H}^{H*}(\underline{n}) = \left(1 + \frac{\omega_H}{\omega_C}(1+\lambda)^{n_C(\eta-1)} + \frac{\omega_F}{\omega_C}\left(\frac{(1+\lambda)^{n_C-n}}{\tau}\right)^{\eta-1}\right)^{-1} \tag{D.2}$$

$$\sigma_F^{H*}(\underline{n}) = \left(1 + \frac{\omega_H}{\omega_F}\left(\frac{\tau}{(1+\lambda)^{-n}}\right)^{\eta-1} + \frac{\omega_C}{\omega_F}\left(\frac{\tau}{(1+\lambda)^{n_C-n}}\right)^{\eta-1}\right)^{-1} \tag{D.3}$$

Proof. The planner seeks to maximize world output, defined by $\boldsymbol{Y}^W = (\boldsymbol{Y}^H)^{\frac{1}{2}}(\boldsymbol{Y}^F)^{\frac{1}{2}}$, and chooses $\mathbb{L}_j = \left\{\ell_{jc}^k : c = H, C_k, F; k = H, F\right\}$ for each industry $j \in [0,1]$ in order to solve:

$$\max_{(\mathbb{L}_j)_{j\in[0,1]}} \left\{ \exp\left[\int_0^1 \ln\left((\omega_H)^{\frac{1}{\eta}}\left(q_{jH}\ell_{jH}^H\right)^{\frac{\eta-1}{\eta}} + (\omega_C)^{\frac{1}{\eta}}\left(q_{jC_H}\ell_{jC_H}^H\right)^{\frac{\eta-1}{\eta}} + (\omega_F)^{\frac{1}{\eta}}\left(\frac{q_{jF}\ell_{jF}^H}{\tau}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} \mathrm{d}j\right]^{\frac{1}{2}} \right.$$
$$\left. \times \exp\left[\int_0^1 \ln\left((\omega_F)^{\frac{1}{\eta}}\left(q_{jF}\ell_{jF}^F\right)^{\frac{\eta-1}{\eta}} + (\omega_C)^{\frac{1}{\eta}}\left(q_{jC_F}\ell_{jC_F}^F\right)^{\frac{\eta-1}{\eta}} + (\omega_H)^{\frac{1}{\eta}}\left(\frac{q_{jH}\ell_{jH}^F}{\tau}\right)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} \mathrm{d}j\right]^{\frac{1}{2}} \right\}$$

subject to:

$$\int_0^1 \left(\ell_{jH}^H + \ell_{jC_H}^H + \ell_{jH}^F\right)\mathrm{d}j \leq 1;$$
$$\int_0^1 \left(\ell_{jF}^F + \ell_{jC_F}^F + \ell_{jF}^H\right)\mathrm{d}j \leq 1;$$
$$\ell_{jH}^H, \ell_{jC_H}^H, \ell_{jF}^H, \ell_{jF}^F, \ell_{jC_F}^F, \ell_{jH}^F \geq 0; \quad \forall j$$

Let $\beta^k \geq 0$ be the Lagrange multiplier on the feasibility constraint for country $k = H, F$. The optimality conditions yield:

$$\ell_{jH}^H = \left(\frac{\frac{1}{2}\boldsymbol{Y}^W}{\beta^H}\right)^\eta \omega_H \left(\frac{q_{jH}}{Y_j^H}\right)^{\eta-1}, \quad \ell_{jC_H}^H = \left(\frac{\frac{1}{2}\boldsymbol{Y}^W}{\beta^H}\right)^\eta \omega_C \left(\frac{q_{jC_H}}{Y_j^H}\right)^{\eta-1}, \quad \ell_{jH}^F = \left(\frac{\frac{1}{2}\boldsymbol{Y}^W}{\beta^H}\right)^\eta \omega_H \left(\frac{q_{jH}/\tau}{Y_j^F}\right)^{\eta-1}$$
$$\ell_{jF}^F = \left(\frac{\frac{1}{2}\boldsymbol{Y}^W}{\beta^F}\right)^\eta \omega_F \left(\frac{q_{jF}}{Y_j^F}\right)^{\eta-1}, \quad \ell_{jC_F}^F = \left(\frac{\frac{1}{2}\boldsymbol{Y}^W}{\beta^F}\right)^\eta \omega_C \left(\frac{q_{jC_F}}{Y_j^F}\right)^{\eta-1}, \quad \ell_{jF}^H = \left(\frac{\frac{1}{2}\boldsymbol{Y}^W}{\beta^F}\right)^\eta \omega_F \left(\frac{q_{jF}/\tau}{Y_j^H}\right)^{\eta-1}$$

Using the formula for H's industry output we obtain:

$$\left(\frac{Y_j^H}{\frac{1}{2}\boldsymbol{Y}^W}\right)^{\eta-1} = \left(\frac{1}{\beta^H}\right)^{\eta-1} \omega_H (q_{jH})^{\eta-1} + \left(\frac{1}{\beta^H}\right)^{\eta-1} \omega_C (q_{jC_H})^{\eta-1} + \left(\frac{1}{\beta^F}\right)^{\eta-1} \omega_F \left(\frac{q_{jF}}{\tau}\right)^{\eta-1} \tag{D.4}$$

Imposing feasibility in H, i.e. $\int_0^1 \left(\ell_{jH}^H + \ell_{jC_H}^H + \ell_{jH}^F\right)\mathrm{d}j$, gives:

$$1 = \left(\frac{\frac{1}{2}\boldsymbol{Y}^W}{\beta^H}\right)^\eta \int_0^1 \left(\omega_H \left(\frac{q_{jH}}{Y_j^H}\right)^{\eta-1} + \omega_C \left(\frac{q_{jC_H}}{Y_j^H}\right)^{\eta-1} + \omega_F \left(\frac{q_{jH}/\tau}{Y_j^F}\right)^{\eta-1}\right)\mathrm{d}j \tag{D.5}$$

The counterparts of (D.4)-(D.5) for F give us four equations and four unknowns, $(\beta^H, \beta^F, Y_j^H, Y_j^F)$. Plugging (D.4) into (D.5) yields:

$$\begin{aligned}(\beta^H)^\eta = \frac{1}{2}\boldsymbol{Y}^W \int_0^1 \Bigg(& \frac{\omega_H \left(q_{jH}\right)^{\eta-1}}{\left(\frac{1}{\beta^H}\right)^{\eta-1} \omega_H (q_{jH})^{\eta-1} + \left(\frac{1}{\beta^H}\right)^{\eta-1} \omega_C (q_{jC_H})^{\eta-1} + \left(\frac{1}{\beta^F}\right)^{\eta-1} \omega_F \left(\frac{q_{jF}}{\tau}\right)^{\eta-1}} \\ & + \frac{\omega_C \left(q_{jC_H}\right)^{\eta-1}}{\left(\frac{1}{\beta^H}\right)^{\eta-1} \omega_H (q_{jH})^{\eta-1} + \left(\frac{1}{\beta^H}\right)^{\eta-1} \omega_C (q_{jC_H})^{\eta-1} + \left(\frac{1}{\beta^F}\right)^{\eta-1} \omega_F \left(\frac{q_{jF}}{\tau}\right)^{\eta-1}} \\ & + \frac{\omega_F \left(\frac{q_{jH}}{\tau}\right)^{\eta-1}}{\left(\frac{1}{\beta^F}\right)^{\eta-1} \omega_F (q_{jF})^{\eta-1} + \left(\frac{1}{\beta^F}\right)^{\eta-1} \omega_C (q_{jC_F})^{\eta-1} + \left(\frac{1}{\beta^H}\right)^{\eta-1} \omega_H \left(\frac{q_{jH}}{\tau}\right)^{\eta-1}}\Bigg)\mathrm{d}j \end{aligned} \tag{D.6}$$

As in the decentralized allocation, we focus on a symmetric solution where $\beta^H = \beta^F = \beta$. Then the Equation (D.6) gives $\beta = \frac{1}{2}\boldsymbol{Y}^W$. Given this, Equation (D.4) gives:

$$Y_j^H = \left[\omega_H (q_{jH})^{\eta-1} + \omega_C (q_{jC_H})^{\eta-1} + \omega_F \left(\frac{q_{jF}}{\tau}\right)^{\eta-1}\right]^{\frac{1}{\eta-1}} \tag{D.7a}$$

and similarly for F:

$$Y_j^F = \left[\omega_F (q_{jF})^{\eta-1} + \omega_C (q_{jC_F})^{\eta-1} + \omega_H \left(\frac{q_{jH}}{\tau}\right)^{\eta-1}\right]^{\frac{1}{\eta-1}} \tag{D.7b}$$

Total world output $\ln \boldsymbol{Y}^W = \frac{1}{2}\ln \boldsymbol{Y}^H + \frac{1}{2}\ln \boldsymbol{Y}^F$ is then given by:

$$\ln \boldsymbol{Y}^W = \frac{1}{2}\left(\int_0^1 \ln Y_j^H \mathrm{d}j + \int_0^1 \ln Y_j^F \mathrm{d}j\right) \tag{D.8}$$

The labor allocation within industries and across countries is, then:

$$\begin{aligned}
\ell_{jH}^H &= \omega_H \left(\frac{q_{jH}}{Y_j^H}\right)^{\eta-1}, \quad \ell_{jC_H}^H = \omega_C \left(\frac{q_{jC_H}}{Y_j^H}\right)^{\eta-1}, \quad \ell_{jH}^F = \omega_H \left(\frac{q_{jH}/\tau}{Y_j^F}\right)^{\eta-1} \\
\ell_{jF}^F &= \omega_F \left(\frac{q_{jF}}{Y_j^F}\right)^{\eta-1}, \quad \ell_{jC_F}^F = \omega_C \left(\frac{q_{jC_F}}{Y_j^F}\right)^{\eta-1}, \quad \ell_{jF}^H = \omega_F \left(\frac{q_{jF}/\tau}{Y_j^H}\right)^{\eta-1}
\end{aligned} \tag{D.9}$$

In particular, labor for use in the domestic market (by the leader and the fringe) is:

$$\ell_{jH}^H = \frac{\omega_H}{\omega_H + \omega_C \left(\frac{q_{jC_H}}{q_{jH}}\right)^{\eta-1} + \omega_F \left(\frac{q_{jF}/\tau}{q_{jH}}\right)^{\eta-1}} \tag{D.10a}$$

$$\ell_{jC_H}^H = \frac{\omega_C}{\omega_H \left(\frac{q_{jH}}{q_{jC_H}}\right)^{\eta-1} + \omega_C + \omega_F \left(\frac{q_{jF}/\tau}{q_{jC_H}}\right)^{\eta-1}} \tag{D.10b}$$

$$\ell_{jF}^F = \frac{\omega_F}{\omega_F + \omega_C \left(\frac{q_{jC_F}}{q_{jF}}\right)^{\eta-1} + \omega_H \left(\frac{q_{jH}/\tau}{q_{jF}}\right)^{\eta-1}} \tag{D.10c}$$

$$\ell_{jC_F}^F = \frac{\omega_C}{\omega_F \left(\frac{q_{jF}}{q_{jC_F}}\right)^{\eta-1} + \omega_C + \omega_H \left(\frac{q_{jH}/\tau}{q_{jC_F}}\right)^{\eta-1}} \tag{D.10d}$$

To find labor use for the export market, first use (D.7b) to note that:

$$\begin{aligned}
\left(\frac{Y_j^F}{q_{jH}/\tau}\right)^{\eta-1} &= \omega_F \left(\frac{q_{jF}}{q_{jH}/\tau}\right)^{\eta-1} + \omega_C \left(\frac{q_{jC_F}}{q_{jH}/\tau}\right)^{\eta-1} + \omega_H \\
\left(\frac{Y_j^H}{q_{jF}/\tau}\right)^{\eta-1} &= \omega_F \left(\frac{q_{jH}}{q_{jF}/\tau}\right)^{\eta-1} + \omega_C \left(\frac{q_{jC_H}}{q_{jF}/\tau}\right)^{\eta-1} + \omega_H
\end{aligned}$$

Therefore, labor use for exports by H and F firms, respectively, is:

$$\ell_{jH}^F = \frac{\omega_H}{\omega_F \left(\frac{q_{jF}}{q_{jH}/\tau}\right)^{\eta-1} + \omega_C \left(\frac{q_{jC_F}}{q_{jH}/\tau}\right)^{\eta-1} + \omega_H} \tag{D.10e}$$

$$\ell_{jF}^{H} = \frac{\omega_F}{\omega_H \left(\frac{q_{jH}}{q_{jF}/\tau}\right)^{\eta-1} + \omega_C \left(\frac{q_{jC_H}}{q_{jF}/\tau}\right)^{\eta-1} + \omega_F} \tag{D.10f}$$

In the BGP, we may identify each industry j by a pair of technology gaps $\underline{n} = (n, n_C)$, holding $\frac{q_{jH}}{q_{jF}} = (1+\lambda)^n$ and $\frac{q_{jH}}{q_{jC_H}} = \frac{q_{jH}}{q_{jC_F}} = (1+\lambda)^{n_C}$. Using these into Equations (D.10a)-(D.10f) completes the proof of Lemma 1. ■

The planner allocates equal labor to all industries, as $\sigma_H^{H*}(\underline{n}) + \sigma_{C_H}^{H*}(\underline{n}) + \sigma_F^{H*}(\underline{n}) = 1$. Labor within each industry is allocated according to a weighted average of relative qualities (the term involving a ratio of ω's), relative productivities (the term involving λ) and trade costs (the term involving τ) across firms. Statically, this allocation differs from the DE solution because of dispersion in markups between firms. Indeed, recall that the DE labor allocation (Equation (16)) holds

$$\ell_c^H(\underline{n}) = \sigma_c^H(\underline{n}) \cdot \left(\frac{\mu_c^H(\underline{n})}{\boldsymbol{\mu}}\right)^{-1} \tag{D.11}$$

where $\boldsymbol{\mu}$ is the aggregate markup, defined in Equation (19). Moreover, by definition of $\sigma_c^H(\underline{n})$ and the formula for equilibrium prices we have:

$$\sigma_H^H(\underline{n}) = \left(1 + \frac{\omega_C}{\omega_H}(1+\lambda)^{-n_C(\eta-1)}\left(\frac{\mu_H^H(\underline{n})}{\mu_{C_H}^H(\underline{n})}\right)^{\eta-1} + \frac{\omega_F}{\omega_H}\left(\frac{(1+\lambda)^{-n}}{\tau}\right)^{\eta-1}\left(\frac{\mu_H^H(\underline{n})}{\mu_F^H(\underline{n})}\right)^{\eta-1}\right)^{-1}$$

and similarly for $\sigma_{C_H}^H(\underline{n})$ and $\sigma_F^H(\underline{n})$. Notice that the only difference between $\ell_c^H(\underline{n})$ from Equation (D.11) and $\ell_c^{H*}(\underline{n})$ from Lemma 1 are the terms involving ratios of markups, i.e. the presence of markup dispersion. Indeed, the two allocations coincide when $\mu_H^H(\underline{n}) = \mu_{C_H}^H(\underline{n}) = \mu_F^H(\underline{n}) = \boldsymbol{\mu}$. Moreover, as $\mu_{C_H}^H(\underline{n}) = 1$ by assumption, then the static labor allocation in the DE coincides with the SP solution when all firms charge zero markups in all industries.

Dispersion in markups generates both within- and across-industry misallocation. To study within-industry misallocation, we can show that industry-level TFP (the ratio of industry output to industry labor use) for the DE can be written in terms of market shares and markups as follows:[1]

$$TFP(\underline{n}) = \left(\frac{\omega_H}{\sigma_H^H(\underline{n})}\right)^{\frac{1}{\eta-1}}\left(\sigma_H^H(\underline{n}) + \sigma_{C_H}^H(\underline{n})\frac{\mu_H^H(\underline{n})}{\mu_{C_H}^H(\underline{n})} + \sigma_F^H(\underline{n})\frac{\mu_H^H(\underline{n})}{\mu_F^H(\underline{n})}\right)^{-1} \tag{D.12}$$

Setting markups to zero and using $\sigma_H^H(\underline{n}) + \sigma_{C_H}^H(\underline{n}) + \sigma_F^H(\underline{n}) = 1$ gives us the TFP level in the SP allocation:

[1] To obtain this formula, we use Equation (4) and normalize $q_H = 1$ to arrive at industry output $Y(\underline{n}) = \left(\frac{\omega_H}{\sigma_H^H(\underline{n})}\right)^{\frac{1}{\eta-1}} \frac{\boldsymbol{\mu}}{\mu_H^H(\underline{n})}$. Moreover, from Equation (D.11) we have that total industry labor use is $L(\underline{n}) = \sum_{c=H,C_H,F} \sigma_c^H(\underline{n}) \left(\frac{\mu_c^H(\underline{n})}{\boldsymbol{\mu}}\right)^{-1}$. Taking the ratio gives Equation (D.12).

$$TFP^*(\underline{n}) = \left(\frac{\omega_H}{\sigma_H^{H*}(\underline{n})} \right)^{\frac{1}{\eta-1}} \tag{D.13}$$

The green surface on the left panel of Figure D.1 plots TFP losses across industries, computed as $\frac{TFP(\underline{n})}{TFP^*(\underline{n})}$, for our set of calibrated parameters. As seen in the figure, TFP in the DE is lowest in industries where markups are more dispersed, i.e. in industries where the technology gap between firms is the largest. TFP losses are entirely driven by within-industry markup dispersion: in industries with large technology gaps, static TFP losses are as high as 8%, where the planner would choose to give all production to one firm but this firms finds it more profitable to charge high markups instead. In contrast, in industries where the markup dispersion is lowest, TFP losses are close to zero. This occurs for a slightly negative value for n, as in this industry the slight technological disadvantage of the Home leader exactly offsets the trade cost disadvantage of the Foreign leader.

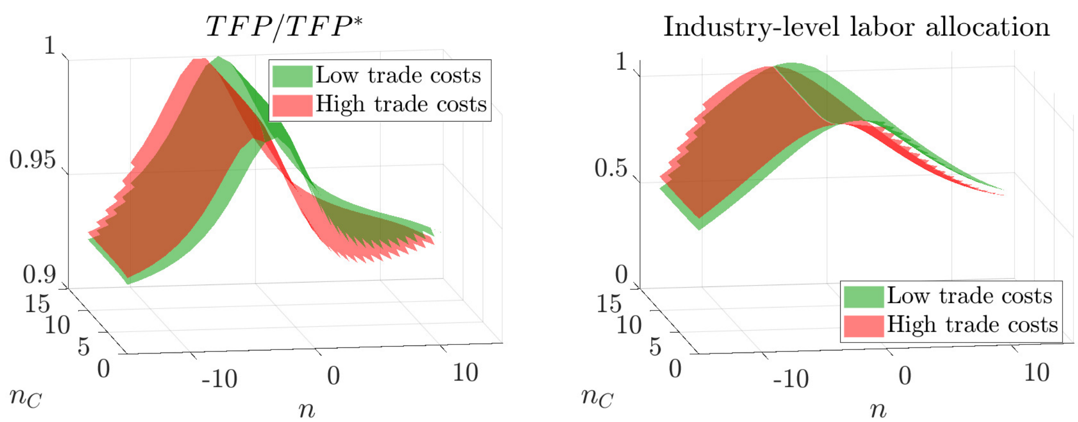

Figure D.1: Comparison between high and low trade costs for the industry TFP relative to efficient level (left) and labor allocation across industries (right). Notes: The parameter values are given in Table 1.

Labor is also misallocated across industries, as seen on the right panel of Figure D.1. As noted before, the planner wants to allocate equal labor to all industries. In the DE, however, industries with higher within-industry misallocation receive too little labor, as technological leaders in these industries are too small relative to the social optimum. This labor is allocated into neck-to-neck industries instead, where firms are inefficiently large.

Trade has interesting effects on the amount of static misallocation and labor reallocation across industries. In Figure 9 we saw that moving from high to low trade costs in BGP depresses markups for the industries with larger technological leads, thereby reducing their levels of misallocation. As seen on the right panel of Figure D.1, labor is reallocated away from industries where the Home leader is lagging the most. This occurs because, as discussed in the previous section, a decrease in trade costs increases the innovation incentives of neck-to-neck firms, leading to a shift in the composition of firms toward states with larger technology gaps.

Computing misallocation as the gap between the DE and SP initial levels of output, or $1 - \boldsymbol{Y}_0/\boldsymbol{Y}_0^*$, we find that opening up to trade (from autarky to the BGP with low trade costs) reduces this gap by 16%. By comparison, Edmond *et al.* (2015) find that opening up to trade reduces misallocation by one-fifth.

D.2 Optimal innovation policies

We now turn to characterizing the optimal innovation choices of the planner. Given the optimal labor allocation (Lemma 1), the planner chooses R&D investment for each firm, $\{z_H^*(n, n_C), x_H^*(n, n_C)\}$, and the distribution of firms across states, $\{\varphi^*(n, n_C)\}$, to maximize social welfare in BGP, given by Equation (29). The following lemma describes the optimal dynamic choices of the planner (once again, we focus on country H as solutions for F are symmetrical).

Lemma 2 *The dynamic choices of the constrained SP in country H satisfy*

$$\frac{\chi_i \psi_i \left[z_H^*(n, n_C)\right]^{\psi_i - 1}}{1 - \boldsymbol{r}^{H*}} = \frac{\chi_e \psi_e \left[x_H^*(n, n_C)\right]^{\psi_e - 1}}{1 - \boldsymbol{r}^{H*}} = \frac{\ln(1+\lambda)}{\rho} + \nu(n+1, n_C+1) - \nu(n, n_C) \quad \text{(D.14)}$$

where $\boldsymbol{r}^{H*}$ is the aggregate R&D share of GDP (from Equation (28)), and $\nu(n, n_C)$ is the shadow value of an innovation.

Proof. We solve for the planner's solution under balanced growth, so aggregate world output must grow at a constant rate, g^W. The objective function is

$$\int_0^{+\infty} e^{-\rho t} \ln(\boldsymbol{C}_t^W) \mathrm{d}t = \frac{1}{\rho}\left(\ln(\boldsymbol{C}_0^W) + \frac{g^W}{\rho}\right)$$

as $\boldsymbol{C}_t^W = \boldsymbol{C}_0^W e^{g^W t}$. The planner chooses:

1. An initial consumption level, $\boldsymbol{C}_0^W$, and a growth rate, g^W.
2. Innovation intensities in each country: $\Big\{\big(z_k(n, n_C), x_k(n, n_C)\big) : (n, n_C) \in \mathbb{Z} \times \mathbb{Z}_+, k = H, F\Big\}$.
3. A stationary distribution of firms across states in each country: $\Big\{\varphi_k(n, n_C) : (n, n_C) \in \mathbb{Z} \times \mathbb{Z}_+, k = H, F\Big\}$.

The planner is subject to the following constraints:

- First, the planner recognizes the symmetric nature of the technology gap distribution: for every leader that is ahead in one country and given industry, there must be a leader that is behind in the other country in the same industry. Formally:

  $$\varphi_H(n, n_C) = \varphi_F(-n, n_C - n)$$

  Given this symmetry, henceforth we will write $\varphi(n, n_C)$ generically to speak about the distribution from the point of view of a domestic firm in H.

- Second, the planner takes into account each country's resource constraint: $\frac{\boldsymbol{C}_t^k}{\boldsymbol{Y}_t^k} + \frac{\boldsymbol{R}_t^k}{\boldsymbol{Y}_t^k} \leq 1$, where:

$$\boldsymbol{r}^k \equiv \frac{\boldsymbol{R}_t^k}{\boldsymbol{Y}_t^k} = \sum_n \sum_{n_C} \varphi(n, n_C) \left[\chi_i \left(z_k(n, n_C)\right)^{\psi_i} + \chi_e \left(x_k(n, n_C)\right)^{\psi_e}\right] \tag{D.15}$$

Therefore, $\ln \boldsymbol{C}_t^W = \ln \boldsymbol{Y}_t^W + \frac{1}{2} \sum_{k=H,F} \ln(1 - \boldsymbol{r}^k)$, where $\ln \boldsymbol{Y}_t^W = \ln \boldsymbol{Y}_0^W + g^W t$, and:

$$\ln \boldsymbol{Y}_0^W = \frac{1}{2} \sum_k \sum_n \sum_{n_C} \varphi(n, n_C) \ln Y^k(n, n_C)$$

from the static planner solution (Equation (D.8)). Using (D.7a)-(D.7b) and normalizing $q_H = 1$, we have:

$$Y^H(n, n_C) = \left[\omega_H + \omega_C \big(1+\lambda\big)^{-n_C(\eta-1)} + \omega_F \left(\frac{(1+\lambda)^{-n}}{\tau}\right)^{\eta-1}\right]^{\frac{1}{\eta-1}}$$

$$Y^F(n, n_C) = \left[\omega_F (1+\lambda)^{-n(\eta-1)} + \omega_C \big(1+\lambda\big)^{(n-n_C)(\eta-1)} + \omega_H \left(\frac{1}{\tau}\right)^{\eta-1}\right]^{\frac{1}{\eta-1}}$$

- Third, the planner takes into account two laws of motion: (i) the law of motion for the φ distribution (written below in the planner's program); and (ii) the law of motion for aggregate output, $\dot{\boldsymbol{Y}}_t^W = \boldsymbol{Y}_t^W g^W$. For the latter, we have:

$$g^W \equiv \frac{\mathrm{d}}{\mathrm{d}t} \ln \boldsymbol{Y}_t^W = \frac{1}{2} \Bigg(\underbrace{\frac{\partial}{\partial t} \int_0^1 \ln Y_j^H \mathrm{d}j}_{\equiv g^H} + \underbrace{\frac{\partial}{\partial t} \int_0^1 \ln Y_j^F \mathrm{d}j}_{\equiv g^F} \Bigg) = \frac{g^H + g^F}{2}$$

where

$$\begin{aligned}
\int_0^1 \ln Y_j^H \mathrm{d}j &= \int_0^1 \left(\frac{1}{\eta-1}\right) \ln \left(\omega_H (q_{jH})^{\eta-1} + \omega_C (q_{jC_H})^{\eta-1} + \omega_F \left(\frac{q_{jF}}{\tau}\right)^{\eta-1}\right) \mathrm{d}j \\
&= \int_0^1 \left(\frac{1}{\eta-1}\right) \left[\ln \left((q_{jH})^{\eta-1}\right) + \ln \left(\omega_H + \omega_C \left(\frac{q_{jC_H}}{q_{jH}}\right)^{\eta-1} + \omega_F \left(\frac{q_{jF}/\tau}{q_{jH}}\right)^{\eta-1}\right)\right] \mathrm{d}j \\
&= \int_0^1 \ln(q_{jH}) \mathrm{d}j + \int_0^1 \left(\frac{1}{\eta-1}\right) \left[\ln \left(\omega_H + \omega_C \left(\frac{q_{jC_H}}{q_{jH}}\right)^{\eta-1} + \omega_F \left(\frac{q_{jF}/\tau}{q_{jH}}\right)^{\eta-1}\right)\right] \mathrm{d}j
\end{aligned}$$

and similarly for F. In a balanced-growth solution, the second additive term on the last line is constant. Thus, following the same logic as in the proof of Lemma 1:

$$g^W = \frac{\ln(1+\lambda)}{2} \sum_k \sum_n \sum_{n_C} \varphi(n, n_C) i_k(n, n_C)$$

We are now ready to solve the dynamic problem:

$$\max_{\substack{\{z_H(n,n_C),x_H(n,n_C)\}\\ \{z_F(n,n_C),x_F(n,n_C)\}\\ \{\varphi(n,n_C)\}}} \left\{ \sum_k \sum_n \sum_{n_C} \varphi(n,n_C) \ln Y^k(n,n_C) + \frac{\ln(1+\lambda)}{\rho} \sum_k \sum_n \sum_{n_C} \varphi(n,n_C) i_k(n,n_C) \right.$$
$$\left. + \sum_k \ln \left(1 - \sum_n \sum_{n_C} \varphi(n,n_C) \left[\chi_i \left(z_k(n,n_C)\right)^{\psi_i} + \chi_e \left(x_k(n,n_C)\right)^{\psi_e} \right] \right) \right\}$$

subject to:

$$0 = -\varphi(n,0) \sum_k i_k(n,0) + \zeta\varphi(n,1) + i_F(n+1,0)\varphi(n+1,0) \tag{D.16a}$$

$$\forall n_C \geq 1: \; 0 = -\varphi(n,n_C) \left(\sum_k i_k(n,n_C) + \zeta \right) + \zeta\varphi(n,n_C+1) + i_H(n-1,n_C-1)\varphi(n-1,n_C-1) + i_F(n+1,n_C)\varphi(n+1,n_C) \tag{D.16b}$$

$$0 \leq \varphi(n,n_C), z_H(n,n_C), x_H(n,n_C), z_F(n,n_C), x_F(n,n_C) \tag{D.16c}$$

$$1 = \sum_n \sum_{n_C} \varphi(n,n_C); \tag{D.16d}$$

where we have defined the total innovation rates of each country as $i_H(n,n_C) \equiv z_H(n,n_C) + x_H(n,n_C) + \xi \mathbb{1}_{n<0}$ and $i_F(n,n_C) \equiv z_F(n,n_C) + x_F(n,n_C) + \xi \mathbb{1}_{n>0}$.

The Lagrangian is:

$$\mathcal{L} = \sum_k \sum_n \sum_{n_C} \varphi(n,n_C) \ln Y^k(n,n_C) + \frac{\ln(1+\lambda)}{\rho} \sum_k \sum_n \sum_{n_C} \varphi(n,n_C) i_k(n,n_C)$$
$$+ \sum_k \ln \left(1 - \sum_n \sum_{n_C} \varphi(n,n_C) \left[\chi_i \left(z_k(n,n_C)\right)^{\psi_i} + \chi_e \left(x_k(n,n_C)\right)^{\psi_e} \right] \right)$$
$$+ \sum_{n=-\infty}^{+\infty} \nu(n,0) \left[-\varphi(n,0) \sum_k i_k(n,0) + \zeta\varphi(n,1) + i_F(n+1,0)\varphi(n+1,0) \right]$$
$$+ \sum_{n=-\infty}^{+\infty} \sum_{n_C=1}^{+\infty} \nu(n,n_C) \Bigg[-\varphi(n,n_C) \left(\sum_k i_k(n,n_C) + \zeta \right) + \zeta\varphi(n,n_C+1)$$
$$+ i_H(n-1,n_C-1)\varphi(n-1,n_C-1) + i_F(n+1,n_C)\varphi(n+1,n_C) \Bigg]$$
$$- \sum_{n=-\infty}^{+\infty} \sum_{n_C=0}^{+\infty} \left[\vartheta^{\varphi}(n,n_C)\varphi(n,n_C) + \sum_k \left(\vartheta^{z_k}(n,n_C) z_k(n,n_C) + \vartheta^{x_k}(n,n_C) x_k(n,n_C) \right) \right]$$

$$+ \phi \left[1 - \sum_{n} \sum_{n_C} \varphi(n, n_C)\right]$$

where $\nu(n, n_C) \geq 0$ is the multiplier on (D.16a)-(D.16b); $\vartheta^{\varphi}(n, n_C), \vartheta^{z_k}(n, n_C), \vartheta^{x_k}(n, n_C) \geq 0$ are the multipliers on (D.16c); and $\phi \geq 0$ is the multiplier on (D.16d).

The first-order conditions for z_k and x_k yield:

$$\frac{\chi_i \psi_i \left[z_k(n, n_C)\right]^{\psi_i - 1}}{1 - \boldsymbol{r}^k} + \frac{\vartheta^{z_k}(n, n_C)}{\varphi(n, n_C)} = \frac{\ln(1+\lambda)}{\rho} + \Delta_k(n, n_C) \tag{D.17a}$$

$$\frac{\chi_e \psi_e \left[x_k(n, n_C)\right]^{\psi_e - 1}}{1 - \boldsymbol{r}^k} + \frac{\vartheta^{x_k}(n, n_C)}{\varphi(n, n_C)} = \frac{\ln(1+\lambda)}{\rho} + \Delta_k(n, n_C) \tag{D.17b}$$

respectively, where we have defined:

$$\Delta_k(n, n_C) \equiv \begin{cases} \nu(n+1, n_C+1) - \nu(n, n_C) & \text{if } k = H \\ \nu(n-1, n_C) - \nu(n, n_C) & \text{if } k = F \end{cases}$$

and $\boldsymbol{r}^k$ is given by Equation (D.15). The first-order condition for $\varphi(n, n_C)$ yields:

$$\sum_k \left(\frac{\chi_i \left[z_k(n, n_C)\right]^{\psi_i}}{1 - \boldsymbol{r}^k} + \frac{\chi_e \left[x_k(n, n_C)\right]^{\psi_e}}{1 - \boldsymbol{r}^k} \right) = \sum_k \ln Y^k(n, n_C) + \frac{\ln(1+\lambda)}{\rho} \sum_k i_k(n, n_C)$$
$$+ \sum_k i_k(n, n_C) \Delta_k(n, n_C) + \zeta \Big(\nu(n, n_C - 1) - \nu(n, n_C) \Big) \mathbb{1}_{n_C \geq 1} - \vartheta^{\varphi}(n, n_C) - \phi \tag{D.18}$$

Equations (D.17a), (D.17b) and (D.18) can then be solved to find the socially optimal innovation policies and invariant distribution. The results in Lemma 2 correspond to the interior equilibrium of the above optimality conditions, where $\vartheta^{z_k} = \vartheta^{x_k} = \vartheta^{\varphi} = 0$. ■

Lemma 2 says that the planner wants to equate the marginal cost of both types of R&D to the marginal gain from innovating. The marginal cost comes from direct R&D costs, and indirect costs from the fact that innovating takes resources away from consumption (as $\boldsymbol{r}^{H*}$ is the R&D share, the term $1 - \boldsymbol{r}^{H*}$ equals the aggregate consumption share of GDP). The social marginal gains, which are independent of whether an innovation is performed by an incumbent or an entrant, are composed of two terms: (i) the discounted permanent increase in productivity, given by $\ln(1+\lambda)/\rho$; and (ii) the change in the shadow value of innovations, given by $\nu(n+1, n_C+1) - \nu(n, n_C)$.

In contrast, the DE solution, given by Equations (24) and (25), equates the private cost of R&D to the private benefit. There are two differences between the social and private return to innovation: (i) firms do not internalize that future innovators will benefit from their own innovations (positive externality), and (ii) firms do not internalize that part of their (private) gains from innovation is associated with a decrease in the value of other firms through business stealing (negative externality).

References

Akcigit, U. and Kerr, W. R. (2018). Growth through heterogeneous innovations. *Journal of Political Economy*, **126** (4), 1374–1443.

Bernard, A. B., Jensen, J. B., Redding, S. J. and Schott, P. K. (2018). Global Firms. *Journal of Economic Literature*, **56** (2), 565–619.

Edmond, C., Midrigan, V. and Xu, D. Y. (2015). Competition, Markups, and the Gains from International Trade. *American Economic Review*, **105** (10), 3183–3221.

Harris, R. and Moffat, J. (2011). *R&D, Innovation and Exporting*. SERC Discussion Papers 0073, Spatial Economics Research Centre, LSE.