Production, Trade, and Cross-Border Data Flows

Qing Chang[1] Lin William Cong[2] Liyong Wang[3] Longtian Zhang[3]

1. School of Social Sciences, Tsinghua University
2. Cornell University SC Johnson College of Business and NBER
3. School of International Trade and Economics, Central University of Finance and Economics

American Finance Association 2026 Annual Meeting

January 4, 2026

Outline

Introduction

Motivating Facts

- Data as the "new oil" in the information age.
- The unrestricted flow of data factor truly creates values.
- Countries across every stage of development have recently erected barriers to them.
  - "data localization": protect local economy, privacy, and national security
- No unified global rules on cross-border data flows.
  - US: supports "free flow of data with trust".
  - EU: prioritizes privacy, restricting data flows via GDPR.
  - China: emphasizes data sovereignty and national security.
- The intense competition between the U.S. and China in AI and digital economy (strategic interactions).

Introduction

Overview

A general equilibrium model of production and trade in a global economy.

- Comparisons:
  - A closed economy;
  - Partially open economies;
    - with only goods trade,
    - with goods trade and unilateral data flows.
  - Open economies with free data flows.
- Data
  - generated as byproducts of the consumption of households.
  - play crucial roles as input factors in production.

Introduction

Key Results

- Welfare analysis:
  - Latecomer's advantage: International data flows significantly improve welfare, especially for countries more backward in data economy.
  - Facilitate trade, when the data divide between the two countries is not very large;
  - Working data (useful information) vs. Raw data
    - Data-intensive country uses more working data,
    - Data-light country contributes more raw data.
- The transition dynamics:
  - After a productivity shock, the cyclicality of working data reverses, from closed to open economies;
  - shocks to data privacy and to flow costs have opposite effects on domestic and foreign data sectors.

- Data divde trap:
  - Though the country with low importance of data can gain large welfare improvements from trading goods and data flows in most cases,
  - the foreign country with high data importance refuses to trade since it faces a welfare loss.
  - The developing country in data economy should keep up with the pace of the developed country.

- Different paths to the economics of data:
  - Jones and Tonetti (2020): Horizontal nonrivalry & production process;
  - Cong, Xie, and Zhang (2021): Dynamic nonrivalry & innovation process;
  - Cong, Wei, Xie, and Zhang (2022): Vertical nonrivalry & both production process and innovation process simultaneously;
  - Xie and Zhang (2023): "producer data" lead to higher growth rate than "consumer data."
- Data do not always lead to sustained economic growth: Farboodi and Veldkamp (2020, 2022); Hou et al. (2025); Veldkamp (2023).

- Data privacy: Abowd and Schmutte (2019); Fainmesser et al. (2019); Ichihashi (2020).
- Data intermediary and markets: Bergemann and Bonatti (2019); Acemoglu et al. (2022); Ichihashi (2021a, b).
- Valuing data: Sun et al. (2021); Farboodi et al. (2022); Veldkamp (2023).

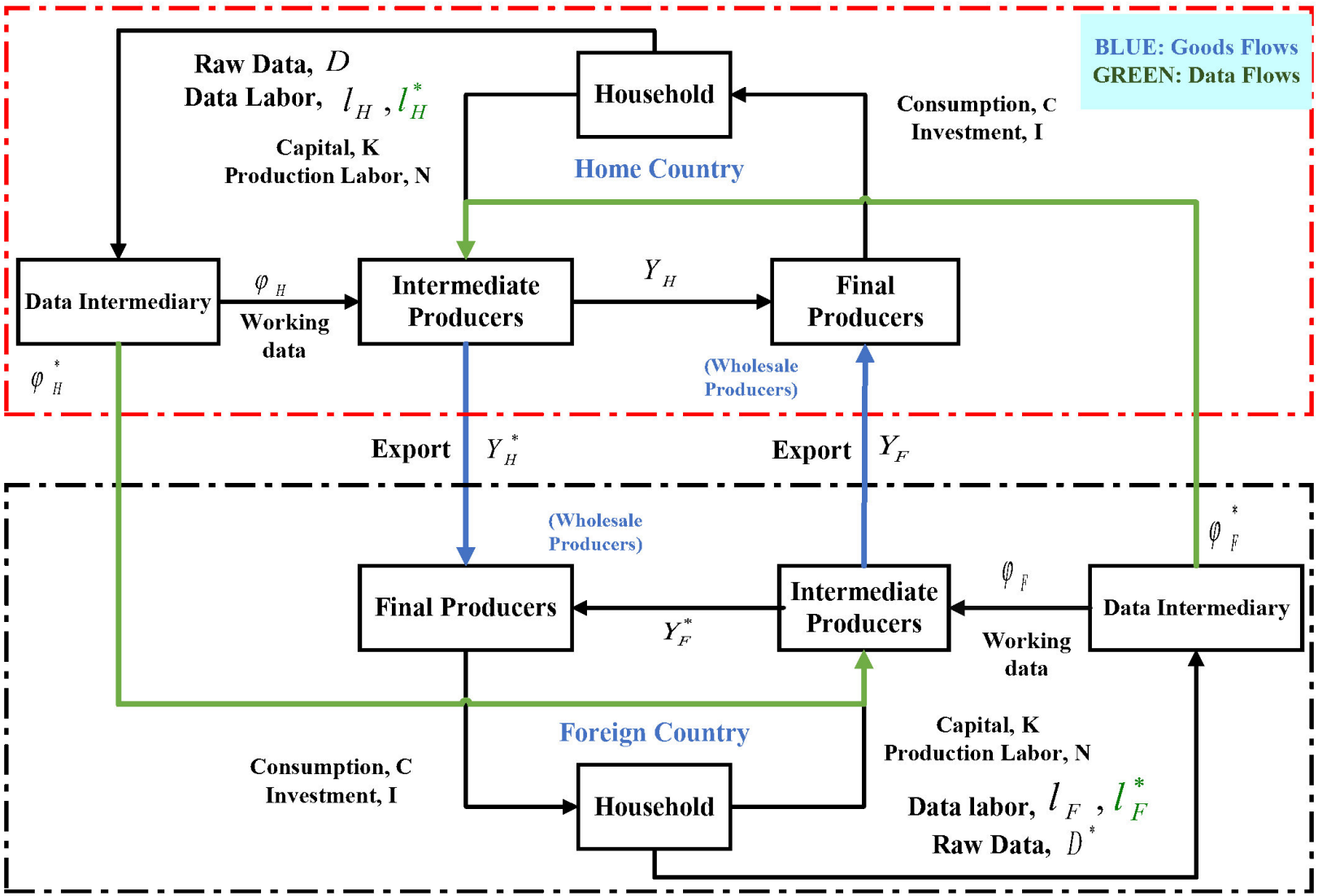
The Model

Setup

- Representative household
  - consumption C_t, labor N_t, and raw data D_t
- Final good producer
  - packs up intermediate goods (from wholesale) to produce output Y_t
- Wholesale good producers
  - assemble domestic and imported intermediate goods ($Y_{H,i,t}$ and $Y_{F,i,t}$) to produce wholesale goods $Y_{i,t}$
- Intermediate good producers (monopolistic with variety i)
  - produce goods $Y_{i,t}$ using labor n_t, capital K_t and data Φ_t
- Data intermediary (monopolist constrained by free entry, i.e. zero profit)
  - buy raw data D_t from household and sell working data φ_t ($\varphi_{H,t}$ and $\varphi^*_{H,t}$) to intermediate producers

A Model of Data Economy

Setup (Cont.)



Utility function:

$$U = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\sigma}}{1-\sigma} - \Omega \frac{N_t^{1+\eta}}{1+\eta} - \Pi(1 + b\mathbb{I})\pi_t D_t^2 \right],$$

- π_t is the disutility shock of consumer's data risk, following AR(1):

$$\ln \pi_t - \ln \pi = \rho_\pi (\ln \pi_{t-1} - \ln \pi) + \sigma_\pi \varepsilon, \quad \varepsilon \sim N(0,1),$$

- $\mathbb{I}$ is an indicator function
  - $\mathbb{I} = 1$ when there are cross-border data flows (exporting).
  - In the closed economy, $\mathbb{I} = 0$.

The utility function:

$$U = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\sigma}}{1-\sigma} - \Omega \frac{N_t^{1+\eta}}{1+\eta} - \Pi(1 + b\mathbb{I})\pi_t D_t^2 \right],$$

- Budget constraint: (We normalize the price of final goods in the home country to 1)

$$\begin{aligned} C_t + I_t + B_{H,t} + B_{F,t} =& w_t N_t + r_t K_t + R_{t-1} B_{H,t-1} + R_{t-1}^* B_{F,t-1} \\ &+ P_{D,t} D_t + F_t. \end{aligned}$$

- Capital accumulation:

$$K_{t+1} = (1 - \delta_K) K_t + I_t.$$

- The final good producer:

$$Y_t = \left(\int_0^1 Y_{i,t}^{\frac{\rho-1}{\rho}} \mathrm{d}i \right)^{\frac{\rho}{\rho-1}},$$

  - $Y_{i,t}$: the wholesale good of variety i (composite intermediate good).
  - Given the prices of each individual variety $P_{i,t}$, the profit maximization problem :

$$\max_{Y_{i,t}} Y_t - \int_0^1 P_{i,t} Y_{i,t} \mathrm{d}i.$$

A Two-country Open Economy

Final Good Producer and Wholesale Producers (Cont.)

- Wholesale producers: assemble goods produced domestically and goods imported from abroad using CES technology:

$$Y_{i,t} = \left(Y_{H,i,t}^{\frac{m-1}{m}} + Y_{F,i,t}^{\frac{m-1}{m}} \right)^{\frac{m}{m-1}}. \tag{1}$$

  - A country may impose restrictions on imported goods to protect its domestic firms (e.g., tariffs).
  - Firms in the home country must pay a markup of d_P to purchase these goods.

- The profit maximization problem for the wholesale producers:

$$\max_{Y_{H,i,t}, Y_{F,i,t}} \; Y_{i,t} - P_{H,i,t} Y_{H,i,t} - d_P P_{F,i,t} Y_{F,i,t},$$

where $Y_{i,t}$ is shown in (1).

A Two-country Open Economy

Intermediate Good Producers

- Each producer produces outputs both for domestic use and for exporting, according to the following technology:

$$Y_{H,i,t} + Y^*_{H,i,t} = A_t(\Phi^{\xi}_{i,t} k_{i,t})^{\alpha} n^{1-\alpha}_{i,t}, \tag{2}$$

- A_t is the productivity level, following AR(1) process:

$$\ln A_t - \ln A = \rho_A(\ln A_{t-1} - \ln A) + \sigma_A \varepsilon, \quad \varepsilon \sim N(0,1).$$

- The accumulation process of the data is shown as:

$$\Phi_{i,t+1} = (1-\delta_\Phi)\Phi_{i,t} + \varphi_t, \tag{3}$$

- φ_t here represents a data composite:

$$\varphi_t = \left[\chi^{\frac{1}{\omega}} \varphi_{H,t}^{\frac{\omega-1}{\omega}} + (1-\chi)^{\frac{1}{\omega}} \varphi_{F,t}^{\frac{\omega-1}{\omega}}\right]^{\frac{\omega}{\omega-1}}. \tag{4}$$

- The trade cost of importing data.
- Suppose that $P_{\varphi,F,t}$ is the price the foreign country sells its working data to the home country, then the price that the home country should in fact pay is $d_D^* f_t P_{\varphi,F,t}$.
- Here, d_D^* is viewed as the trade cost of imported foreign working data, and f_t is the shock of this cost, following AR(1) process:

$$\ln f_t - \ln f = \rho_f(\ln f_{t-1} - \ln f) + \sigma_f \varepsilon, \quad \varepsilon \sim N(0,1).$$

A Two-country Open Economy

Data Intermediary

- A data intermediary in each country buys raw data from household in domestic country and then sells working data to intermediate producers in both countries.
- It also employ labors ($l_{H,t}$ and $l^*_{H,t}$) to do the collecting and cleaning works, to produce working data for domestic usage and exporting.
- The working data generation functions in the home country:

$$\varphi_{H,t} = BD_t^{\gamma} l_{H,t}^{1-\gamma}, \tag{5}$$

and

$$\varphi^*_{H,t} = BD_t^{\gamma} (l^*_{H,t})^{1-\gamma}. \tag{6}$$

Equilibrium Definition

An allocation in which (both in the home country and foreign country):

- Households maximize the utility,
- Producers and data intermediaries maximize the profit,
- All markets clear.

  (Final goods market, wholesale goods market, intermediate goods market, labor market, capital market, assets market, and data market.)

Alternative Models

Goods Trade Economy, Partially Open Economies, and Closed Economy

- Goods trade model:
  - Intermediate good producers: $\varphi_{H,t} = \varphi_t$.
  - The data intermediary:

$$\max_{D_{H,t}, l_{H,t}} \int_0^1 P_{i,\varphi,H,t}\varphi_{H,t}\mathrm{d}i - P_{D,H,t}D_t - w_t l_{H,t},$$

subject to the working data generation function:

$$\varphi_{H,t} = BD_t^{\gamma} l_{H,t}^{1-\gamma},$$

and the zero-profit condition:

$$\int_0^1 P_{i,\varphi,H,t}\varphi_{H,t}\mathrm{d}i - P_{D,H,t}D_t - w_t l_{H,t} = 0.$$

- Closed economy and economies with unilateral data flow for comparison (without goods traded or data flows).

Steady States Solution

Parameters

Parameters	Meaning	Value	Source
β	Subjective discount factor	0.99	Standard
σ	Reciprocal of elasticity of intertemporal substitution	2	Standard
η	Reciprocal of Frisch labor supply elasticity	1.3	Standard
$1-\alpha$	Contribution of labor in good productions	$2/3$	Standard
δ_k	Capital depreciation rate	0.025	Standard
Ω	Weight on leisure in the utility function	1.315	Christensen and Dib (2008)
ρ	Elasticity of substitution (varieties)	21	Fernandez-Villaverde et al. (2015)
m	Elasticity of substitution (domestic and imported)	5	Alessandria et al. (2021)
ρ_A, ρ_e, ρ_f	Persistence of exogenous shocks	0.95	Alessandria et al. (2013)
δ_Φ	Data depreciation rate	0.025	Discretionary
χ	Availability of domestic data	0.5	Discretionary
B	Efficiency term in working data generation	1	Discretionary
γ	Contribution of raw data in working data generation	0.5	Discretionary

Steady States Solution

Welfare Analyses

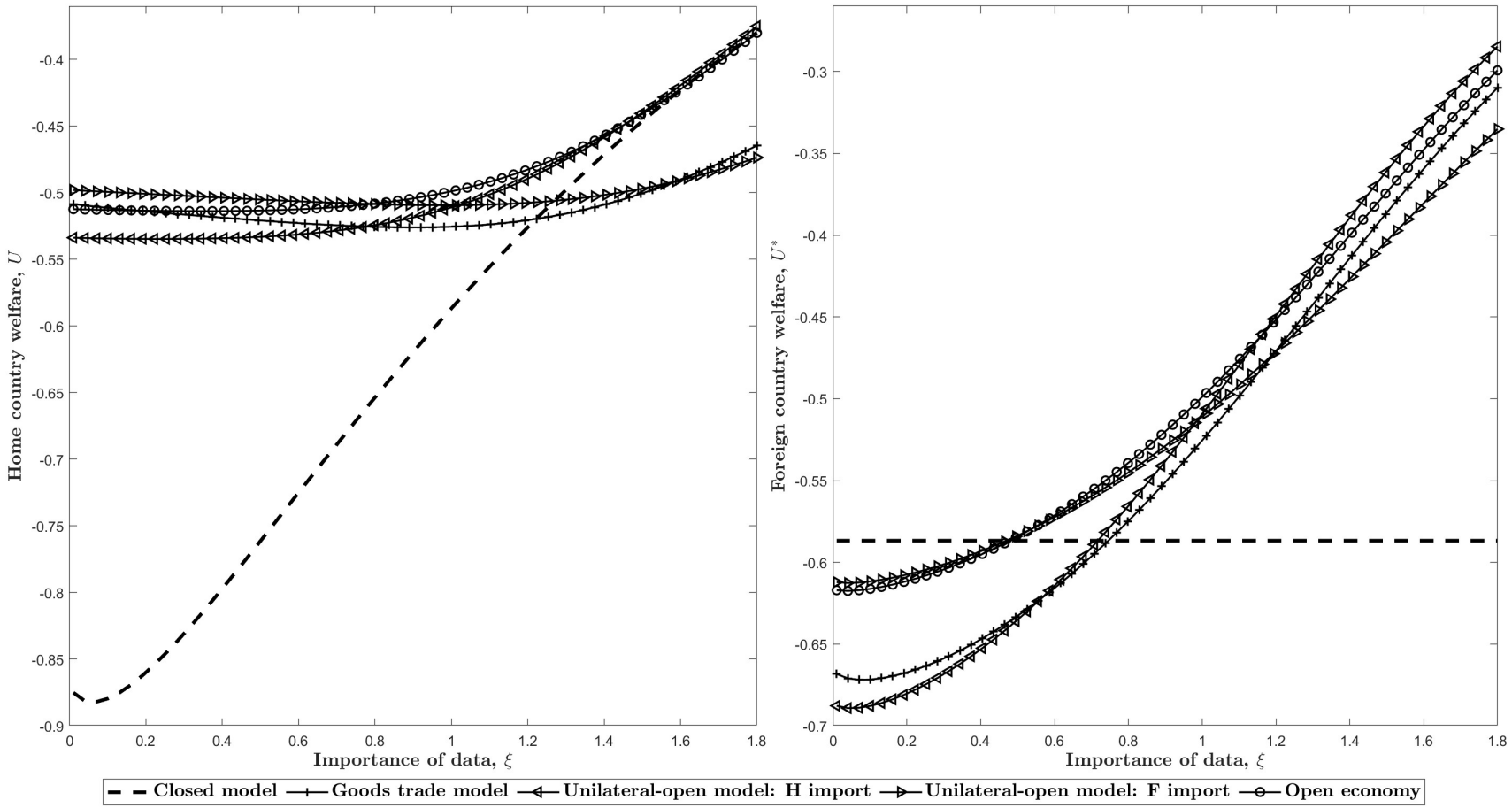


Figure: Welfare Levels with Different Importance of Data in the Two Countries (Fixing $\xi^* = 1$)

Steady States Solution

Welfare Analyses (Cont.)

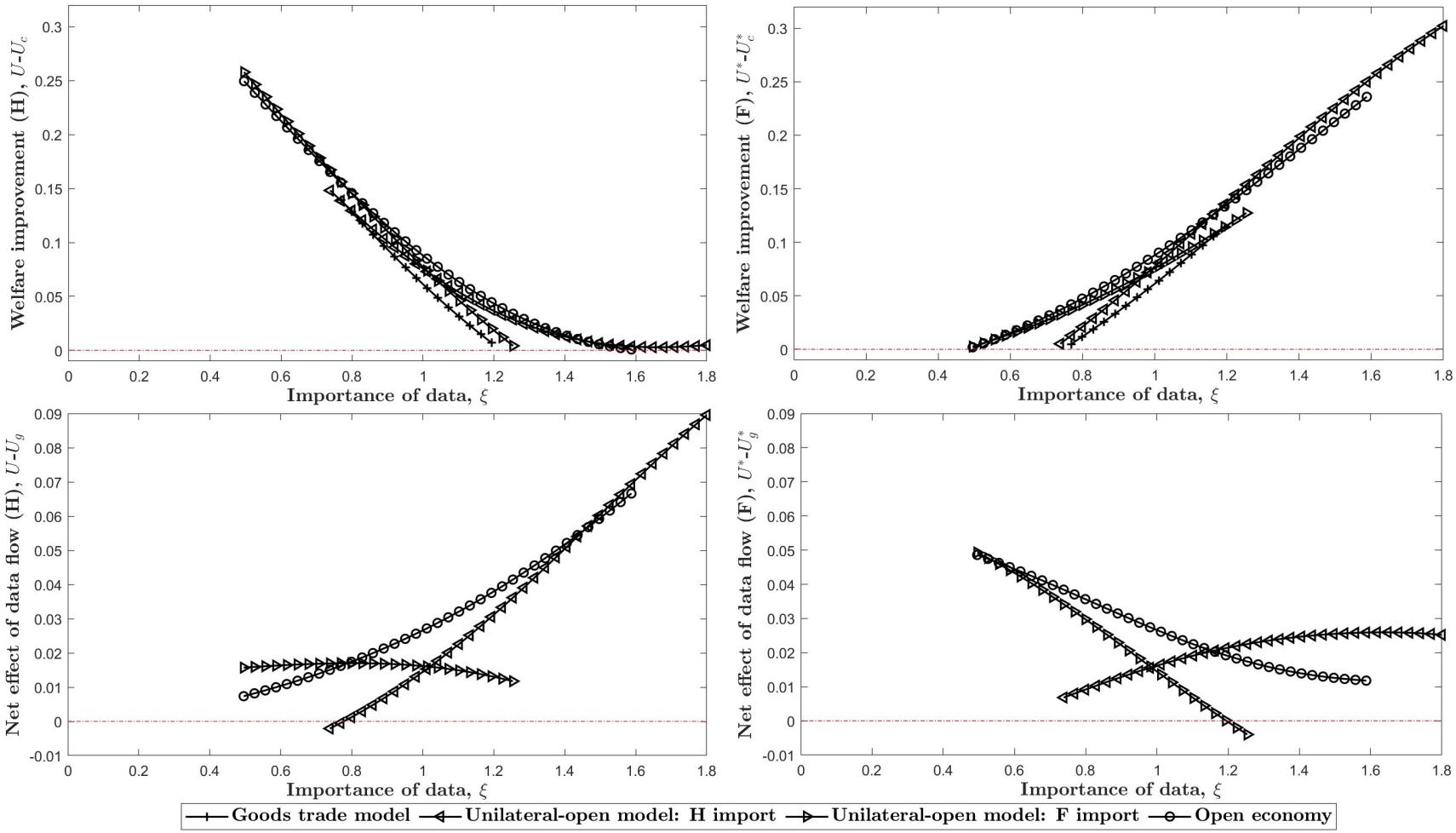


Figure: Welfare Improvements from Trade and Data Flows under Different Importance of Data (Fixing $\xi^* = 1$)

Steady States Solution

Welfare Analyses (Cont.)

- The importance of data in the foreign country is fix at $\xi^* = 1.0$.
  - In the home country, welfare improvements are very large when ξ is small.
  - However, in the foreign country, we see welfare loss in these cases.
    - The foreign country will refuse to trade.
- Intuitions
  - Benefits of trade: market expansion and demand increasing
  - Costs for high ξ country: the price of its goods become cheaper
    - High ξ country export more (cheap), but import less (expensive).
    - Goods from different countries are substitutes.
    - In extreme cases where data divide is too large, welfare loss.

Steady States Solution

Welfare Analyses (Cont.)

	Model	Foreign Country: Importance of Data, ξ^*		
		$\xi^* = 0.5$	$\xi^* = 1.0$	$\xi^* = 1.5$
Home Country: Available Interval of Trade, ξ. ($U - U_c > 0$ and $U^* - U_c^* > 0$)	Goods Trade Model	$[0, 0.79]$	$[0.75, 1.22]$	$[1.31, 1.67]$
	Unilateral-Open Model: F import	$[0, 0.80]$	$[0.47, 1.27]$	$[0.88, 1.75]$
	Unilateral-Open Model: H import	$[0, 1.03]$	$[0.72, 1.75]$	$[1.23, 1.75]$
	Open Economy: $d_D^* = 1$	$[0, 1.02]$	$[0.48, 1.61]$	$[0.93, 1.75]$
	Open Economy: $d_D^* = 1.5$	$[0, 0.91]$	$[0.49, 1.39]$	$[0.94, 1.75]$
	Open Economy: $d_D^* = 2$	$[0, 0.87]$	$[0.50, 1.33]$	$[0.96, 1.75]$
Home Country: Positive Net Effect Interval of Data Flow, ξ. ($U - U_g > 0$ and $U^* - U_g^* > 0$)	Unilateral-Open Model: F import	$[0, 0.73]$	$[0.47, 1.19]$	$[0.88, 1.61]$
	Unilateral-Open Model: H import	$[0.29, 1.03]$	$[0.78, 1.75]$	$[1.37, 1.75]$
	Open Economy: $d_D^* = 1$	$[0.13, 1.02]$	$[0.48, 1.61]$	$[0.93, 1.75]$
	Open Economy: $d_D^* = 1.5$	$[0.15, 0.91]$	$[0.49, 1.39]$	$[0.94, 1.75]$
	Open Economy: $d_D^* = 2$	$[0.17, 0.87]$	$[0.50, 1.33]$	$[0.96, 1.75]$

Steady States Solution

Welfare Analyses (Cont.)

- Further observations
  - Data flows help mitigate the welfare loss of high ξ country.
    - Available interval becomes larger.
  - Positive net effect of data flows in most of the cases.
    - Unilateral data flows to foreign country also have positive net effects.
    - Unilateral data flows to home country enlarges the interval.
  - Increasing trade costs of data reduces the feasible interval.
    - The interval of positive net effect of data flows also become narrower.

Steady States Solution

Gaps between Asymmetric Countries

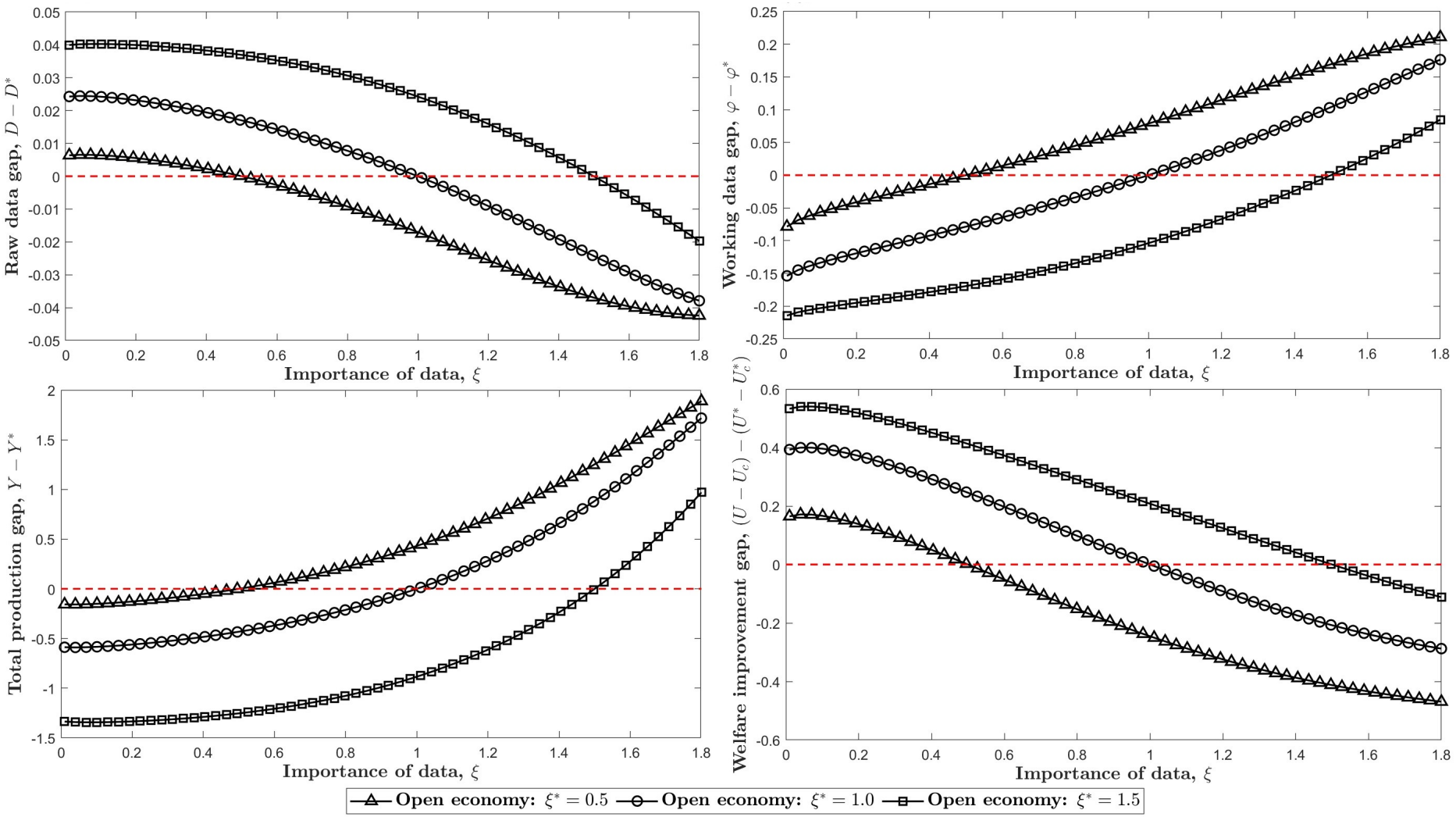


Figure: Gaps between Asymmetric Countries with Different ξ

Steady States Solution

Gaps between Asymmetric Countries (Cont.)

- Positive raw data gap in the home country (low ξ).
  - Home country provides more raw data.
  - Larger privacy costs.
- Positive working data gap and production gap in the foreign country (high ξ).
  - Foreign country uses more working data to produce,
  - produce more final goods.
- Positive welfare improvement gap in the home country (low ξ).
  - However, the economy is not realizable if data divide is too large.

Steady States Solution

Policy Implications

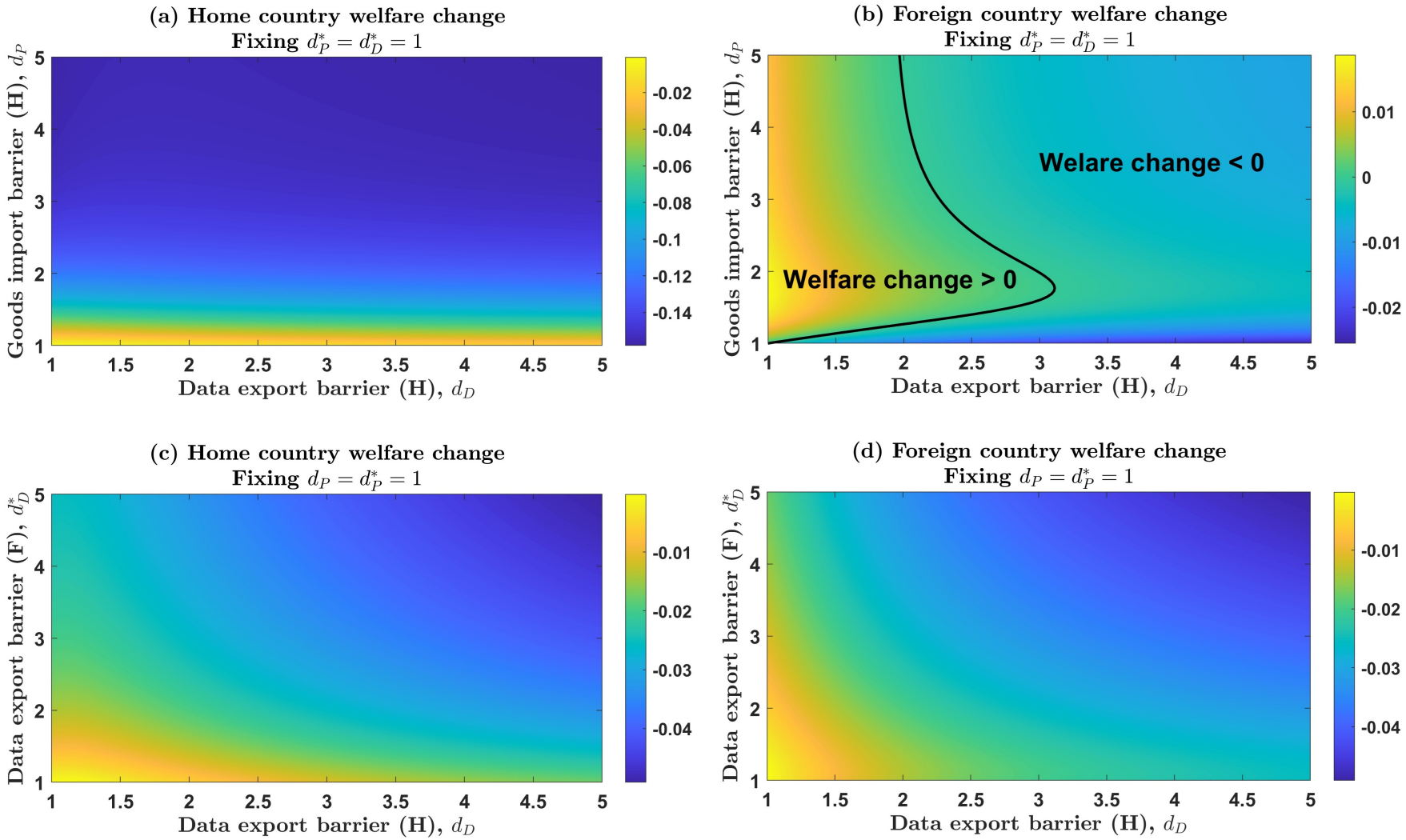


Figure: Welfare Changes With Different Combinations of Policies

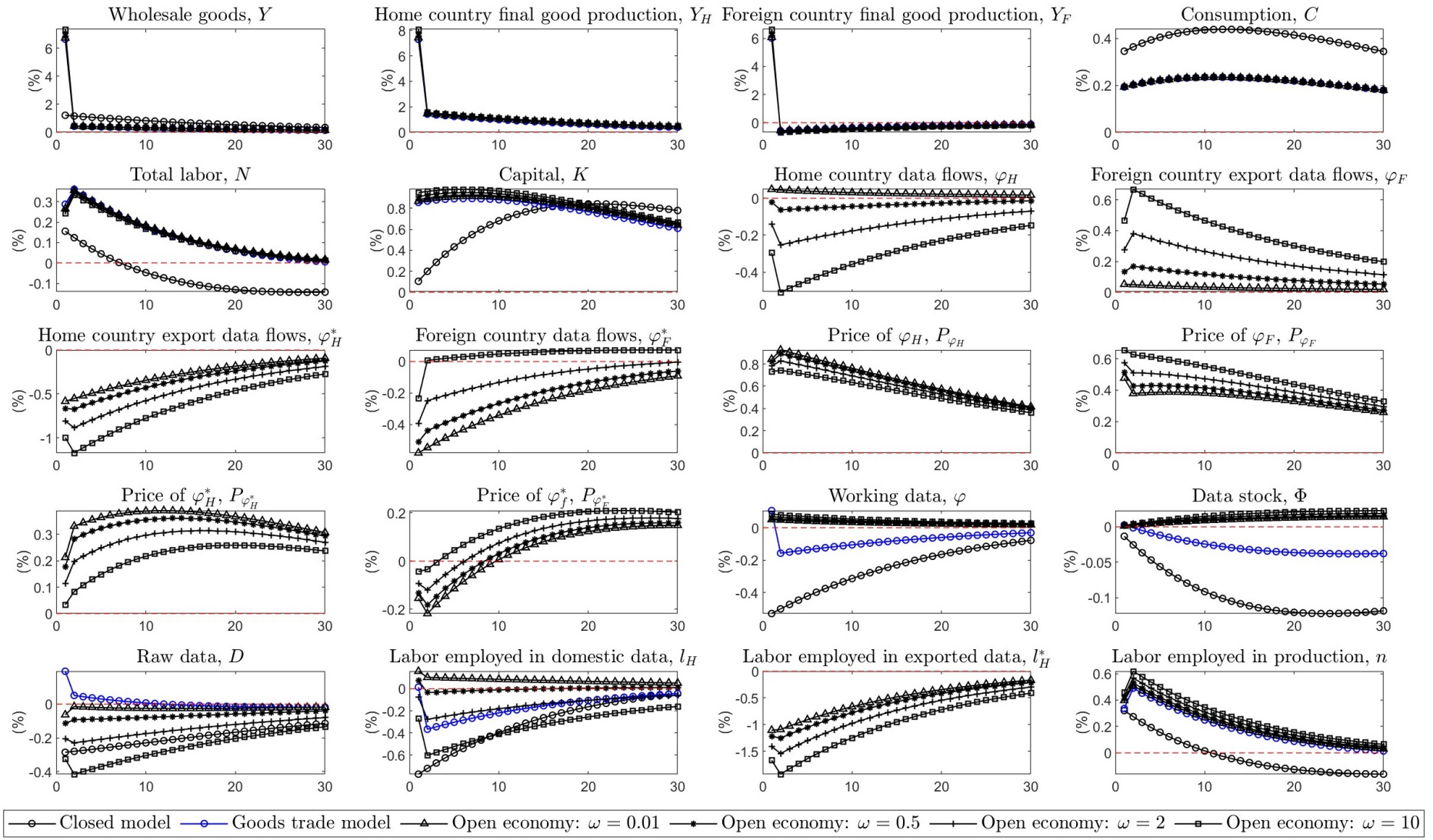
Steady States Solution

Policy Implications (Cont.)

- A country incurs welfare loss if it imposes high barriers (both trade and data flow).
  - In some scenarios, its counterpart may even benefit.
- The impact of data trade policy is relatively insignificant compare to the goods trade policy.
  - It is always optimal to raise a low good trade barrier.
  - When the counterpart country imposes a high data export barrier, it is optimal for the home country to adopt a low level.
  - Considering other perspectives, policy coordination among countries is very important.

Transition Dynamics

Analysis on Symmetric Countries—Productivity Shock A_t



Transition Dynamics

Analysis on Symmetric Countries

Main Observations.

- Productivity shock.
  - Working data (used by firms) cyclicality reverses when we change from closed model to open economy.
  - Intuition: When a productivity shock occurs in the home country, data concentrate in the high-productivity country. Less exporting, more importing....
- Disutility shock.
  - The closed model and the goods trade model behave similarly for data-related variables.
- Price shock.
  - Acts similarly with disutility shock in the open economies, but has different mechanisms (supply side vs. demand side)

Conclusion

- We contribute to the emerging literature on the data economy by extending the discussion to the international context in a dynamic general equilibrium framework.
  - Latecomer's advantage: welfare improvements from trade are large in the data-inefficient country.
  - However, trade liberalization may come to a halt when the data divide between two countries is too large.
  - Working data tends to concentrate in the data-intensive country.
  - A reversed cyclical pattern in working data following aggregate productivity shocks.
- Policy guidance on the development of data-related industries, cross-border data flows, and the mitigation of aggregate shocks.

Thanks for your attention!

For the full paper, please see NBER working paper No. 31416.