The Twin Peaks of the Export Intensity Distribution

Teaching Materials

Fabrice Defever[1], Alejandro Riaño[2]

[1] University of Lille, CEP (LSE), and CESifo
[2] City, University of London, GEP, and CESifo

November 22, 2021

Exporters...

1. are a minority within a given country/industry ✓

Exporters...

1. are a minority within a given country/industry ✓

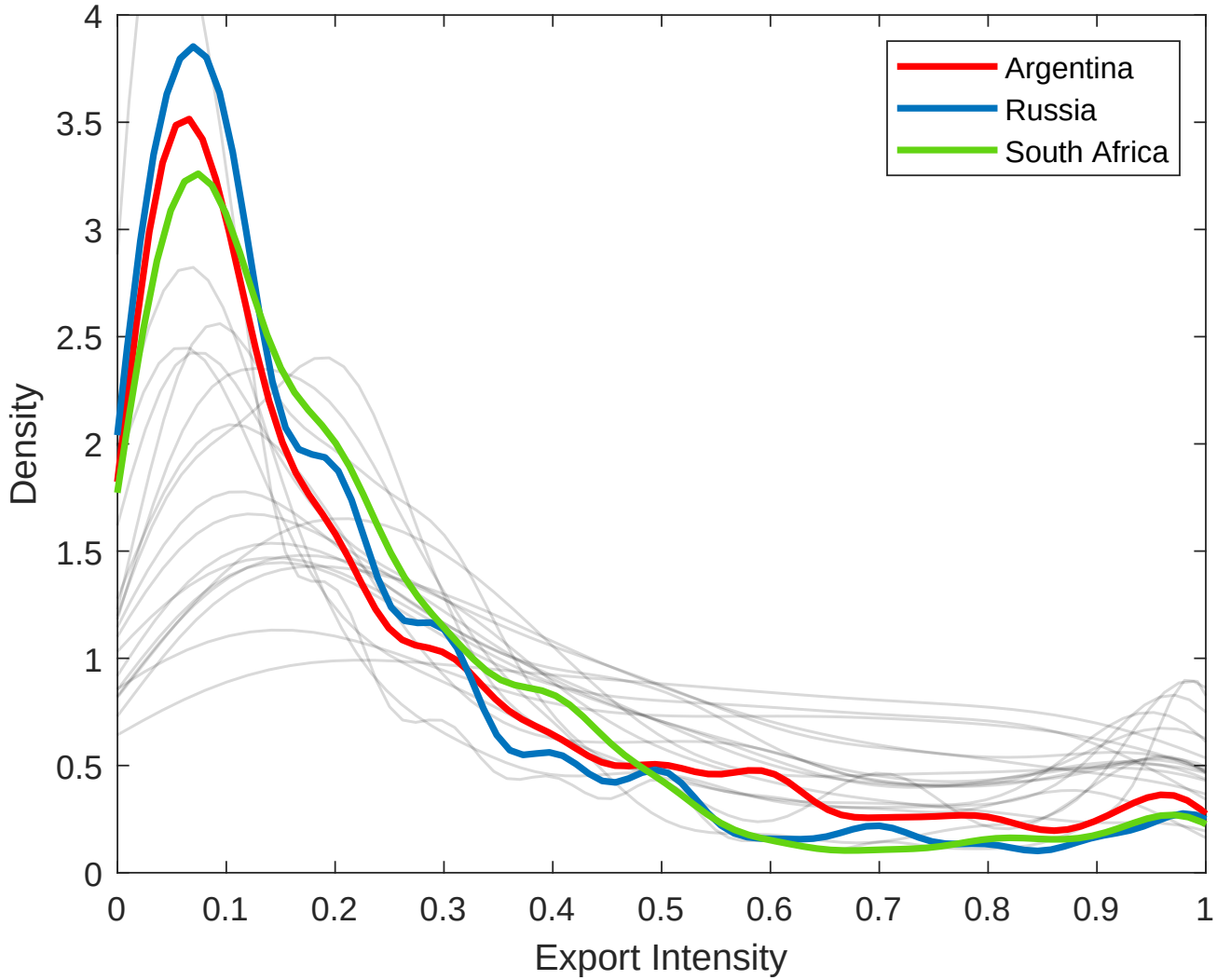
2. are 'better' than non-exporters (in terms of size, productivity, skill intensity,...) ✓

Exporters...

1. are a minority within a given country/industry ✓
2. are 'better' than non-exporters (in terms of size, productivity, skill intensity,...) ✓
3. sell most of their output domestically (?)

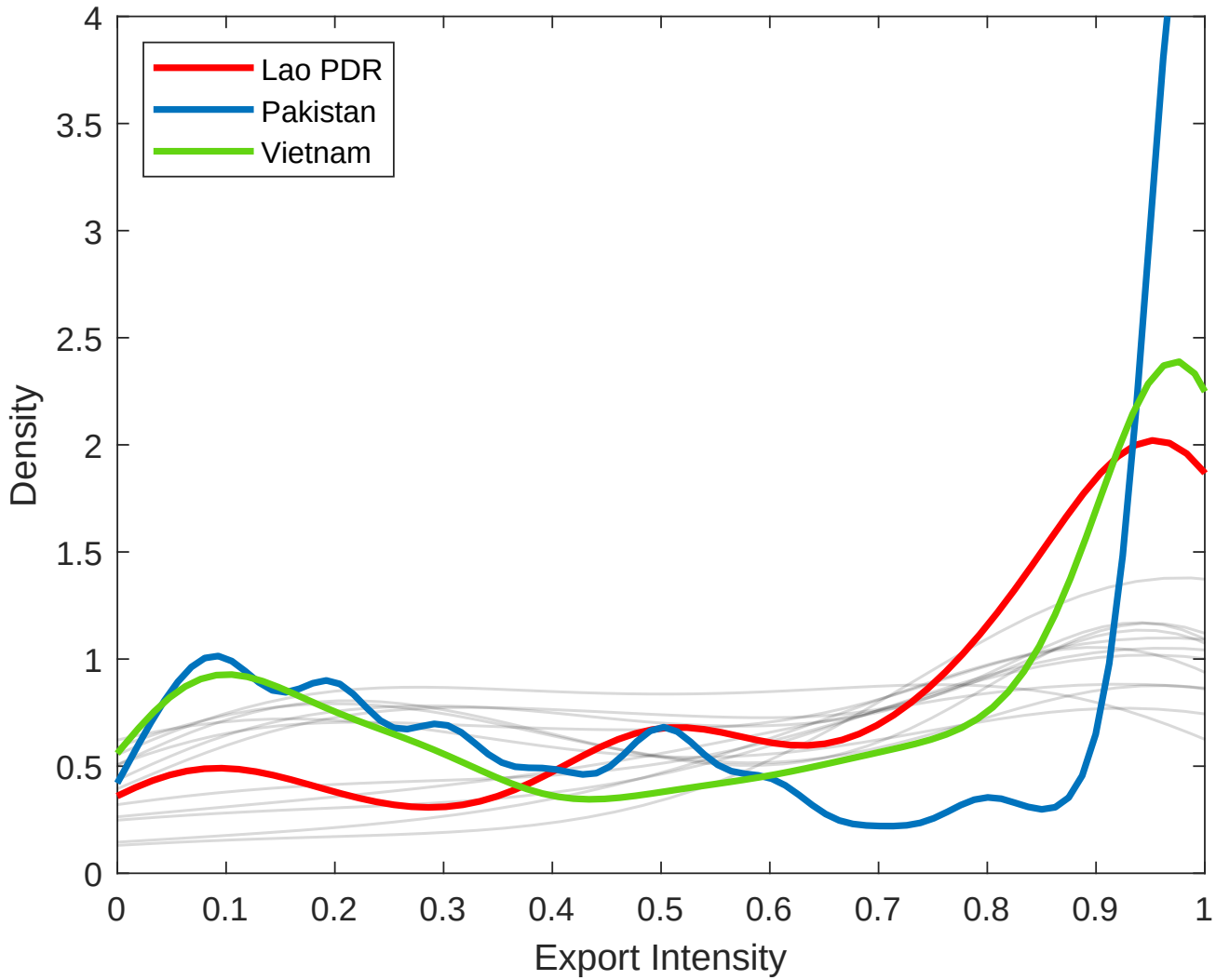
Received wisdom

Most exporters in a country sell the majority of their output domestically and only a small minority of them concentrate their sales abroad



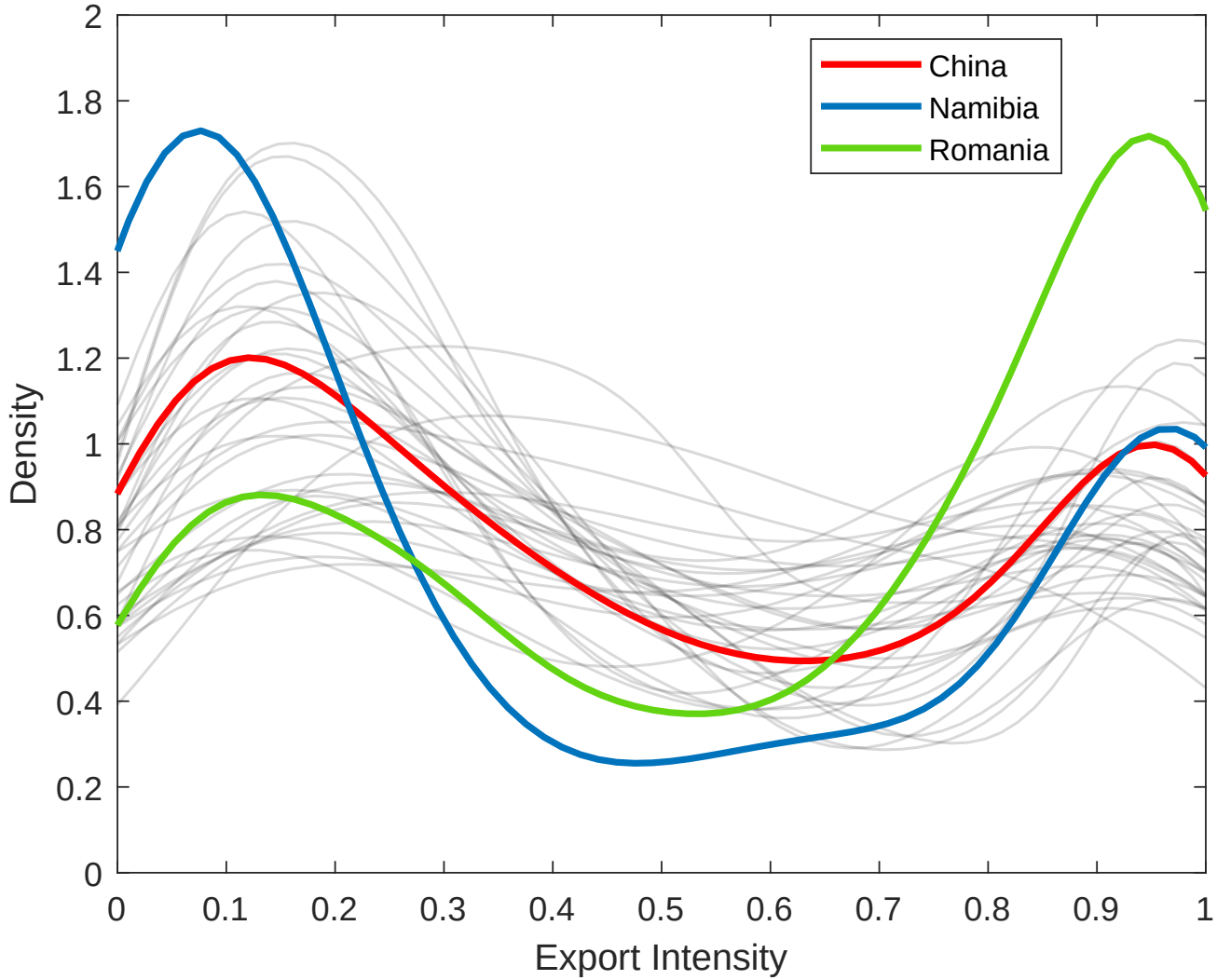
Sometimes we observe the opposite

Most exporters are high-intensity ones



Twin peaks

In 2/3 of the countries in our data we see a high concentration of firms on both ends of the export intensity distribution



In this paper we show that...

1. Export intensity distributions vary a lot across the world
   - Bimodal distributions are the rule rather than the exception
   - The mode(s) are located near 0, 1, or in both ends of the distribution

2. CES model of trade + firm-destination-specific revenue shifters $\rightarrow$ generate observed variation in export intensity distributions across countries when the firm-destination component of sales has sufficiently high dispersion
   - Derive a closed-form pdf for export intensity that generate bimodality when firm-destination-specific revenue shifters are distributed lognormal
   - Estimate the model's structural parameters using readily-available data

Main takeaways

1. Differences in countries' size relative to the rest of the world account for most of the observed variation in the distribution of export intensity across the countries in our data

2. In our model relatively small and large countries have export intensity distributions that **'look'** unimodal—in the sense that a statistical test does not reject unimodality, while countries of intermediate size display prominent twin peaks

3. While policies that incentivize firms to export a high share of their output (e.g. export processing regimes, subsidies subject to export share requirements, special economic zones) account for a substantial share of the variation in the dispersion of firm-destination revenue shifters, they cannot fully account for the widespread prevalence of twin peaks around the world

Why is the Export Intensity Distribution Important?

- The **mean** export intensity has been used extensively to pin down the magnitude of costs that impede international flows of goods → the variable cost associated with exporting is the same for all firms selling in a given destination
- Several policies that distort the relative incentives to export vis-à-vis selling domestically are **firm-specific**: allocation of export quota rights (Khandelwal et. al. 2013); subsidies targeted to firms on the basis of their location, foreign ownership, subject to export share requirements (Farole and Akinci 2011; Defever & Riaño 2017; Defever et al. 2019, 2020); subsidies to products of "strategic importance" (Westphal 1990; Kalouptisidi 2018)
- To quantify the impact of these policies on aggregate outcomes such as productivity and welfare, it is necessary to infer what the distribution of export intensity would have been in the absence of distortions, analogously to the misallocation literature (Restuccia & Rogerson 2008; Guner et al. 2008; Hsieh & Klenow 2009; Brooks & Wang 2017; Defever & Riaño 2017)
- Recent work has shown that firms' response to external shocks—most notably to real exchange rate (RER) depreciations—is substantially heterogeneous across the distribution of export intensity (Alfaro et al. 2017; Kohn et al. 2020)
- The shape of the export intensity distribution also affects the sales diversification benefits that firms achieve from exporting (Riaño 2011; Vannoorenberghe 2012)

Data and Stylized Fact

Data

- 72 countries drawn from World Bank's Enterprise Surveys (WBES) with at least 97 manufacturing exporters (ISIC Rev. 3.1 sectors 15-37), 2002-2016
- Data are representative of the private, non-agricultural economy
- Unit of observation is the establishment—a physical location where business is carried out or industrial operations take place which should have its own management and control over its own workforce
- Vast majority of establishments surveyed report to be single-establishment firms, hereafter we refer to them as 'firms'
- WBES provides information on firms main sector of operation, age, total sales, export intensity, foreign-ownership status, labor productivity, and the share of material inputs accounted for by imports.
- Some survey waves provide information on the first year a firm began exporting, the number of products it produces (at the 4-digit ISIC industry level) and bilateral export sales to specific destinations
- Export intensity is defined as the share of sales that a firm exported directly or indirectly through an intermediary in a fiscal year, and therefore takes values in the interval $(0, 1]$

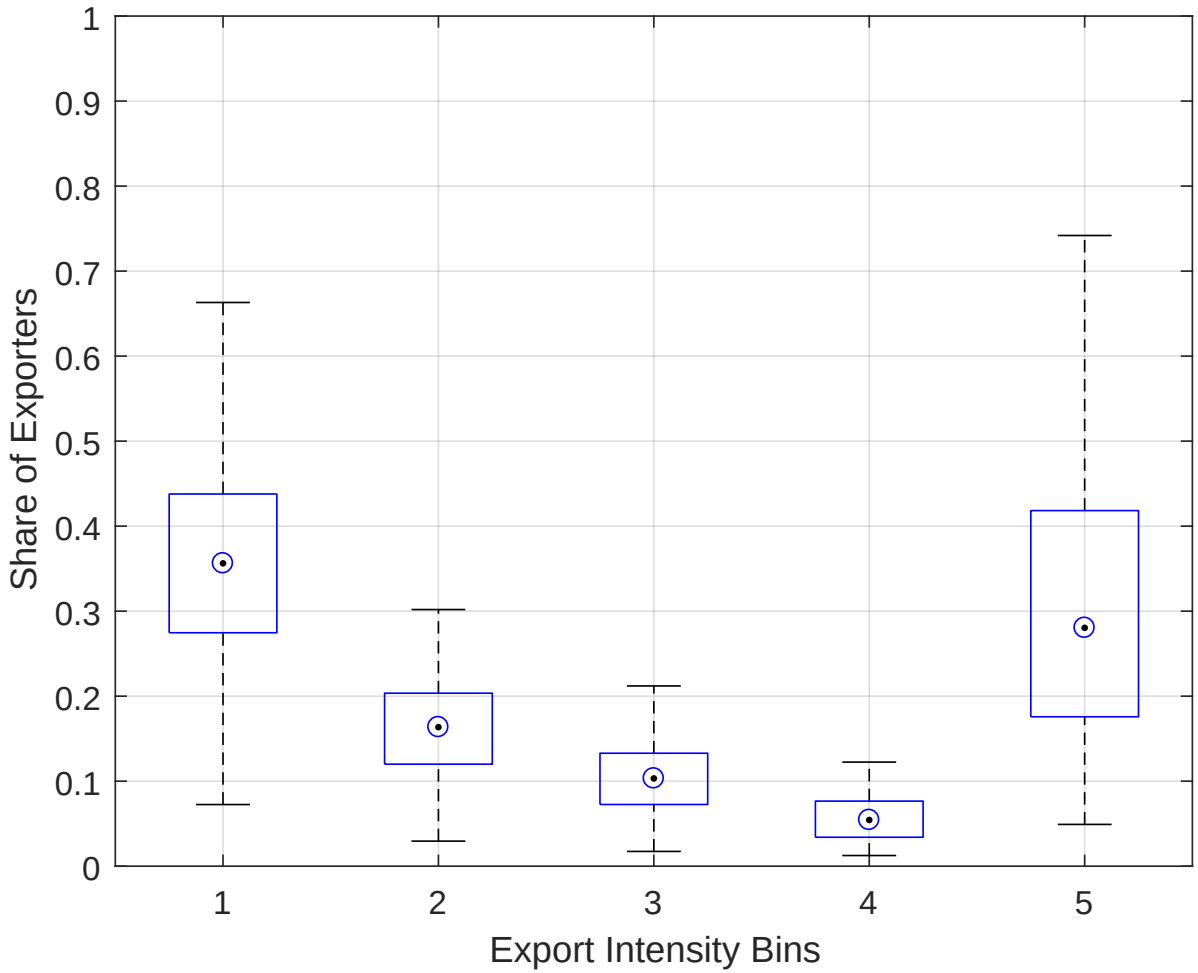
Summary statistics

- In the WBES data:
  - exporters are on average larger and more productive than non-exporters
  - there is no strong correlation between productivity and export intensity
  - high-intensity exporters are more likely to be foreign-owned and import a higher share of their inputs

	Employment	Output	Output per worker	% Foreign-owned	% Imported inputs
	(1)	(2)	(3)	(4)	(5)
Non-Exporters	0.5	0.5	0.8	6.0	23.7
Exporters					
Export intensity:					
$\in (0.0, 0.2]$	1.7	2.1	1.4	17.4	37.7
$\in (0.2, 0.4]$	1.5	1.8	1.3	18.8	35.6
$\in (0.4, 0.6]$	1.7	1.9	1.3	21.0	35.2
$\in (0.6, 0.8]$	2.3	2.3	1.4	23.0	35.4
$\in (0.8, 1.0]$	2.2	2.3	1.6	31.1	41.0

Premia relative to the average firm in each country-year survey wave

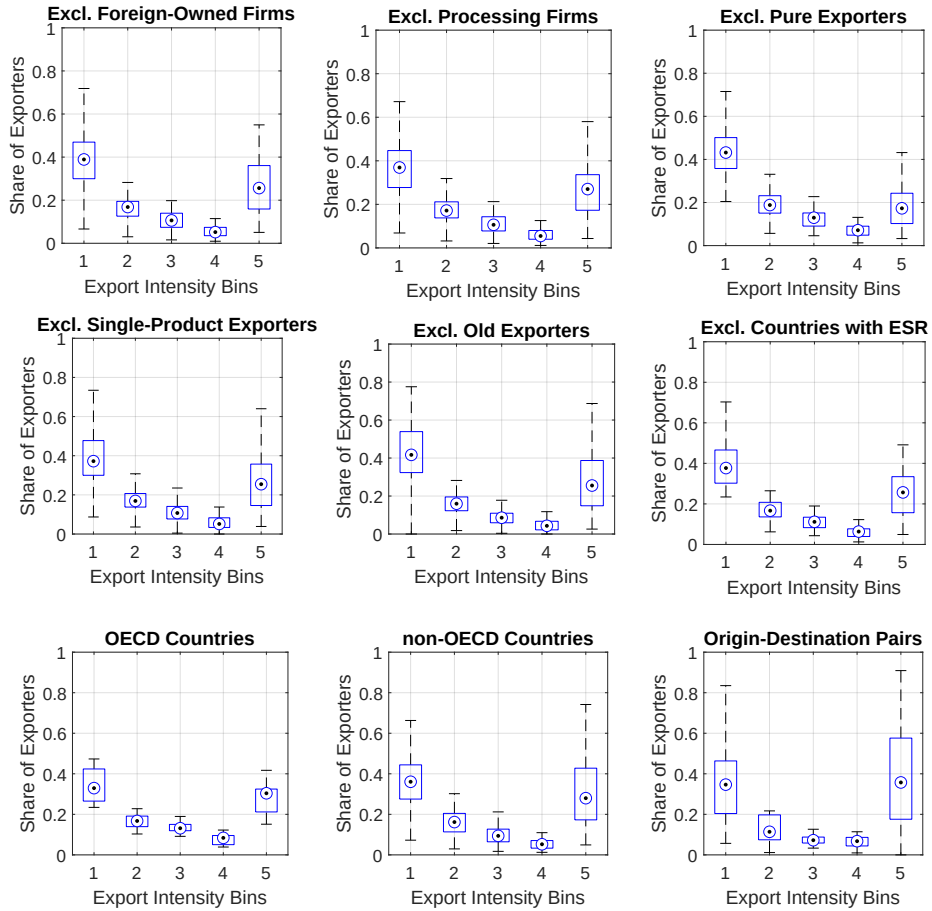
Raw data



- We calculate the share of exporters across 5 export intensity bins ($(0, 0.2], \ldots, (0.8, 1]$) in each country and present the distribution of these shares in each bin
- In most countries exporters concentrate in the first and last export intensity bins—i.e. they either sell most of their output domestically or abroad

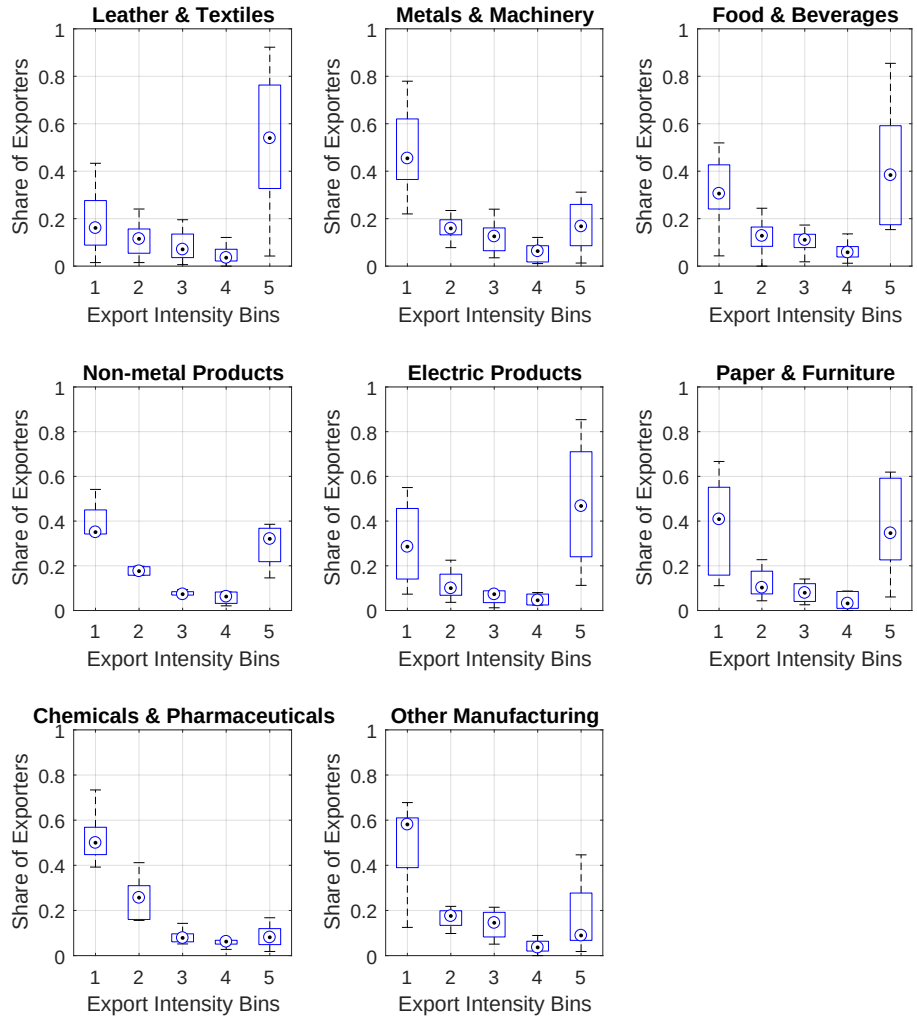
Are twin peaks a product of composition effect?

- is the distribution at the country level is mixture of 'standard' low-intensity exporters and a significant number of firms that export most of their output?



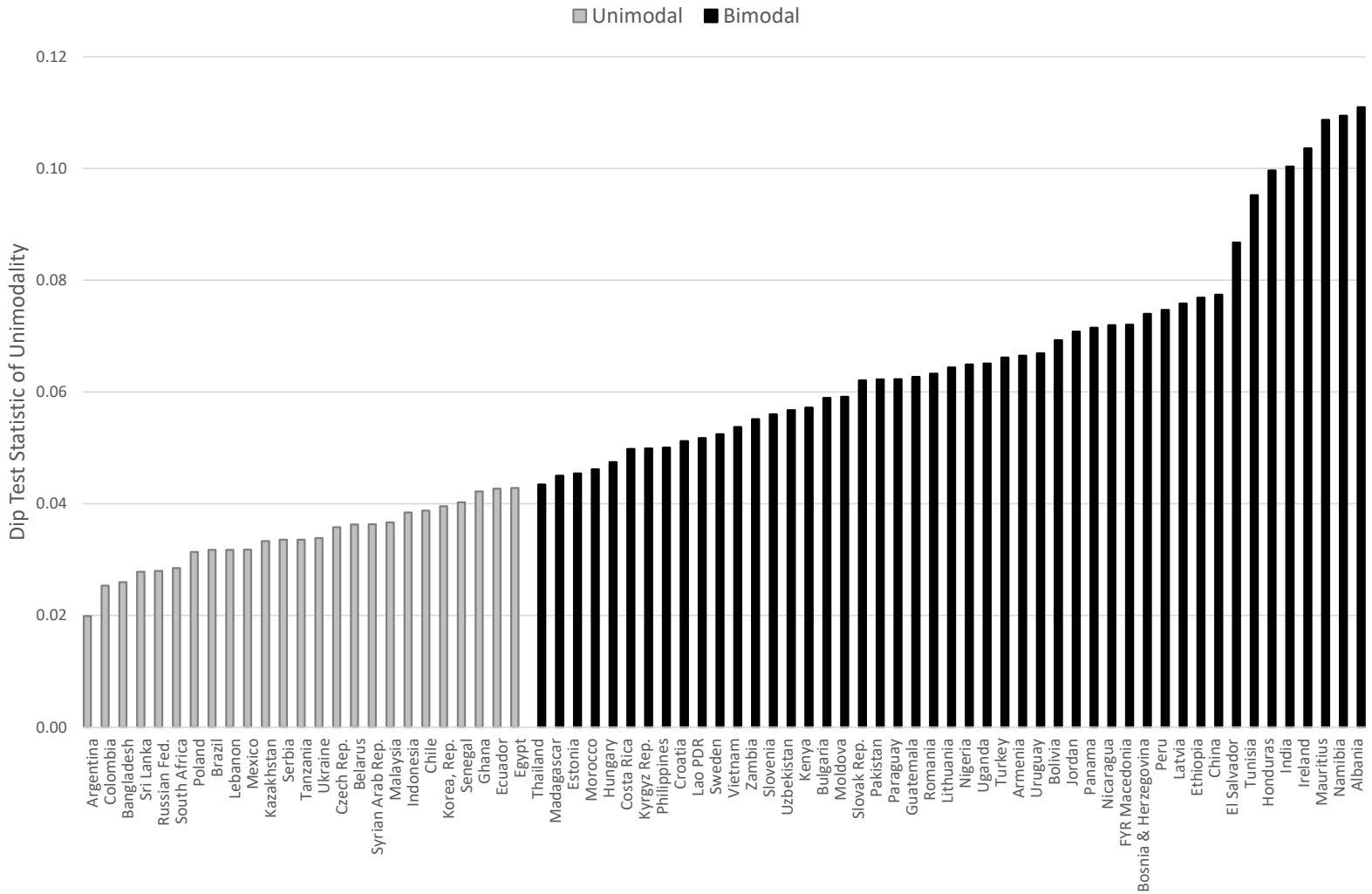
- **it doesn't seem to be the case**

Could twin peaks be due to sectoral differences?



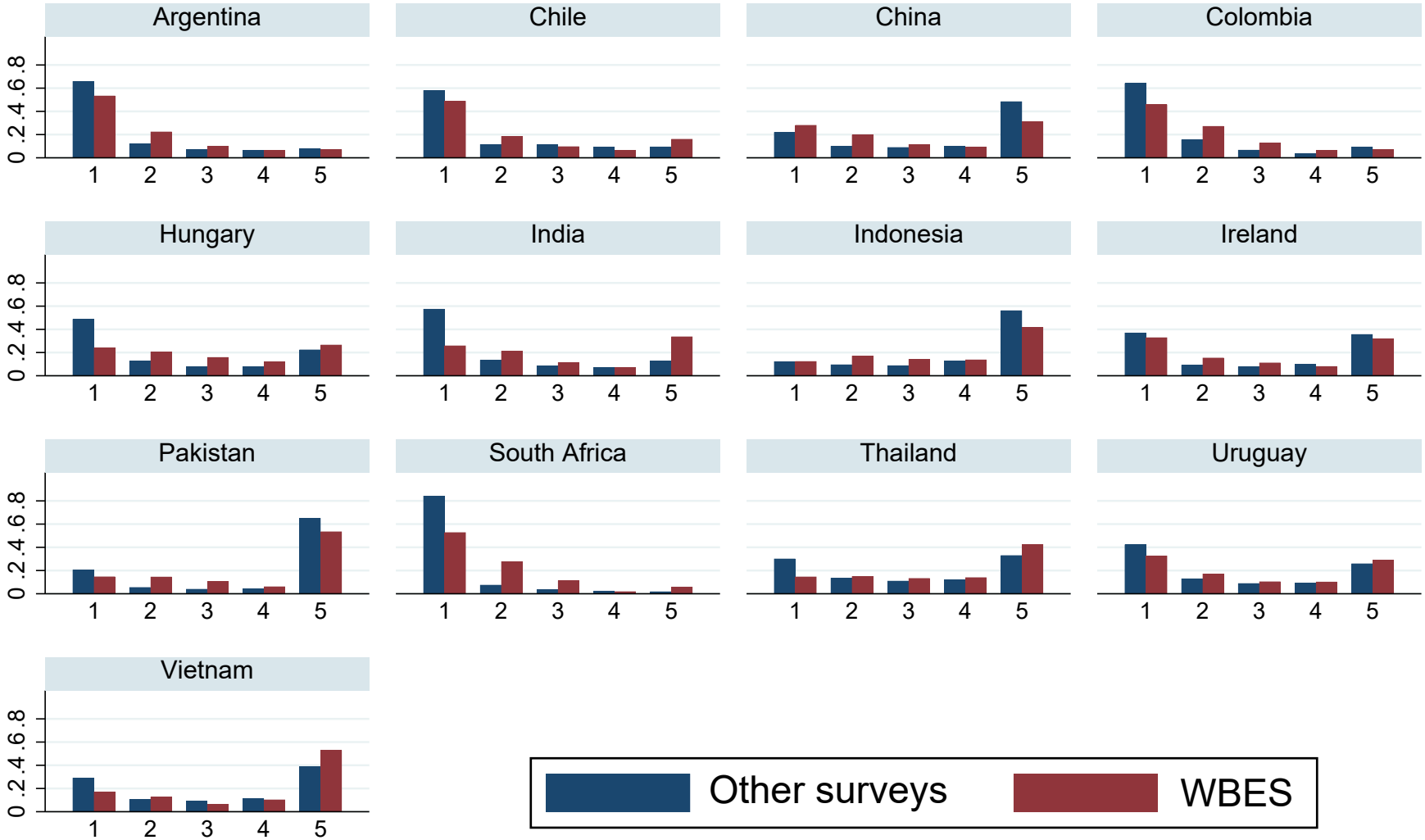
- doesn't seem to be the case either
- **most exporters operate at either a very low or very high export intensity regardless of how we slice the data**

Bimodality is a salient feature of export intensity distributions



- **For 47 out of 72 countries we reject the null hypothesis of unimodality using the Hartigan & Hartigan (1985) dip test**

Are twin peaks an artifact of the WBES data?



- **No. if anything the share of exporters in the lowest and highest export intensity bins tends to be higher in more representative surveys than in WBES**

Model

Theoretical framework

- Two countries $i \in \{d, x\}$ where a representative consumer has CES preferences
- Monopolistic competition with large number of heterogeneous firms as in Melitz (2003)
- Sales of firm ω in market i can be written as:

$$r_i(\omega) = s_i \cdot \Phi(\omega) \cdot z_i(\omega)$$

- s_i is market i's size, which is common across all firms selling there (e.g. market i's total income and price level, home's wage)
- $\Phi(\omega)$ varies between firms but is the same across destinations (e.g. total factor productivity or the quality of a firm's product)
- $z_i(\omega)$ includes factors that are firm-destination specific (e.g. cross-country differences in tastes, policies that incentivize firms to sell more in a given market such as export processing regimes or SEZ targeted at exporters, the extent of a firm's network of customers, or its participation in global value chains,...)
- We assume that $\{z_i(\omega)\}_{i \in \{d,x\}}$ are distributed lognormal ($\mathcal{LN}$) independently across destinations, with underlying mean 0 and variance σ_{zi}^2
- $\rightarrow Z_i(\omega) \equiv s_i z_i(\omega) \sim \mathcal{LN}\left(\ln(s_i), \sigma_{zi}^2\right)$

Export intensity

- Export intensity—the share of sales accounted for by exports—is given by:

$$E \equiv \frac{r_x}{r_d + r_x} = \frac{s_x z_x(\omega)}{s_d z_d(\omega) + s_x z_x(\omega)} = \frac{Z_x(\omega)}{Z_d(\omega) + Z_x(\omega)}$$

- Without $z_i(\omega)$ and 2 countries $\rightarrow$ E is degenerate at the level $\frac{s_x}{s_d + s_x}$
- With more than 2 countries $\rightarrow$ the export intensity distribution inherits the properties of the productivity distribution
- Let $h^{\mathcal{LN}}(e)$ denote the probability density function of export intensity

Closed-form expression for $h^{\mathcal{LN}}(e)$

Proposition

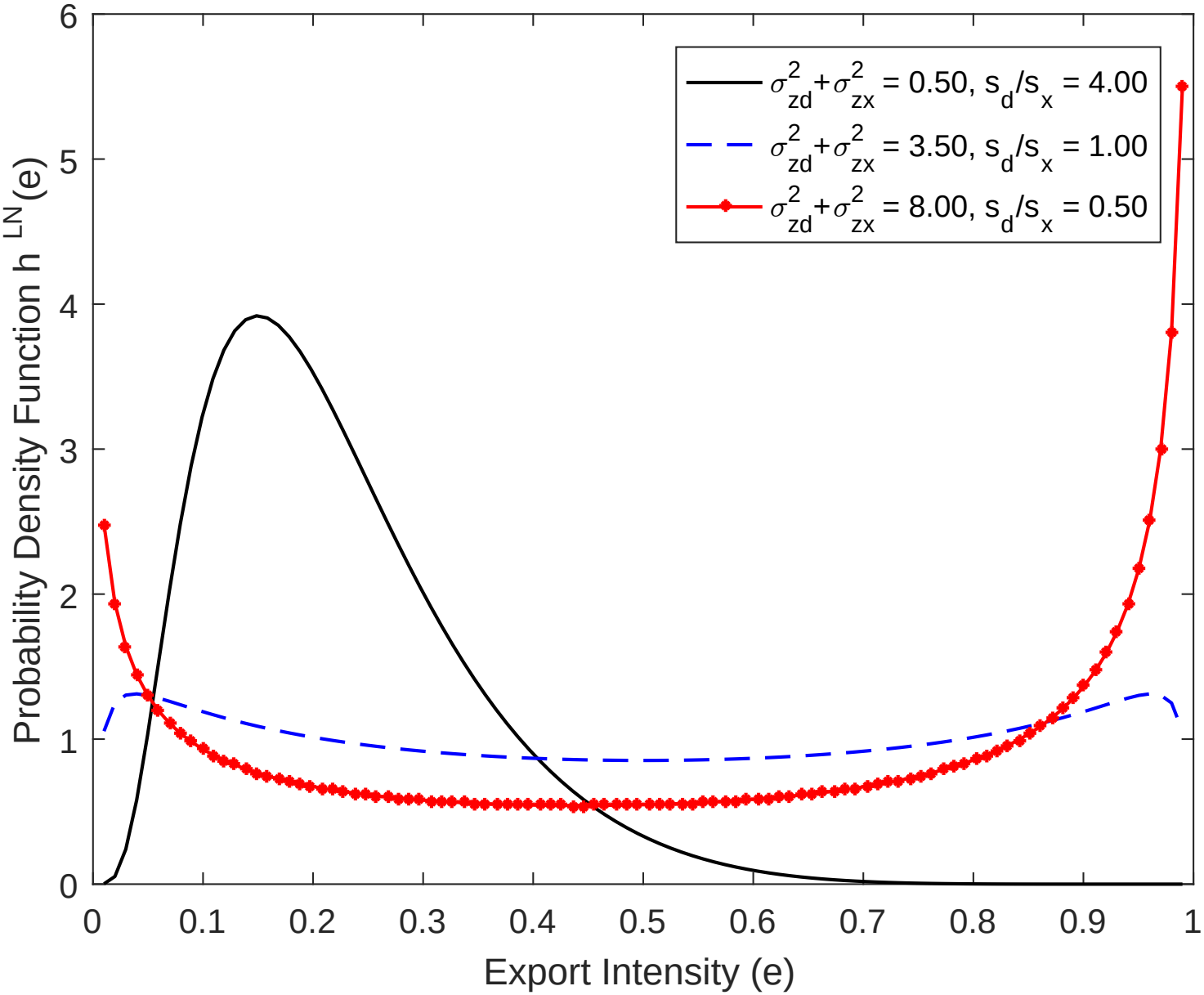
Assume that firm-destination-specific revenue shifters $\{z_i(\omega)\}_{i\in\{d,x\}}$ *are distributed lognormal* ($\mathcal{LN}$) *that are independent across destinations, with underlying mean 0 and variance* σ_{zi}^2*, so that* $z_i(\omega) \sim \mathcal{LN}\left(0, \sigma_{zi}^2\right)$*, and therefore,* $Z_i(\omega) \equiv s_i z_i(\omega) \sim \mathcal{LN}\left(\ln(s_i), \sigma_{zi}^2\right)$*. Then the probability density function of export intensity is given by:*

$$h^{\mathcal{LN}}(e) = \frac{1}{[e(1-e)]\sqrt{2\pi(\sigma_{zd}^2+\sigma_{zx}^2)}} \times \exp\left[-\frac{\left(\ln\left(\frac{e}{1-e}\right)+\ln\left(\frac{s_d}{s_x}\right)\right)^2}{2(\sigma_{zd}^2+\sigma_{zx}^2)}\right], \quad e \in (0,1).$$

$h^{\mathcal{LN}}(e)$ is characterized by two parameters:

- s_d/s_x: the relative size of the domestic market compared to the foreign one, aka the scale parameter
- $\sigma_{zd}^2+\sigma_{zx}^2$: the sum of the variances of domestic and export revenue shifters, aka the shape parameter

Examples of $h^{\mathcal{LN}}(e)$ for different parameter values



Conditions for bimodality

- $h^{\mathcal{LN}}(e)$ follows what is called a logit-normal distribution

Proposition

The distribution of export intensity is bimodal if revenue shifters follow a lognormal distribution as specified in Proposition 1 and the following two conditions are satisfied:

$$\sigma_{zd}^2 + \sigma_{zx}^2 > 2,$$

and

$$|\ln(s_d/s_x)| < \left(\sigma_{zd}^2 + \sigma_{zx}^2\right)\sqrt{1 - \frac{2}{\sigma_{zd}^2 + \sigma_{zx}^2}} - 2\tanh^{-1}\left(\sqrt{1 - \frac{2}{\sigma_{zd}^2 + \sigma_{zx}^2}}\right).$$

- The two modes lie in the interior of the support but do not have a closed-form solution
- The major mode is located near 0 when $s_d/s_x > 1$, and near 1 in the converse case; if $s_d/s_x = 1$, then the distribution is symmetric around 0.5.

These properties are proved by Johnson (1949)

Intuition

- When $\sigma_{zd}^2 + \sigma_{zx}^2$ is sufficiently high, the distribution of export intensity exhibits twin peaks
- Because the revenue shifters are independent across destinations, the likelihood that firms face a very high demand in only one of the two markets they serve—thereby generating export intensities close to either 0 or 1—is higher when $\sigma_{zd}^2 + \sigma_{zx}^2$ is high
- Increasing $\sigma_{zd}^2 + \sigma_{zx}^2$ makes the twin peaks more prominent by shifting probability mass towards the boundaries of the support
- The second condition for bimodality defines a U-shaped curve in the space $\left\{\left(\sigma_{zd}^2 + \sigma_{zx}^2\right), e^{med}\right\}$, which determines the level of the variance of revenue shifters necessary to produce bimodality given the relative market size
- For countries that are either very small or very large vis-à-vis the foreign market, and therefore have substantial probability mass near 0 or 1 respectively, the necessary cutoff for the shape parameter to produce a bimodal distribution is higher than for countries for which s_d/s_x is closer to 1

Alternative assumptions

- Bimodality obtains with other distributions too. For firm-destination-specific revenue shifters that follow gamma and Fréchet, we can also obtain closed-form expressions for the pdf of export intensity; beta, chi-squared and F distributions also generate bimodal (but not closed-form) export intensity distributions

- Higher $\sigma_{zd}^2 + \sigma_{zx}^2$ also generates bimodality when firms select into exporting based on the realization of their export revenue shifter, $z_x(\omega)$. The truncation generated by the fixed cost of exporting precludes us from obtaining a closed form expression for the pdf of export intensity

- The result that higher dispersion of firm-destination-specific revenue shifters results in a bimodal export intensity distribution also obtains when revenue shifters are correlated across markets and when firms can export to more that one destination

Recovering relative market size directly from the data

Proposition

If revenue shifters follow a lognormal distribution as specified in Proposition 1, then the median export intensity, e^{med}, is given by:

$$e^{med} = \frac{s_x}{s_d + s_x},$$

which is independent of the shape parameter $\sigma_{zd}^2 + \sigma_{zx}^2$.

- There is a one-to-one relationship between relative market size and the median export intensity
- This allows us to recover relative market size for each country in a model-consistent way without having to calibrate other parameters of the model

Estimation and Results

Identification

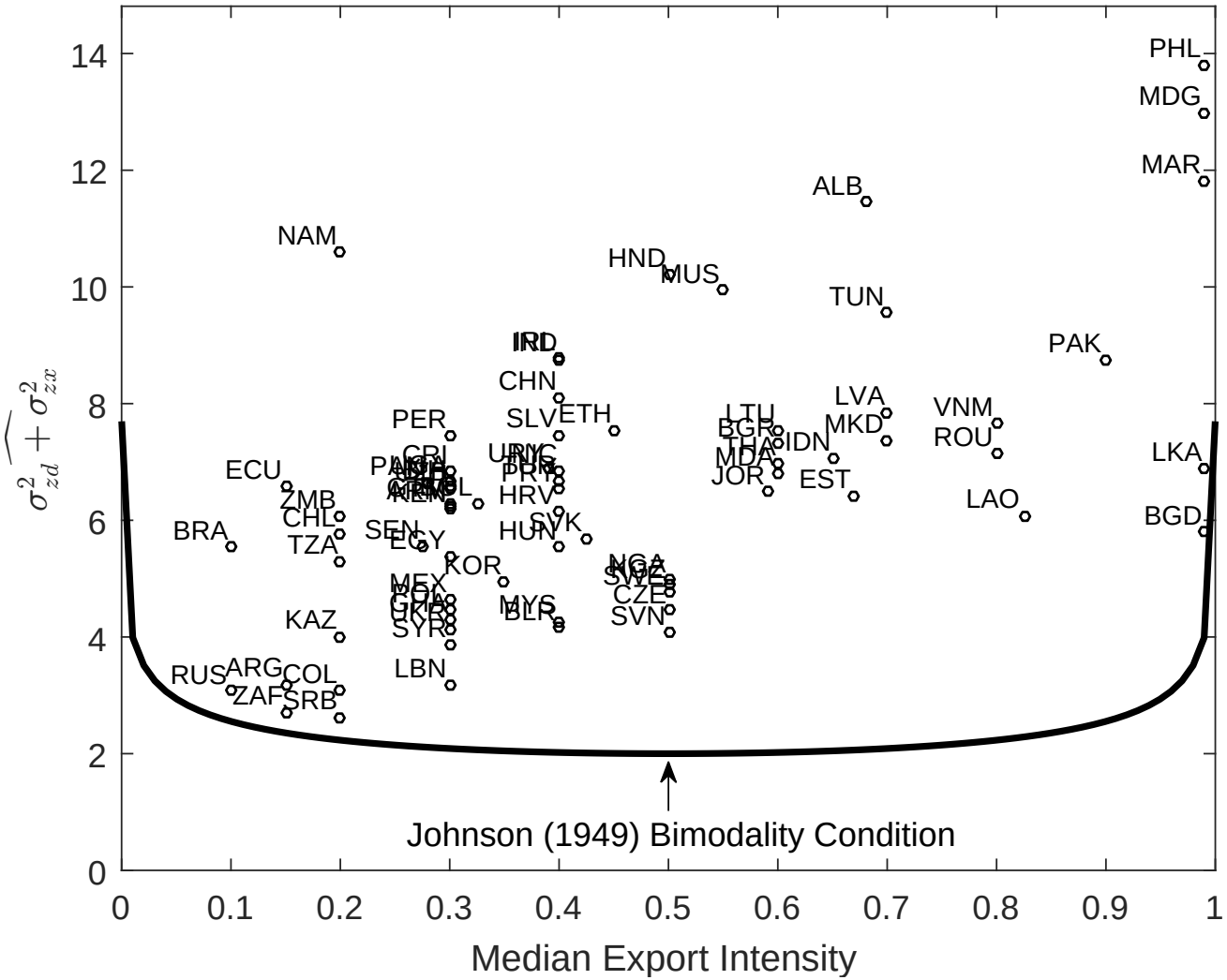
- We recover s_d/s_x directly from using each country's median export intensity, e^{med}:

$$\left(\frac{s_d}{s_x}\right) = \frac{1 - e^{med}}{e^{med}}$$

- Conditional on s_d/s_x, we estimate $\sigma_d^2 + \sigma_x^2$ by maximum likelihood
- The value of $\sigma_d^2 + \sigma_x^2$ is identified by whether the mass of the export intensity distribution is concentrated in the interior of the support or near its boundaries
- If the dispersion of revenue shifters is low $\rightarrow$ the distribution of export intensity is unimodal with most exporters exhibiting an intensity close to $s_x/(s_d + s_x)$
- Alternatively, if $\sigma_d^2 + \sigma_x^2$ is high there would be large clusters of exporters with intensities near 0 and 1

Benchmark estimates

- We estimate $\sigma_d^2 + \sigma_x^2$ independently for each country

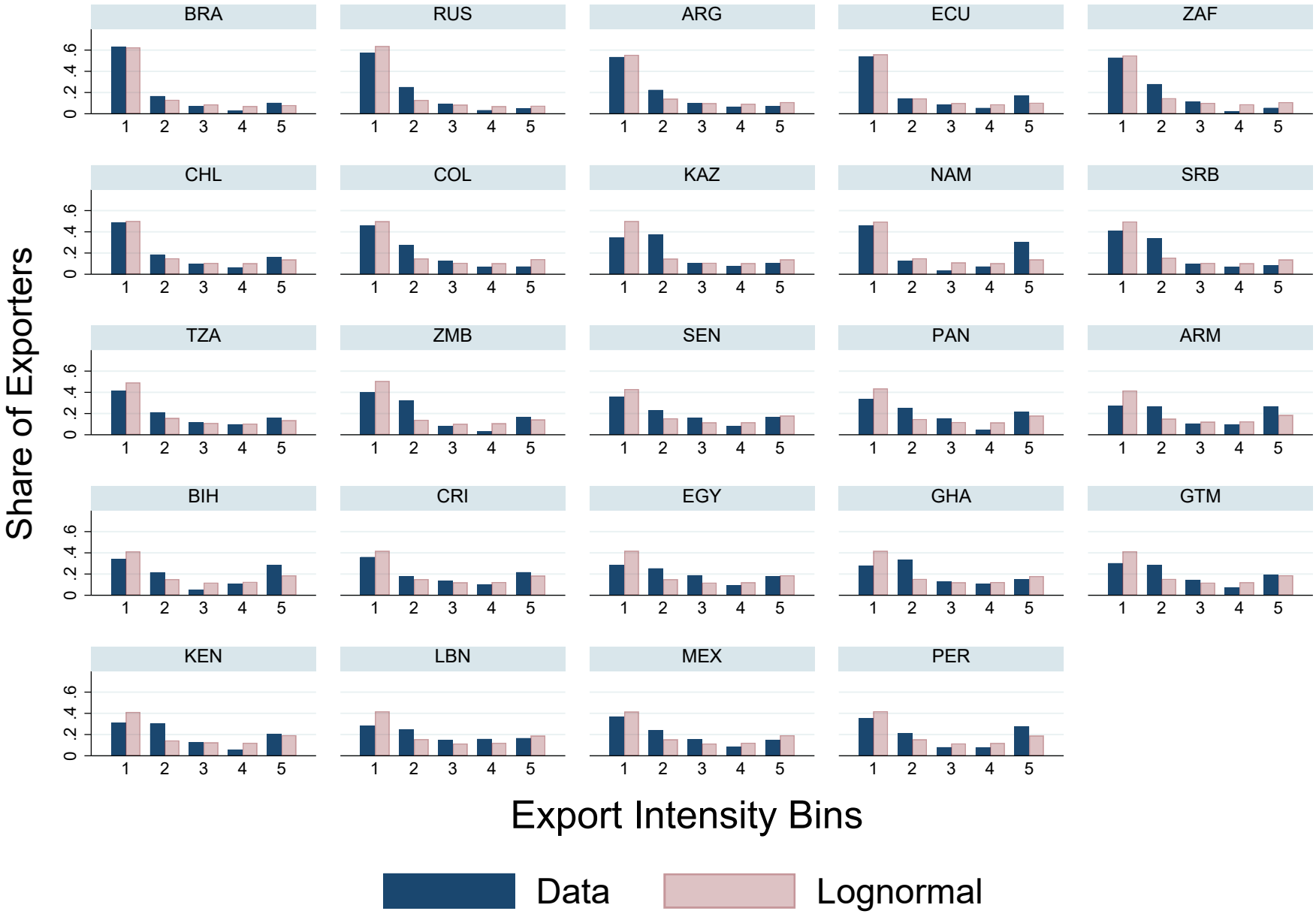


- **The conditions for bimodality are satisfied in every country!**

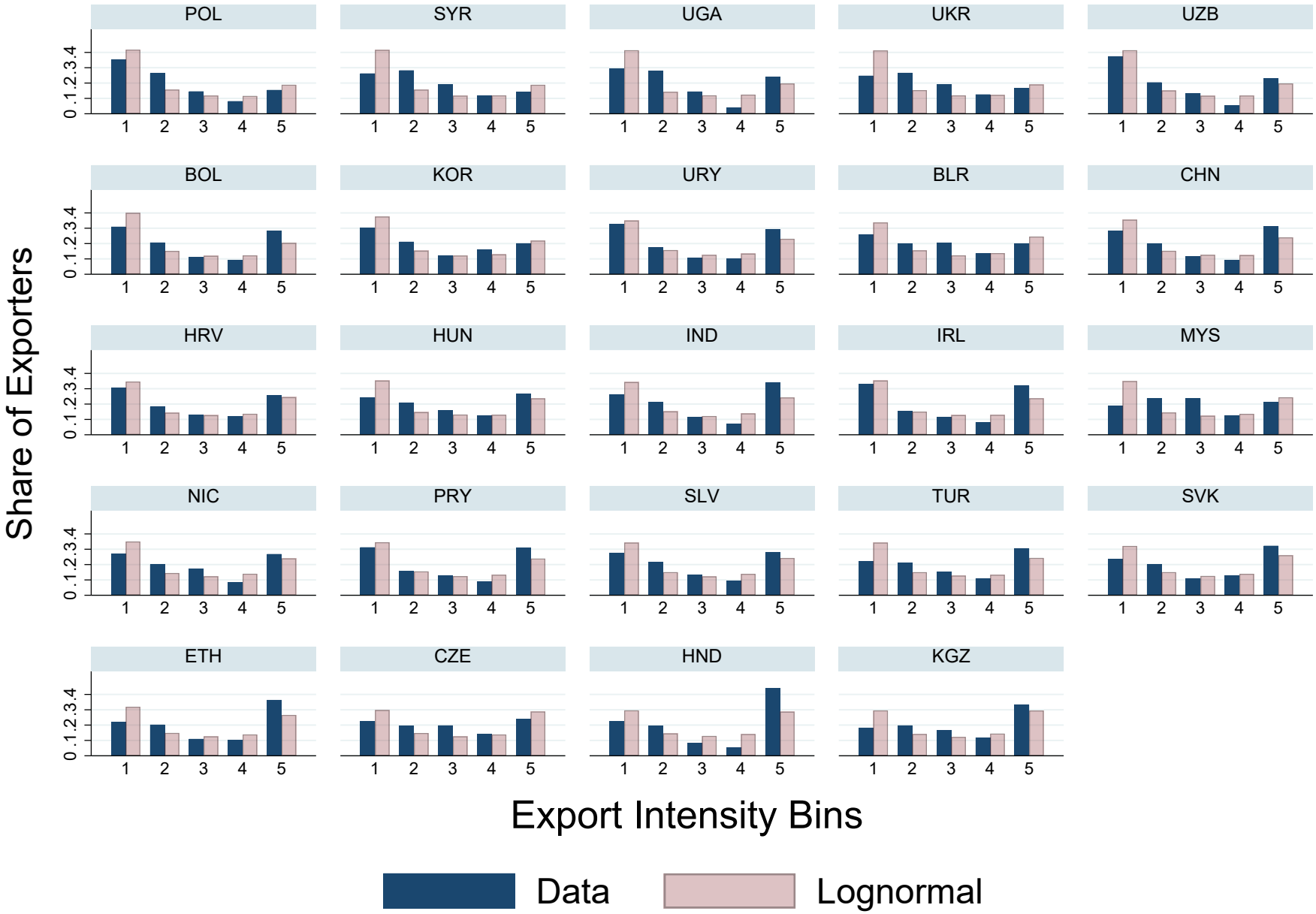
Bimodality and relative market size

- **How do we reconcile the results showing that the distribution of export intensity in all countries is bimodal with the fact that we cannot reject the null hypothesis of unimodality for 1/3 of the countries in our data?**
- Given a sufficiently high variance of revenue shifters, differences in countries' relative market size vis-à-vis the rest of the world explain both the variation in the distributions of export intensity around the world remarkably well and the reason why the dip test fails to reject unimodality for some countries
- Firms in all countries draw both of their revenue shifters from a distribution with the same shape parameter
- All the variation in the distribution of export intensity across countries is due to differences in their market size relative to the rest of the world
- We estimate the 'restricted' single-shape parameter model by pooling together data across all countries and weight each observation by the inverse of the number of exporters in each country to ensure that each country receives the same weight in the estimation

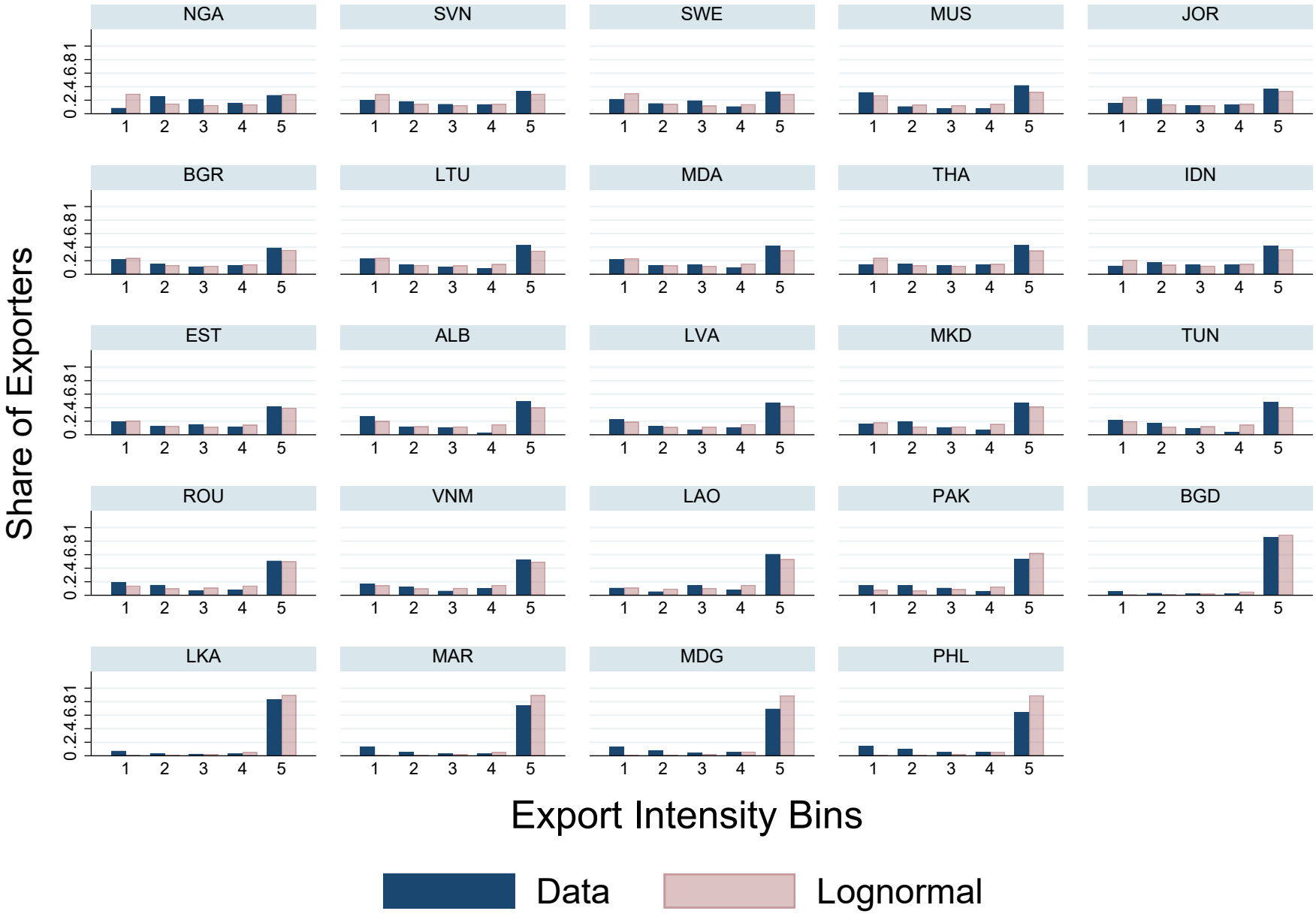
Model fit


BRA
RUS
ARG
ECU
ZAF
CHL
COL
KAZ
NAM
SRB
TZA
ZMB
SEN
PAN
ARM
BIH
CRI
EGY
GHA
GTM
KEN
LBN
MEX
PER
0 .2 .4 .6
1 2 3 4 5
Share of Exporters
Export Intensity Bins
Data
Lognormal

Model fit, cont'd

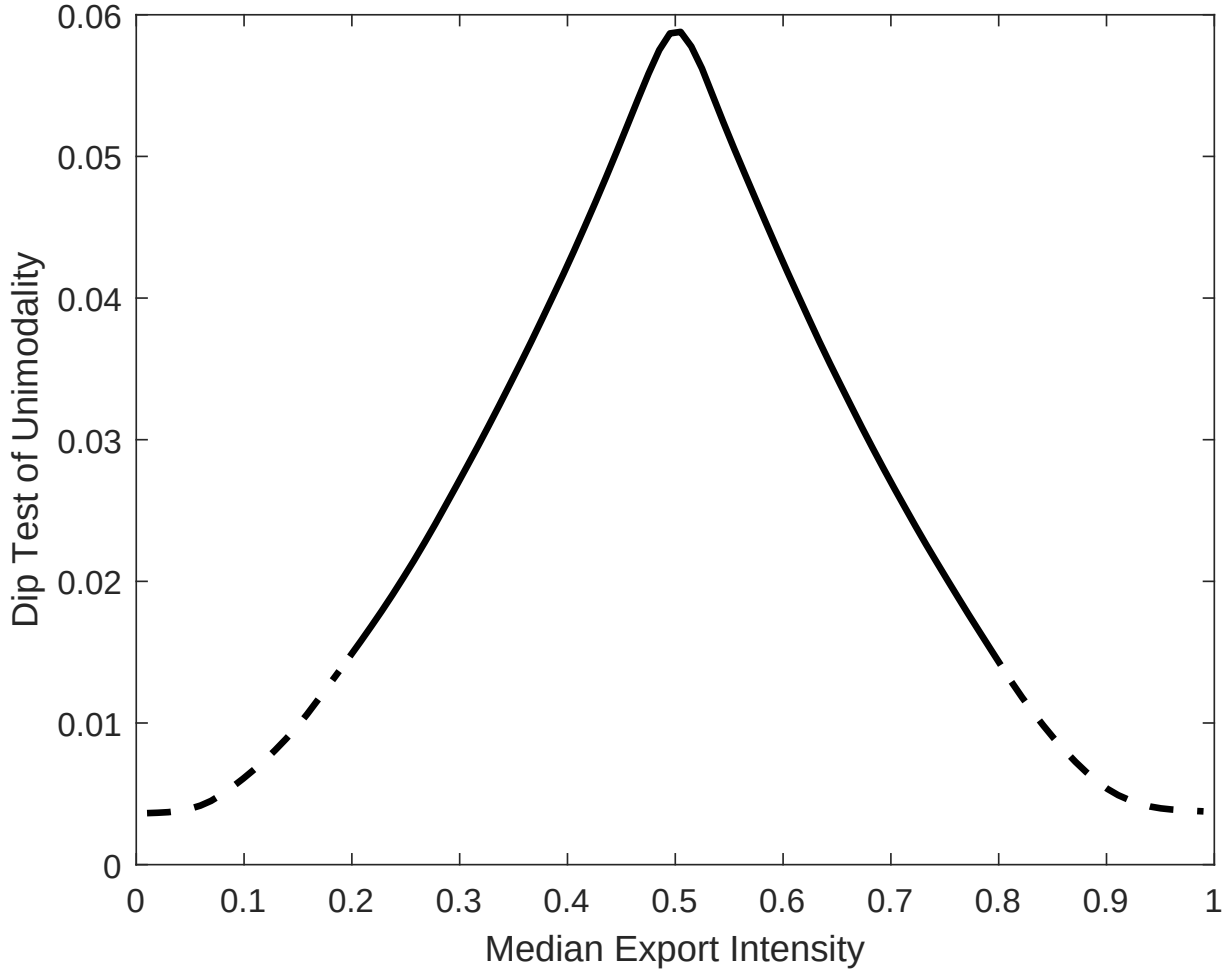


Model fit, cont'd



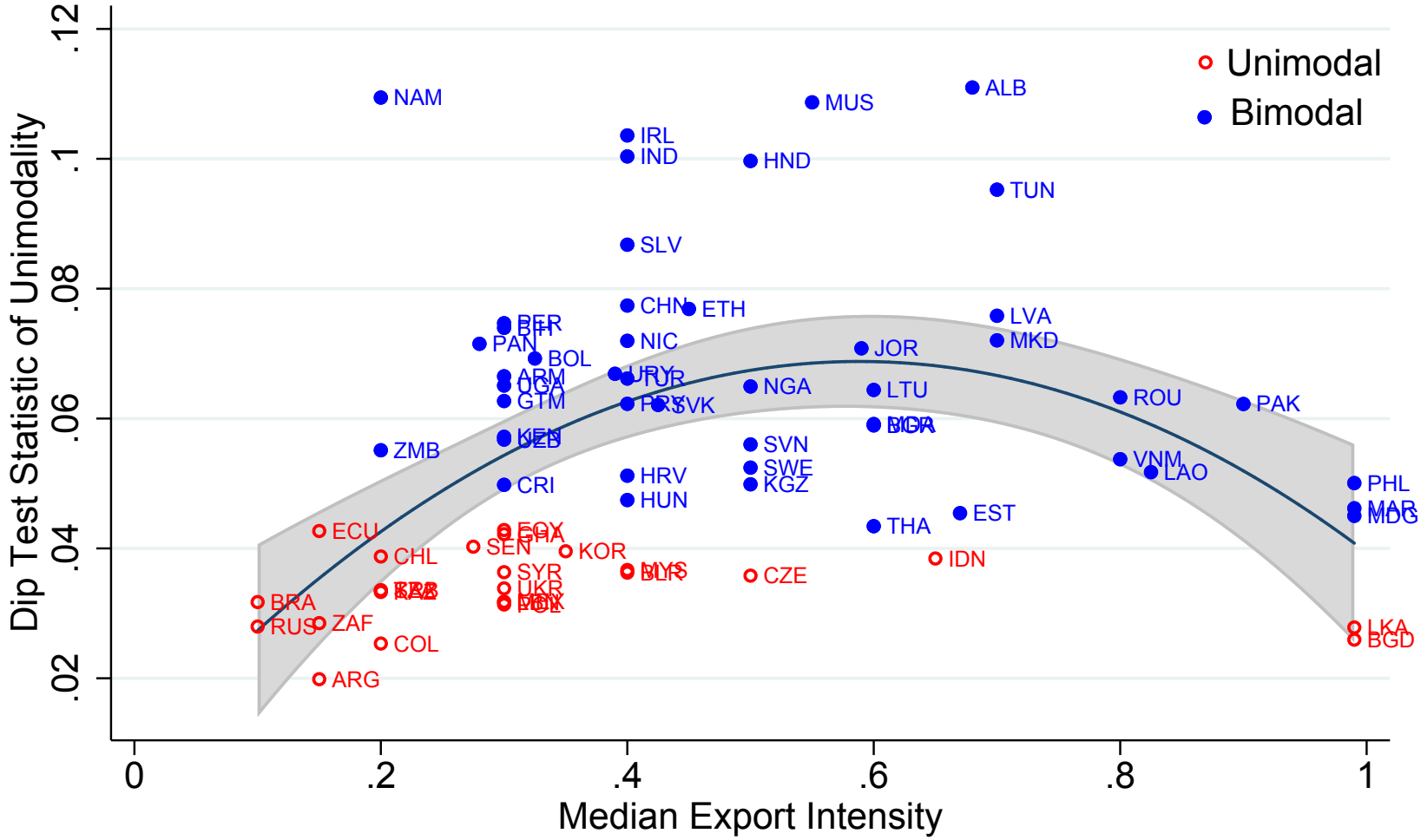
Takeaway

- The point-estimate for the single shape parameter across all countries is 6.489
- The Vuong (1989) likelihood ratio test reveals that we cannot discriminate between the full and restricted models in 51 out of 72 countries in our data
- Relying only on variation in countries' relative market size and a unique shape parameter, our model reproduces the wide range of shapes observed in the distribution of export intensity across the world extremely well
- The correlation between the observed and predicted shares of exporters across export intensity bins of 0.88
- Our model is able to generate unimodal distributions where the majority of exporters exhibit either very low or very high export intensity just as well as distributions featuring prominent



- when a country's domestic market is either very small or very large relative to the export market, the height of the minor mode shrinks so much that the distribution appears unimodal
- conversely, in countries for which the size of domestic and export markets are similar, the distribution of export intensity displays prominent twin peaks

- Our model predicts an inverse-U relationship between a measure of bimodality (the dip statistic) and a country's relative size



- **...which also bears in the data!**

Accounting for the High Dispersion of Revenue Shifters

- We re-estimate the single-shape-parameter model excluding different subsets of observations one at a time
- This allows us to verify that the conditions for bimodality are satisfied in each subsample
- Since we are estimating the sum of the variances of revenue shifters, comparing the different estimates permits us to gauge the contribution of each group of observations to the benchmark shape parameter estimated using the whole data
- We only keep countries or country pairs with at least 50 exporters and each observation is weighted so that each country or each dyad receives an equal weight in the estimation

Accounting for the High Dispersion of Revenue Shifters

	Full Sample	Excluding					Countries		
		Foreign-owned	Processing	Pure exporters	Old	Single-prod.	without ESR	OECD	non-OECD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Total Export Intensity at the Country-level									
$\sigma^2_{zd} + \sigma^2_{zx}$	6.489	6.044	6.056	3.810	7.106	5.863	5.762	5.497	6.670
	(0.050)	(0.053)	(0.051)	(0.033)	(0.096)	(0.070)	(0.069)	(0.111)	(0.056)
Countries	72	72	72	72	58	61	39	11	61
Obs.	33,224	26,217	28,375	26,647	10,915	13,849	13,691	4,916	28,308
Panel B: Bilateral Export Intensity									
$\sigma^2_{zd} + \sigma^2_{zx}$	8.351	8.663	7.986	3.890	9.985	7.505	7.468	11.250	6.419
	(0.246)	(0.328)	(0.264)	(0.147)	(0.544)	(0.342)	(0.712)	(0.515)	(0.247)
Country pairs	20	15	16	14	17	17	3	8	11
Obs.	2,299	1,396	1,832	1,410	674	965	220	953	1,346

Interpretation

- The estimated sum of the variances of revenue shifters remains sufficiently high to generate bimodal export intensity distributions in all cases
- Excluding one group of exporters at a time results in similar reductions of the sum of the variances of revenue shifters of the order of 7 to 11%
- Excluding pure exporters reduces the dispersion of revenue shifters by 42%—crucially, the dispersion of revenue shifters we estimate is still sufficiently high to generate twin peaks
- Dispersion of revenue shifters is 20% higher among exporters in developing countries compared to that from developed ones
- Twin peaks are more prominent among exporters selling to OECD countries than among those selling their output to other developing markets
- **No single factor can fully account for the dispersion of revenue shifters necessary to engender twin peaks!**

Putting all together

- To what extent is $\sigma_d^2 + \sigma_x^2$ estimated at the country level related to observable country characteristics?

Dependent Variable	$\sigma_{zd}^2 + \sigma_{zx}^2$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log GDP per capita	-0.224*					-0.051	0.104
	(1.99)					(0.44)	(1.28)
Subsidies with ESR		0.348***				0.320**	0.177*
		(3.06)				(2.59)	(1.98)
Share of foreign-owned firms			0.403***				0.276***
			(2.75)				(2.90)
Share of processing exporters				0.498***			0.217**
				(4.15)			(2.30)
Share of pure exporters					0.683***		0.526***
					(3.44)		(2.84)
Observations	72	72	72	72	72	72	72
R-squared	0.050	0.121	0.163	0.248	0.467	0.123	0.617

- proxies for policies that incentivize firms to export a high share of their output explain 12-47% of the cross-country variation in $\sigma_d^2 + \sigma_x^2$
- the share of pure exporters operating in a country accounts for the lion's share of the variation
- all proxies for individual policies remain significant, even when considered jointly
- **however, policies cannot fully account for the prevalence of export intensity distributions that exhibit twin peaks**

Conclusion

Conclusion

1. The distribution of export intensity among exporters varies a lot across countries
   - It is not generally true that most exporters in a country sell most of their output domestically
   - In many countries a large share of very high and low-intensity exporters operate alongside each other
2. If firms differ substantially in how successful they are in selling in different markets → a workhorse model of trade with heterogeneous firms can explain the cross-country variation in export intensity distributions very well
   - The shape of the distribution is determined by a country's relative size vis-à-vis the rest of the world
3. Presence of multinational affiliates, export processing, export subsidies, level of development, industrial composition account for some of the dispersion that produces bimodality—**but a lot of it remains unexplained!**