Phyloeconomic Trade Diversity: a Study of the Impact of Diversification on Firms' Export Volatility for 63 Developing Countries

Matthieu Crozet[1], Ana Margarida Fernandes[2], Thierry Mayer[3], Emmanuel Milet[4], Daria Taglioni[2]

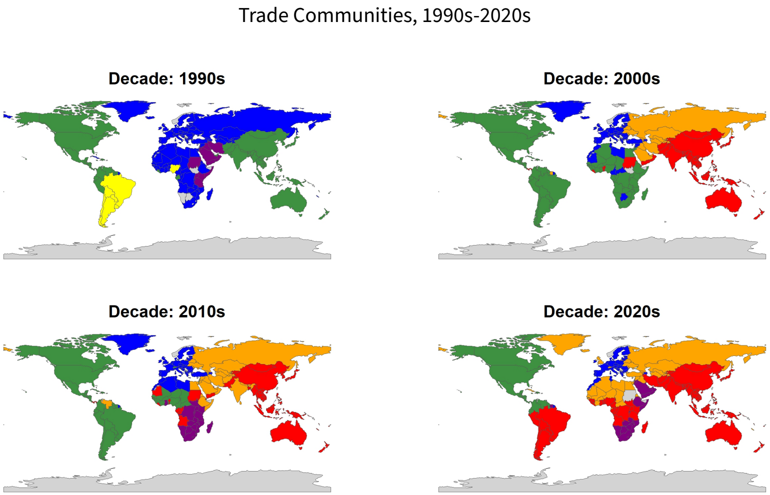
December 2025 - 4th BoI–ECB— Conference on "Trade, value chains and financial linkages in the global economy", Rome

[1]Université Paris-Saclay and CEPII
[2]World Bank
[3]Science Po, Paris
[4]University of Geneva

The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank of Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the countries they represent. All errors are our responsibility.

Context, Motivation & Puzzle

Trade volumes remain resilient, but markets are more fragmented



Source: Artuc, Goldberg, Lasso, Taglioni (2025, work in progress). **Notes:** Authors' calculation using UN COMTRADE total bilateral data. Community detection with Greedy Modularity Algorithm (networkX, Hagberg et al., 2008). Each color represents a trade community.

Diversification: does destination market similarity matter?

- This paper asks whether **diversification still stabilizes firm exports** when markets may become increasingly **dissimilar**.
- Standard intuition:
  - More markets $\Rightarrow$ better risk sharing
  - More destinations and more balanced portfolios $\Rightarrow$ less volatility.
- But is something missing?
  - The standard intuition implicitly assumes that **markets are interchangeable**.
  - Meanwhile, we also see that firms exporting to **more heterogeneous sets of destinations** can experience **higher volatility**.

Literature on why trade diversification matters for firms

+ Larger firms with more customers have less volatile output (Herskovic et al. 2020; di Giovanni et al. 2014, 2018; Kramarz et al. 2020)
+ Firms with more diversified input sources have better performance (Goldberg et al. 2010; Topalova and Khandelwal 2011; Pierola et al. 2017)
+ Exporting to foreign countries allows firms to hedge against idiosyncratic demand shocks (as in classical asset portfolio theory; Esposito 2022)

- Some firms—especially smaller exporters—experience higher export volatility when diversifying across destinations (Vannoorenberghe 2012; Vannoorenberghe et al. 2016)

Measuring trade diversification

- Most empirical studies proxy export diversification using:
  - the **number of destination markets**
  - **concentration indices** (Herfindahl–Hirschman, Theil)

- These measures do not condition on the correlation structure of shocks across markets.

- This limitation becomes more relevant in a fragmented or geopolitically polarized trade environment (non-random changes to markets' exposition to correlated vs. uncorrelated shocks).

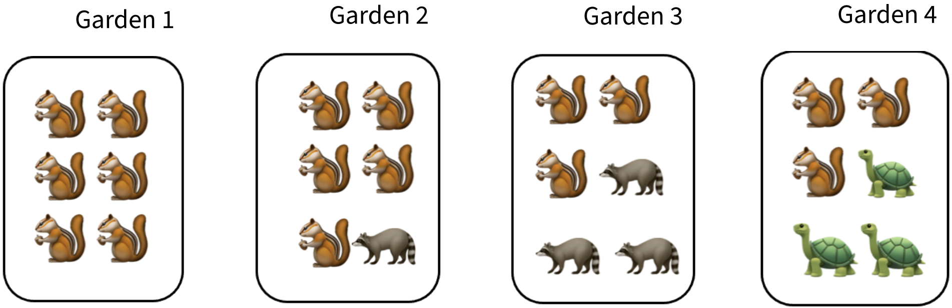
$\Rightarrow$ We need diversification measures that encompass all three margins: how rich, how balanced, and how similar/dissimilar trading partners are.

This paper

- Export diversification has **three distinct dimensions**: **richness**, **evenness**, and **disparity**, which are typically bundled together in the literature
- To separate these margins, we adapt **phylogenetic measures of biodiversity** and construct a firm-level indicator of **phyloeconomic trade diversity**
- Using customs data for over **900,000 firms in 63 emerging economies**, we ask which diversification margin drives **export volatility**
- **Key finding:** once margins are separated, they work in **opposite directions**:
  - More destinations and more balanced portfolios **reduce** volatility
  - Conditional on richness and balance levels, greater destination **dissimilarity increases volatility**

Bio-diversity and economic-diversity

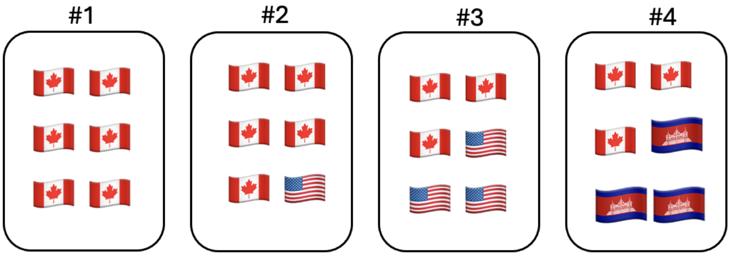
From biodiversity to export portfolios: three dimensions of diversification



- Biodiversity combines:
  - **Richness** (how many species)
  - **Evenness** (how balanced the ecosystem is)
  - **Disparity** (how different species are)

~~Phylogenetic~~ Phyloeconomic diversity across firms

- Trade analogy: export portfolios differ in **how many**, **how balanced**, and **how dissimilar** destination markets are.



- Diversity of exports of firm #4 is greatest because the economic disparity between Canada and Cambodia is larger than the disparity between Canada and the US.
- Two firms can export to the same number of markets with the same concentration, yet face different risk because destinations differ in **disparity**.

Phyloeconomic diversity:
a measurement framework to
separate diversification margins

Phyloeconomic diversity index Φ_i

- We adapt a phylogenetic diversity index (Leinster & Cobbold, 2012) to trade.
- Firm *i*'s diversification across destination markets is an average of the similarity between countries in firm *i*'s portfolio of trade partners:

$$\Phi_i = \left(\sum_{c=1}^{C} p_{ic} \, (Zp)_{ci} \right)^{-1}, \qquad (Zp)_{ci} = \sum_{c'=1}^{C} z_{cc'} \, p_{ic'} .$$

- c is a country and i is a trading firm
- p_{ic} = share of destination country c in firm i trade (exports or imports)
- $z_{cc'}$ = similarity between destination countries c and c' (inverse function of bilateral "distance" between countries)

How Φ_i captures export diversification (relative to standard measures)

- Standard measures:
  - **Number of destinations** captures **richness**
  - **Concentration / entropy measures** (HHI, Shannon, Theil) capture **evenness**

- Φ_i integrates all three dimensions of diversification:
  - **Richness**: more destination markets
  - **Evenness**: more balanced export shares
  - **Disparity**: similarity/ dissimilarity in destination markets

- Formally, $\Phi_i \geq 1$, is **increasing** in the number of partners ($\#Dest_i$) and **decreasing** in export concentration (HHI_i), **but varies across firms with same** $\#Dest_i$ **and** $\textbf{HHI}_i$.

- $\Phi_i = 1$ for a single-destination exporter.

Operationalizing similarity across markets

- We construct Φ using two alternative similarity matrices Z:
  - **Geography:** similarity falls with bilateral distance
  - **Business cycles:** similarity rises with correlation in GDP per-capita growth
- Results are robust across both definitions.

▸ A1: Index Details ▸ A2: Similarity Definition ▸ A3: Geo ▸ A4: Cycles

Data

Data: firm-level customs exports for 64 economies

- Universe of transaction-level customs data from the Exporter Dynamics Database (Fernandes, Freund, Pierola, 2016).
- Coverage: **63** emerging/developing economies; **1996–2019** (varying by country).
- Variables: exporting firm unique identifier, destination country, HS 6-digit, product, export value, year
- We aggregate those to **firm** $\times$ **HS2** $\times$ **destination** $\times$ **year**; $N \approx$ **915k** firm-product pairs.
- We compute Φ at the firm-product-year level using geography- and cycle-based Z matrices.

Understanding sources of variation in diversity indicators

	Adjusted R^2 from variance decomposition			
Matrix Z:	Geographic Proximity		Business cycle	
Sample:	All firms	Firms with $\Phi > 1$	All firms	Firms with $\Phi > 1$
Fixed effects	(1)	(2)	(3)	(4)
Firm	0.344	0.475	0.307	0.401
Year	0.005	0.015	0.001	0.002
Country	0.035	0.102	0.021	0.070
Product	0.012	0.023	0.007	0.011
Year + Country + Product	0.045	0.117	0.028	0.081
Country-Year	0.037	0.108	0.023	0.076
Product-Year	0.020	0.043	0.010	0.017
Country-Product	0.069	0.144	0.047	0.102
Country-Year + Product-Year + Country-Product	0.072	0.151	0.049	0.109

- Variation in diversity indicators is primarily driven by **firm-level heterogeneity**, rather than country, product, or time effects.
- At the same time, diversity varies **within firms over time**, which we exploit for **within-firm identification** in the regressions.

Does diversification reduce export volatility? Firm-level evidence

Econometric specification

$$\ln(Vol_{i,T}) = \alpha + \beta \ln(\Phi_{i,T_0}) + \delta \ln(HHI_{i,T_0}) + \gamma \ln(\#Dest_{i,T_0}) + \eta \ln(Dist_{i,T_0}) + FE + \varepsilon_{iT}.$$

- T = 3-year volatility windows (benchmark), from 1996-1998 to 2014-2016.
- FE: country$\times$product$\times$period (between) and + firm (within).
- $Dist_{i,T_0}$: average distance of firm i to its destination markets (so to measure dissimilarity beyond remoteness).

Measuring export volatility

- For each firm, de-trend exports within each period:

$$x_{it} = \alpha_i + \beta_i Trend_t + \varepsilon_{it}.$$

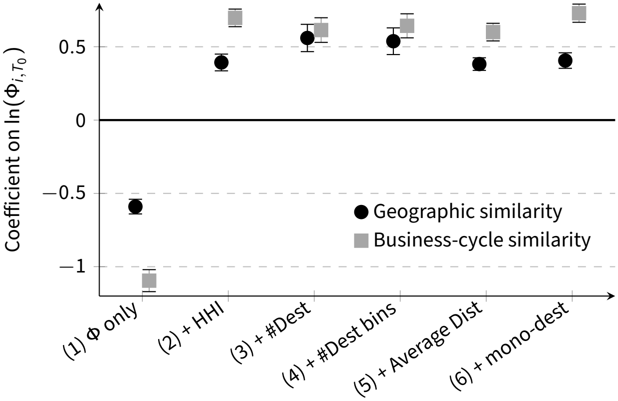
- Volatility: computed as Root Mean Square Error (RMSE) around the firm-specific trend, scaled by average firm exports ($\bar{x}_i$) over the period T:

$$Vol_{i,T} = \frac{RMSE_{i,T}}{\bar{x}_i}.$$

- Requires firms observed **continuously** within each period (requirement that eliminates one-time exporters and smaller firms).

Disentangling the diversification margins: between-firm evidence

Baseline results — across firms within country×product×period



Specification: Export volatility and diversity (between-firm)

(95% confidence intervals shown).

Core findings

Holding constant the number of destinations, export concentration (HHI), **and average distance to markets**, greater destination dissimilarity (Φ) **increases** export volatility.

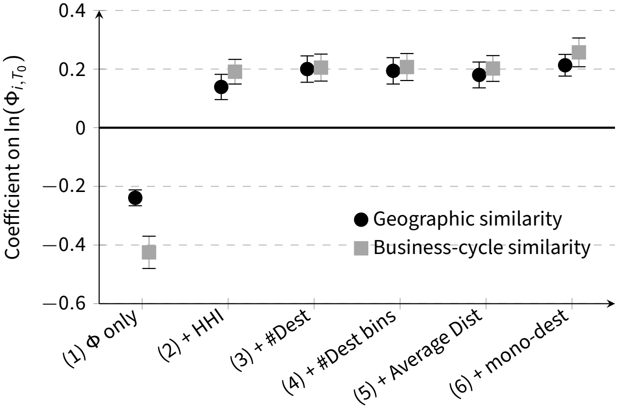
Richer and more even portfolios **reduce** volatility.

Results robust to business cycle-based similarity.

▸ Tables

Disentangling the diversification margins: within-firm evidence

Baseline results — firm fixed effects (changes within firm over time)



Specification: Export volatility and diversity (within-firm)

(95% confidence intervals shown).

Netting-out firm traits (size, capabilities, baseline riskiness)

The sign flip survives **firm fixed effects**.

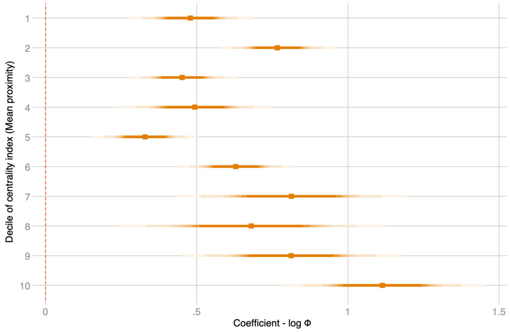
Additional insight: when a firm increases **destination dissimilarity** (higher Φ), its export volatility **rises**, even after controlling for changes in #destinations, concentration (HHI), and average distance to markets.

Interpretation: the sign flip is not driven by cross-firm selection into "exotic" portfolios; it reflects **within-firm portfolio reconfiguration**.

Results are robust to different similarity concepts.

Tables

Where diversification becomes destabilizing: the role of market access



- Geography links disparity across destinations to distance from origin (ϕ is gravity-consistent).
- **Effect strongest for firms in central countries**, where diversification mechanically implies reaching farther, more peripheral, riskier markets.
- For firms in peripheral countries, marginal increases in disparity do not change risk too much.

Why can higher disparity increase volatility?
Evidence from simulations

Why the sign flips: hedging vs. risk-taking

- Think of total exports as a portfolio:

$$Var(X_i) = \sum_c p_{ic}^2 Var(x_{ic}) + \sum_c \sum_{c' \neq c} p_{ic} p_{ic'} Cov(x_{ic}, x_{ic'}).$$

- Disparity can reduce covariance (hedging) …
- **But** reaching dissimilar markets often means reaching **more distant** markets, where firm-destination demand is harder to stabilize (higher variance).
  - For our core finding to be verified, the greater variance effect needs to dominate
  - Intuition relies on role of gravity forces and trade costs that increase with distance from the home country

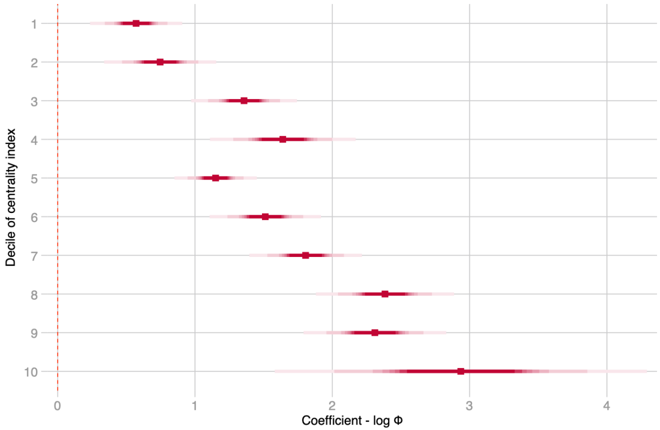
Simulations: when does disparity hedge vs. amplify risk?

- We simulate a Melitz-type model with **firm-destination demand shifters** and destination attractiveness parameters.
- Three shock environments:
  - No spatial structure (benchmark)
  - Spatially correlated macro shocks (disparity hedges)
  - **Higher volatility in distant markets** (disparity amplifies volatility)
- **Key point:** the paradox emerges when **demand volatility increases with distance** — consistent with exporter dynamics evidence.

Key point explained

- Positive coefficient on Φ in firm volatility regression results from two ingredients:
  - For a given number of destinations, firms with more diverse destinations tend to choose destinations more distant from home
  - Regardless of home, sales in destinations more distant from home are more volatile and face larger shocks (to firm-destination demand)
    - Establishing stable exports is harder in more distant markets: e.g., Chaney (2014), Morales et al. (2019)

Why destination dissimilarity can increase export volatility: evidence from simulations



Simulation: Scenario 3 (paper Figure 14c))

Mechanism

To reach more **dissimilar** destinations, firms must reach **more distant and difficult markets**.

Firm–destination demand shocks are more volatile in such markets, raising overall export volatility.

This mechanism reproduces the **positive coefficient on Φ** conditional on richness and evenness.

Consistent with literature: establishing stable exports is harder in more distant markets: e.g., Chaney (2014), Morales et al. (2019)

Implications for diversification strategies under fragmentation

- Diversification is **not one-size-fits-all**: richness and balance hedge risk, but destination dissimilarity can raise volatility.
- **Who** firms diversify towards matters as much as **how many** markets they serve.
- Diversification becomes stabilizing when accompanied by **lower trade costs and lower uncertainty** in distant markets (e.g. trade facilitation, finance, insurance).

Takeaway

- Export diversification has **three margins**: **richness**, **evenness**, and **disparity**.
- Richer and more balanced portfolios **reduce** export volatility, but greater destination dissimilarity **increases** it.
- **Bottom line:** diversification **without reducing trade frictions** can increase volatility.

Appendix: Methods and robustness

Appendix A1: Zoom in on the Φ_i indicator

$$\Phi_i = \left(\sum_{c=1}^{C} p_{ic} \left[(Zp)_{ci} \right] \right)^{-1}, \qquad (Zp)_{ci} = \sum_{c'=1}^{C} z_{cc'}\, p_{ic'}$$

$$\Phi_i^{-1} = \begin{pmatrix} p_1 & p_2 & \cdots & p_{c'} & \cdots & p_N \end{pmatrix} \begin{pmatrix} 1 & z_{1,2} & \cdots & z_{1,c'} & \cdots & z_{1,N} \\ z_{2,1} & 1 & \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ z_{c',1} & \cdots & \cdots & 1 & \cdots & z_{c'N} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ z_{N,1} & \cdots & \cdots & z_{N,c'} & \cdots & 1 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{c'} \\ \vdots \\ p_N \end{pmatrix}$$

- Z contains similarities between destination pairs; $(Zp)_{ci}$ is the weighted average similarity of c relative to the firm's portfolio.

back

Appendix A2: Measuring similarity across markets (Z) and role of home country

- $z_{cc'}$ measures similarity across destination countries c and c'.
- “Similarity” can be defined along multiple dimensions:
  - Geographic proximity (inverse of physical distance) ▸ geo
  - Business-cycle synchronicity (correlation in GDP-per-capita growth) ▸ cycle
  - (Other options: income distance, UN voting, language/culture, etc.)

- Φ_i is based on similarity **between partner markets**, not similarity to the home market.
  - Including the home market would require firm-level domestic sales.
  - Excluding it improves cross-country comparability of Φ_i.

▸ back

Appendix A3: Z matrix based on geography

- Similarity falls with bilateral distance:

$$z_{cc'}^{dist} = \exp\left(-\frac{dist_{cc'}}{\overline{dist}}\right),$$

- $\overline{dist}$ is the average distance across country pairs; $z_{cc'}^{dist} \in [0, 1]$ and $z_{cc}^{dist} = 1$.

back

Appendix A4: Z matrix based on business-cycle synchronicity

- Similarity rises with correlation in GDP-per-capita growth:

$$z_{cc'}^{bc} = \frac{1 + \mathrm{Corr}(g_{ct}, g_{c't})}{2},$$

- g_{ct} is annual growth in GDP per capita over the chosen sample window; $z_{cc'}^{bc} \in [0, 1]$.

back

Appendix B1: Baseline regression tables (between-firm)

	(1)	(2)	(3)	(4)	(5)	(6)
Z = Geographic similarity						
$\ln(\Phi_{i,T_0})$	-0.590^a	0.393^a	0.560^a	0.538^a	0.382^a	0.406^a
	(0.050)	(0.057)	(0.093)	(0.091)	(0.043)	(0.053)
$\ln(HHI_{i,T_0})$		0.454^a	0.005	0.036	0.005	-0.006
		(0.018)	(0.029)	(0.027)	(0.025)	(0.023)
$\ln(\#Dest_{i,T_0})$			-0.508^a			-0.389^a
			(0.017)			(0.012)
$\ln(Dist_{i,T_0})$					0.078^a	0.068^a
					(0.014)	(0.011)
FE hs2$\times$cty$\times$T	✓	✓	✓	✓	✓	✓
#Dest bins				✓		✓
Sample	#Dest>1	#Dest>1	#Dest>1	#Dest>1	#Dest>1	all firms
Obs.	670060	670060	670060	670060	670060	1304967
Z = Business-cycle similarity						
$\ln(\Phi_{i,T_0})$	-1.095^a	0.697^a	0.614^a	0.643^a	0.600^a	0.729^a
	(0.075)	(0.060)	(0.084)	(0.082)	(0.060)	(0.062)
$\ln(HHI_{i,T_0})$		0.465^a	-0.016	0.020	0.013	0.017
		(0.021)	(0.031)	(0.029)	(0.028)	(0.027)
$\ln(\#Dest_{i,T_0})$			-0.495^a			-0.388^a
			(0.016)			(0.011)
$\ln(Dist_{i,T_0})$					0.096^a	0.077^a
					(0.016)	(0.013)
FE hs2$\times$cty$\times$T	✓	✓	✓	✓	✓	✓
#Dest bins				✓		✓
Sample	#Dest>1	#Dest>1	#Dest>1	#Dest>1	#Dest>1	all firms
Obs.	668217	668217	668217	668217	668217	1296460

Notes: clustered SEs at country$\times$period and hs2$\times$period. Signif.: $^a p < 0.01$, $^b p < 0.05$, $^c p < 0.1$.

back

Appendix B2: Baseline regression tables (within-firm)

	(1)	(2)	(3)	(4)	(5)	(6)
Z = Geographic similarity						
$\ln(\Phi_{i,T_0})$	-0.239^a	0.139^a	0.200^a	0.194^a	0.180^a	0.213^a
	(0.027)	(0.043)	(0.045)	(0.045)	(0.044)	(0.037)
$\ln(HHI_{i,T_0})$		0.167^a	0.033^b	0.046^a	0.044^a	0.052^a
		(0.015)	(0.014)	(0.014)	(0.014)	(0.013)
$\ln(\#Dest_{i,T_0})$			-0.251^a			-0.194^a
			(0.015)			(0.012)
$\ln(Dist_{i,T_0})$					0.027^a	0.020^a
					(0.009)	(0.006)
FE hs2$\times$cty$\times$T	✓	✓	✓	✓	✓	✓
FE firm	✓	✓	✓	✓	✓	✓
#Dest bins				✓		✓
Obs.	380144	380144	380144	380144	380144	752241
Z = Business-cycle similarity						
$\ln(\Phi_{i,T_0})$	-0.425^a	0.191^a	0.205^a	0.207^a	0.202^a	0.257^a
	(0.055)	(0.042)	(0.046)	(0.046)	(0.044)	(0.049)
$\ln(HHI_{i,T_0})$		0.167^a	0.025^c	0.039^a	0.039^a	0.050^a
		(0.011)	(0.015)	(0.014)	(0.014)	(0.013)
$\ln(\#Dest_{i,T_0})$			-0.248^a			-0.193^a
			(0.014)			(0.012)
$\ln(Dist_{i,T_0})$					0.029^a	0.024^a
					(0.009)	(0.008)
FE hs2$\times$cty$\times$T	✓	✓	✓	✓	✓	✓
FE firm	✓	✓	✓	✓	✓	✓
#Dest bins				✓		✓
Obs.	379167	379167	379167	379167	379167	747168

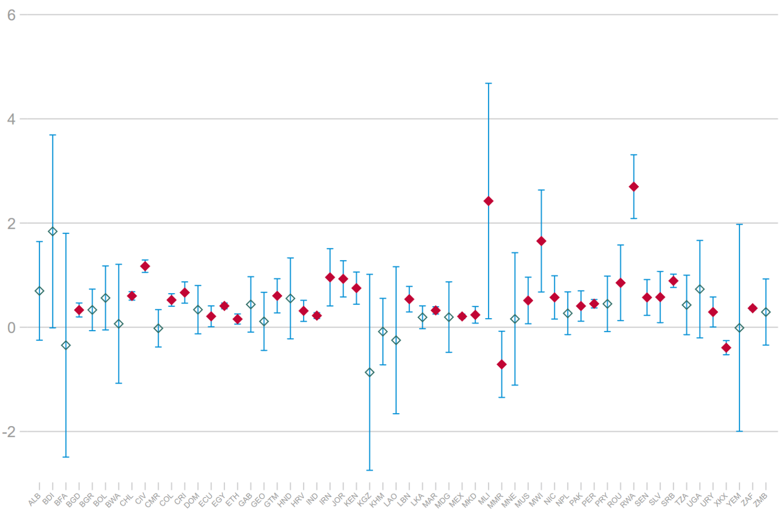
Notes: clustered SEs at country$\times$period and hs2$\times$period. Signif.: $^a p < 0.01$, $^b p < 0.05$, $^c p < 0.1$.

back

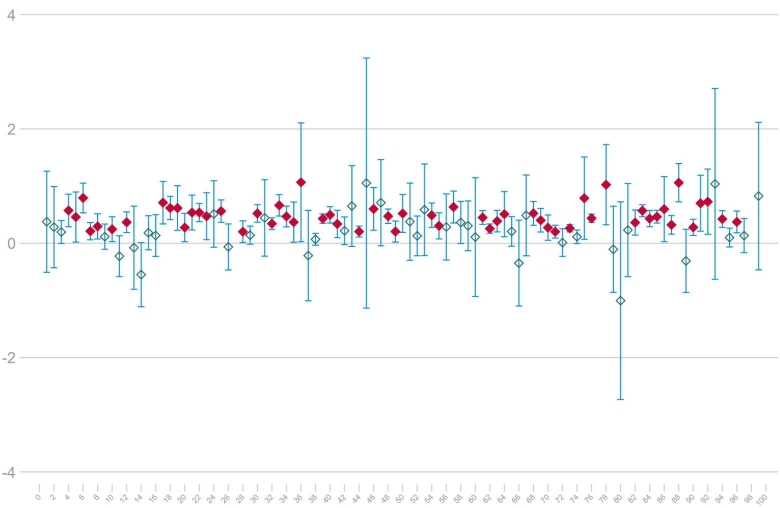
Appendix C0: Robustness and heterogeneity (navigation)

- Baseline robustness checks (tables): ▸ between-firm ▸ within-firm
- Country-by-country and product-by-product coefficients: ▸ by country ▸ by product
- Heterogeneity across subsamples: ▸ heterogeneity

Appendix C1: Between-firm estimates by country (geography)

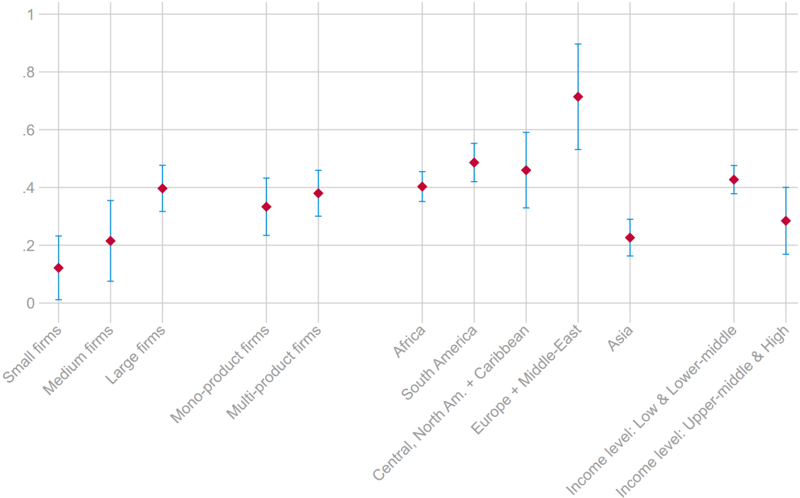


Appendix C2: Between-firm estimates by HS2 product (geography)



Appendix C3: Heterogeneity in the Φ–volatility relationship

Coefficients on ln(Φ) across subsamples



Appendix C4: Robustness checks – between-firm (Table 4)

	(1)	(2)	(3)	(4)	(5)
Dep var:	Vol	$\ln(\nu)$	$\ln(Vol)$	$\ln(Vol_{\Delta T})$	$\ln(Vol)$
T:	3y	3y	5y	3y	3y
Sample:	#Dest>1	#Dest>1	#Dest>1	#Dest>1	#Dest>4
Z = Geographic similarity					
$\ln(\Phi)$	0.058^a	0.485^a	0.382^a	0.363^a	0.370^a
	(0.010)	(0.043)	(0.040)	(0.056)	(0.035)
$\ln(HHI)$	-0.012	0.025	0.025	-0.046	0.073^a
	(0.010)	(0.024)	(0.029)	(0.029)	(0.028)
$\ln(Dist)$	0.024^a	0.075^a	0.075^a	0.087^a	0.078^a
	(0.004)	(0.013)	(0.014)	(0.014)	(0.024)
Z = Business-cycle similarity					
$\ln(\Phi)$	0.058^a	0.485^a	0.605^a	0.541^a	0.763^a
	(0.010)	(0.043)	(0.066)	(0.082)	(0.069)
$\ln(HHI)$	-0.012	0.025	0.045	-0.039	0.092^a
	(0.010)	(0.024)	(0.039)	(0.034)	(0.030)
$\ln(Dist)$	0.024^a	0.075^a	0.098^a	0.104^a	0.118^a
	(0.004)	(0.013)	(0.012)	(0.017)	(0.026)

Notes: clustered SEs at country$\times$period and hs2$\times$period. $^a p < 0.01$, $^b p < 0.05$, $^c p < 0.1$. ▸ back

Appendix C5: Robustness checks – within-firm (Table 5)

	(1)	(2)	(3)	(4)	(5)
Dep var:	Vol	$\ln(\nu)$	$\ln(Vol)$	$\ln(Vol)$	$\ln(Vol)$
T:	3y	3y	5y	3y	3y
Sample:	#Dest>1	#Dest>1	#Dest>1	#Dest>1	#Dest>4
Z = Geographic similarity					
$\ln(\Phi)$	0.005	0.160^{a}	0.129^{a}	0.260^{a}	0.104
	(0.008)	(0.040)	(0.034)	(0.061)	(0.064)
$\ln(HHI)$	-0.003	0.039^{b}	0.047^{a}	-0.010	0.071^{a}
	(0.005)	(0.016)	(0.012)	(0.022)	(0.020)
$\ln(Dist)$	0.006^{a}	0.024^{a}	0.001	0.038^{a}	0.025
	(0.002)	(0.007)	(0.009)	(0.013)	(0.017)
Z = Business-cycle similarity					
$\ln(\Phi)$	0.010^{b}	0.201^{a}	0.136^{a}	0.293^{a}	0.092
	(0.005)	(0.025)	(0.044)	(0.087)	(0.094)
$\ln(HHI)$	-0.001	0.053^{a}	0.045^{a}	-0.014	0.063^{a}
	(0.004)	(0.014)	(0.015)	(0.028)	(0.022)
$\ln(Dist)$	0.007^{a}	0.015^{c}	0.010	0.045^{a}	0.018
	(0.002)	(0.008)	(0.007)	(0.015)	(0.020)

Notes: clustered SEs at country × period and hs2 × period. $^{a}p < 0.01$, $^{b}p < 0.05$, $^{c}p < 0.1$. ▸ back

Appendix C6: Additional robustness (Tables 8–9)

- Additional checks (paper Appendix): winsorizing volatility, controlling for initial exports, restricting to large country$\times$product cells, dropping largest countries.

	(1) Wins.	(2) +ln X_0	(3) $>$500 firms	(4) Drop big C
Between-firm, Z = Geographic				
ln(Φ)	0.272^a	0.246^a	0.343^a	0.475^a
	(0.030)	(0.052)	(0.041)	(0.032)
Within-firm, Z = Geographic				
ln(Φ)	0.130^a	0.119^a	0.190^b	0.149^a
	(0.030)	(0.046)	(0.087)	(0.041)

(Full tables for both Z matrices in paper Appendix Tables 8–9.) back