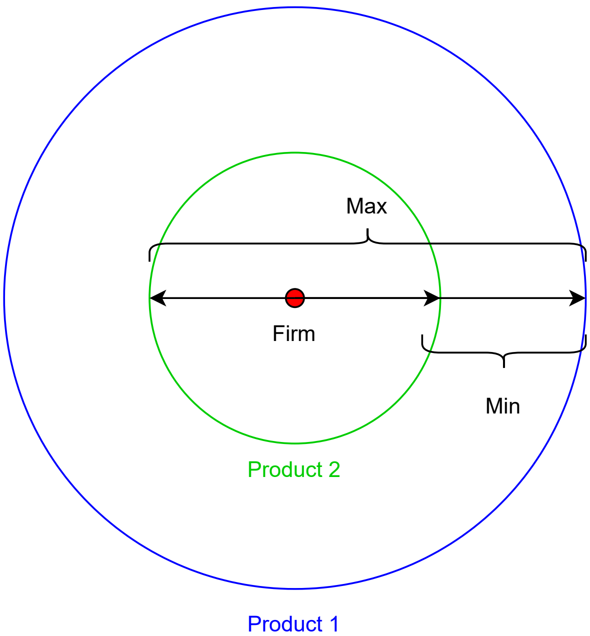
Supplemental Appendix

B Geometric interpretation of distance measures

To illustrate the first step in our algorithm geometrically, consider a firm ω producing two products 1 and 2 with advantages $a_{\omega,1}$ and $a_{\omega,2}$. The distance of the firm to the two products equal these advantages and they can be visualized spatially by placing each product around the firm as in Figure B.1, where product 1 is on the circle of radius $a_{\omega,1}$ and product 2 is on the circle of radius $a_{\omega,2}$. One can imagine if the firm produced product 2 perfectly with $a_{\omega,2} = 0$ (i.e., with an advantage at the zero-th percentile of costs), the distance from the firm to product 1 would be the distance between products 1 and 2, labeled $d_{1,2}$. An application of the triangle inequality in this example provides numerical bounds on this distance

$$|a_{\omega,1} - a_{\omega,2}| \leq d_{1,2} \leq a_{\omega,1} + a_{\omega,2}. \tag{B.1}$$

Figure B.1: Product Distance Bounds in Co-Production



To illustrate this, consider the example in Figure B.1. Conditional on observing the firm's advantages in products 1 and 2, we can apply the triangle inequality to recover bounds of the possible distance between the two products, $d_{1,2}$. The minimum distance arises when both products lie on the same line extending in one direction from the firm's location where both products are located to the right. In this case, the distance between the products equals the absolute difference in their distances to the firm, i.e., $|a_{\omega,1} -$

$a_{\omega,2}|$. Conversely, the maximum distance occurs when the two products lie on opposite directions through the firm with product 1 to the right and product 2 to the left, so that the distance between them equals the sum of the two radii, $a_{\omega,1} + a_{\omega,2}$.

$$d_{\nu,\tau} \equiv \sum_{\omega} \frac{a_{\omega,\nu} + a_{\omega,\tau} + |a_{\omega,\nu} - a_{\omega,\tau}|}{2N_{\nu,\tau}} = \sum_{\omega} \frac{\max\{a_{\omega,\nu}, a_{\omega,\tau}\}}{N_{\nu,\tau}}. \tag{B.2}$$

Equation (B.2) is an average between the upper and lower distance bounds of Equation (B.1) and therefore consistent with Equation (B.1) taken as an average.[28] It is mathematically equivalent to the average of the maximum of the two advantages.

C Estimation

C.1 Product Adoption by Proximity

As a robustness check, we estimate a version of the product adoption equation that uses rank-based measures of distance. Specifically, we estimate

$$\text{Intro}_{\omega,\nu,t} = \beta \cdot \text{Rank}_{\omega,\nu,t-1} + \phi_{\omega,t} + \rho_{\nu,t} + \epsilon_{\omega,\nu,t}, \tag{C.1}$$

where $\text{Rank}_{\omega,\nu,t-1}$ denotes the rank of products not produced by firm ω based on counterfactual advantage (CN distance or product relatedness) in year $t-1$. The closest product receives rank 1.

Table C.1 presents the baseline estimates. Across all specifications, we find that product introduction rates decline with distance rank.[29]

The results are robust to controlling for CN distance fixed effects (Table C.2) and to restricting the sample to firm-product pairs with zero RCA-based relatedness (Table C.3).

[28] Given the bounds in Equation (B.1), one might be tempted to use order statistics to obtain a tighter fit. For example, instead of averaging, one could take the *minimum* of the upper bounds and the *maximum* of the lower bounds across observations to construct a narrower range of potential distances. In the absence of measurement error, this approach would indeed be preferable. However, in the presence of errors, the minimum of the upper bounds can be smaller than the maximum of the lower bounds, especially as the number of observations grows. Empirically, we observe such a pattern in a non-negligible fraction of our sample. For this reason, we adopt a more robust approach that averages the distance bounds across firms to mitigate the influence of errors and outliers.

[29] For the CN classification, product pairs within the same distance bin are randomly ranked due to the discrete nature of the taxonomy. Despite this limitation, the rank-based approach still captures meaningful variation in product proximity.

Table C.1: Product Introduction and Product Rankings

	Dependent Variable: Dummy=1 for Product Introduction			
	(CES)	(DGKP)	(CN)	(RCA)
Lagged Rank (=1 closest)	-0.005***	-0.005***	-0.008***	-0.009***
	(0.001)	(0.001)	(0.000)	(0.000)
Product-Time FE	Yes	Yes	Yes	Yes
Firm-Time FE	Yes	Yes	Yes	Yes
R^2	0.07	0.07	0.07	0.07
# Obs.	645290	645290	645290	645290

Results from OLS estimation of (C.1). Cluster standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%.

Table C.2: Product Introduction and Product Rankings

	Dependent Variable: Dummy=1 for Product Introduction		
	(CES)	(DGKP)	(RCA)
Lagged Rank (=1 closest)	-0.002***	-0.003***	-0.006***
	(0.001)	(0.001)	(0.000)
Product-Time FE	Yes	Yes	Yes
Firm-Time FE	Yes	Yes	Yes
CN Distance FE	Yes	Yes	Yes
R^2	0.08	0.08	0.08
# Obs.	645290	645290	645290

Results from OLS estimation of (C.1). Cluster standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%.

Table C.3: Product Introduction and Product Rankings

	Dependent Variable: Dummy=1 for Product Introduction		
	(CES)	(DGKP)	(CN)
Lagged Rank (=1 closest)	-0.002***	-0.002***	-0.005***
	(0.000)	(0.000)	(0.000)
Product-Time FE	Yes	Yes	Yes
Firm-Time FE	Yes	Yes	Yes
Lagged Relatedness =	0	0	0
R^2	0.08	0.08	0.08
# Obs.	542445	542445	542445

Results from OLS estimation of (C.1). Cluster standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%.

C.2 Product Drops

In this section, we quantify the ability of the product space to predict product drops, which we define as non-production following production at entry. As the measure of distance within a firm is based on advantages that we observe, the estimation of the full product space is not necessary. As a result, we consider this exercise as a sanity check for our measure of absolute advantages. For each firm, we select the products that the firm produces in its first year in the data. For instance, if a firm enters the data set in 2000, the sample of products we consider are the products in the cluster of the firm that the firm produces in 2000. For each product in the years after entry, we compute a production indicator $\text{Drop}_{\omega,\nu,t}$ for product ν, firm ω, and year t, which equals 1 if the product is not produced (i.e., dropped) and zero otherwise.

We estimate the following regression

$$\text{Drop}_{\omega,\nu,t} = \beta \cdot \ln\left(1 + \tilde{a}_{\omega,\nu,t-1}\right) + \phi_{\omega,t} + \rho_{\nu,t} + \varepsilon_{\omega,\nu,t} \tag{C.2}$$

where the main independent variable is the lagged transformed counterfactual advantage $T(\tilde{a}_{\omega,\nu,t-1}) = \ln(1 + \tilde{a}_{\omega,\nu,t-1})$, as in our baseline product space.

Results are shown in Table C.4 and C.5. Products with less advantage are more likely to be dropped, as the coefficient on the lagged advantage is positive and statistically significant in each specification. As shown for the case of product introduction, using a discrete classification cannot shed light on which products, within a certain aggregation of codes, are more likely to get dropped, while our measure based on advantage is able to distinguish between these products.

Table C.4: Product Drop and Advantage

	Dependent Variable: Dummy=1 for Product Drop			
	(CES)	(DGKP)	(CN)	(RCA)
Lagged Advantage	6.419***	1.550***	0.155***	1.918***
	(0.571)	(0.139)	(0.032)	(0.249)
Product-Time FE	Yes	Yes	Yes	Yes
Firm-Time FE	Yes	Yes	Yes	Yes
R^2	0.60	0.60	0.60	0.60
# Obs.	22642	22642	22642	22642

Results from OLS estimation of (C.2). Cluster standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%.

Table C.5: Product Drop and Advantage

	Dependent Variable: Dummy=1 for Product Drop		
	(CES)	(DGKP)	(RCA)
Lagged Advantage	4.737***	1.195***	1.589***
	(0.619)	(0.149)	(0.248)
Product-Time FE	Yes	Yes	Yes
Firm-Time FE	Yes	Yes	Yes
CN Distance FE	Yes	Yes	
R^2	0.60	0.60	0.60
# Obs.	22642	22642	22642

Results from OLS estimation of (C.2). Cluster standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%.

C.3 Export Shocks and Product Adoption

Table C.6: Export Shocks - IV = Autor et al. (2013) Countries

	Dependent Variable: Dummy=1 for Product Introduction							
	(CES)	(CES)	(DGKP)	(DGKP)	(CN)	(CN)	(RCA)	(RCA)
Lagged Advantage	47.600	546.423***	25.540	101.694***	3.720	18.018	-60.616***	-50.157***
	(66.152)	(158.317)	(17.992)	(38.854)	(5.447)	(12.369)	(4.094)	(11.136)
Lagged Advantage	-10.772*	-54.372***	-3.429**	-9.974***	-0.561	-1.806*	-3.737***	-25.362
X Exports	(5.650)	(13.698)	(1.522)	(3.309)	(0.465)	(1.067)	(1.088)	(21.181)
Firm-Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm-Product FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prod-Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Spec.	OLS	IV	OLS	IV	OLS	IV	OLS	IV
# Obs.	13782	13782	13782	13782	13782	13782	13782	13782
F-Stat		191.63		288.95		228.46		23.54

Results from OLS and IV estimation of (28). Cluster standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%. We instrument Log Exports$_{\nu t}$ $\times \ln\left(1+\tilde{a}_{\omega,\nu,t-1}\right)$ with Log Exports IV$_{\nu t}$ $\times \ln\left(1+\tilde{a}_{\omega,\nu,t-1}\right)$ where Log Exports IV$_{\nu t}$ = Log $\sum_{k\in K}$ Exports$_{k\nu t}$ where K denotes the same set of countries as Autor et al. (2013): Australia, Finland, Germany, Japan, New Zealand, Spain, Switzerland, and USA. F-Stat denotes the Kleibergen-Paap (KP) rank Wald F-statistic.

Table C.7: Export Shocks - IV = EU Countries

	Dependent Variable: Dummy=1 for Product Introduction							
	(CES)	(CES)	(DGKP)	(DGKP)	(CN)	(CN)	(RCA)	(RCA)
Lagged Advantage	47.600	509.799***	25.540	116.219***	3.720	14.576	-60.616***	-43.784***
	(66.152)	(140.652)	(17.992)	(35.409)	(5.447)	(11.370)	(4.094)	(15.438)
Lagged Advantage	-10.772*	-51.171***	-3.429**	-11.222***	-0.561	-1.506	-3.737***	-38.540
X Exports	(5.650)	(12.167)	(1.522)	(3.016)	(0.465)	(0.983)	(1.088)	(30.896)
Firm-Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm-Product FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prod-Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Spec.	OLS	IV	OLS	IV	OLS	IV	OLS	IV
# Obs.	13782	13782	13782	13782	13782	13782	13782	13782
F-Stat		257.25		419.77		304.73		11.75

Results from OLS and IV estimation of (28). Cluster standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%. We instrument Log Exports$_{\nu t}$ $\times \ln(1+\tilde{a}_{\omega,\nu,t-1})$ with Log Exports IV$_{\nu t}$ $\times \ln(1+\tilde{a}_{\omega,\nu,t-1})$ where Log Exports IV$_{\nu t}$ = Log $\sum_{k\in K}$ Exports$_{k\nu t}$ where K denotes the set of EU countries: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, and Sweden. F-Stat denotes the Kleibergen-Paap (KP) rank Wald F-statistic.

C.4 First Stage Regressions

Table C.8: Export Shocks - IV. Autor et al. (2013) Countries (First Stage)

	Dependent Variable: Lagged Advantage X Exports			
	(CES)	(DGKP)	(CN)	(RCA)
Lagged Advantage X Exports IV	0.889***	0.908***	0.872***	0.040***
	(0.004)	(0.004)	(0.006)	(0.003)
Firm-Time FE	Yes	Yes	Yes	Yes
Firm-Product FE	Yes	Yes	Yes	Yes
Prod-Time FE	Yes	Yes	Yes	Yes
# Obs.	13782	13782	13782	13782
R^2	1.00	0.99	1.00	0.93

Results from OLS estimation of the first stage of (28). Cluster standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%. Log Exports IV$_{\nu t}$ = Log $\sum_{k\in K}$ Exports$_{k\nu t}$ where K denotes the same set of countries as Autor et al. (2013): Australia, Finland, Germany, Japan, New Zealand, Spain, Switzerland, and USA.

Table C.9: Export Shocks - IV. EU Countries (First Stage)

	Dependent Variable: Lagged Advantage X Exports			
	(CES)	(DGKP)	(CN)	(RCA)
Lagged Advantage X Exports IV	0.857***	0.871***	0.838***	0.038***
	(0.004)	(0.004)	(0.005)	(0.003)
Firm-Time FE	Yes	Yes	Yes	Yes
Firm-Product FE	Yes	Yes	Yes	Yes
Prod-Time FE	Yes	Yes	Yes	Yes
# Obs.	13782	13782	13782	13782
R^2	1.00	0.99	1.00	0.93

Results from OLS estimation of the first stage of (28). Cluster standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%. Log Exports IV$_{\nu t}$ = Log $\sum_{k\in K}$ Exports$_{k\nu t}$ where K denotes the set of EU countries: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, and Sweden.

C.5 Advantage transformed in Levels and IHS

In this section, we replicate the baseline results presented in the main text but using transformations corresponding to Levels or Inverse Hyperbolic Sine (IHS) for robustness. Summary statistics for the distances measured from product-to-product and firm-to-product are provided in Table C.10 for Levels and Table C.11 for IHS. The results for product introduction with lagged distance modulating various fixed effects and sub-samples are all qualitatively similar to those above, so we omit them for brevity.

Table C.10: Product Space Summary Statistics (Levels Transformation, Year 2000)

	Product-to-Product		**Firm-to-Product**	
	CES	DGKP	CES	DGKP
Average	0.87	2.74	0.53	1.39
Std. Dev.	0.49	2.16	0.22	0.71
5th Perc.	0.21	0.52	0.22	0.38
10th Perc.	0.30	0.79	0.27	0.59
25th Perc.	0.48	1.00	0.37	0.91
50th Perc.	0.80	1.92	0.50	1.31
75th Perc.	1.11	3.99	0.65	1.76
90th Perc.	1.53	5.93	0.81	2.19
95th Perc.	1.75	6.91	0.91	2.51

The table reports the distribution of product-to-product and firm-to-product distances in the year 2000. The sample comprises all product-to-product and all firm-to-product distances across all clusters. For confidentiality reasons, we divide distances data in 100 bins, compute the average distance within bin, and report these values in the rows with percentiles.

Table C.11: Product Space Summary Statistics (IHS Transformation, Year 2000)

	Product-to-Product		**Firm-to-Product**	
	CES	DGKP	CES	DGKP
Average	0.53	0.97	0.32	0.75
Std. Dev.	0.29	0.40	0.12	0.27
5th Perc.	0.15	0.36	0.15	0.26
10th Perc.	0.21	0.50	0.18	0.40
25th Perc.	0.31	0.72	0.24	0.59
50th Perc.	0.45	0.88	0.31	0.75
75th Perc.	0.68	1.16	0.39	0.90
90th Perc.	0.91	1.43	0.47	1.04
95th Perc.	1.06	1.66	0.52	1.13

The table reports the distribution of product-to-product and firm-to-product distances in the year 2000. The sample comprises all product-to-product and all firm-to-product distances across all clusters. For confidentiality reasons, we divide distances data in 100 bins, compute the average distance within bin, and report these values in the rows with percentiles.

Tables C.12 and C.13 examine the impact of export demand shocks on product introduction, with the respective first stage regressions detailed in Tables C.14 and C.15. The estimates show that the qualitative conclusions are the same as in the baseline specification.

Table C.12: Export Shocks and Product Adoption (Levels Transformation)

	Dependent Variable: Dummy=1 for Product Introduction			
	(CES)	(CES)	(DGKP)	(DGKP)
Lagged Advantage	10.117	59.895***	1.462	11.970*
	(10.153)	(19.114)	(4.096)	(6.356)
Lagged Advantage X Exports	-1.889**	-6.320***	-0.383	-1.271**
	(0.878)	(1.689)	(0.335)	(0.530)
Firm-Time FE	Yes	Yes	Yes	Yes
Firm-Product FE	Yes	Yes	Yes	Yes
Prod-Time FE	Yes	Yes	Yes	Yes
Spec.	OLS	IV	OLS	IV
# Obs.	13782	13782	13782	13782
F-Stat		526.48		595.68

Results from OLS and IV estimation of (28). Clustered standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%. We instrument Log Exports$_{\nu t}$ $\times$ $\tilde{a}_{\omega,\nu,t-1}$ with Exports IV$_{\nu t}$ $\times$ $\tilde{a}_{\omega,\nu,t-1}$ where Exports IV$_{\nu t}$ = Log $\sum_{k \neq DNK}$ Exports$_{k\nu t}$ is the total exports of all countries except Denmark. F-Stat denotes the Kleibergen-Paap (KP) rank Wald F-statistic.

Table C.13: Export Shocks and Product Adoption (IHS Transformation)

	Dependent Variable: Dummy=1 for Product Introduction			
	(CES)	(CES)	(DGKP)	(DGKP)
Lagged Advantage	69.421	302.434***	15.288	70.766***
	(47.155)	(84.589)	(13.665)	(23.694)
Lagged Advantage X Exports	-10.349**	-30.700***	-2.203*	-6.968***
	(4.023)	(7.319)	(1.155)	(2.017)
Firm-Time FE	Yes	Yes	Yes	Yes
Firm-Product FE	Yes	Yes	Yes	Yes
Prod-Time FE	Yes	Yes	Yes	Yes
Spec.	OLS	IV	OLS	IV
# Obs.	13782	13782	13782	13782
F-Stat		456.18		577.29

Results from OLS and IV estimation of (28). Clustered standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%. We instrument Log Exports$_{\nu t}$ $\times \ln\left(\tilde{a}_{\omega,\nu,t-1} + \sqrt{1+\tilde{a}^2_{\omega,\nu,t-1}}\right)$ with Exports IV$_{\nu t}$ $\times \ln\left(\tilde{a}_{\omega,\nu,t-1} + \sqrt{1+\tilde{a}^2_{\omega,\nu,t-1}}\right)$ where Exports IV$_{\nu t}$ = Log $\sum_{k \neq DNK}$ Exports$_{k\nu t}$ is the total exports of all countries except Denmark. F-Stat denotes the Kleibergen-Paap (KP) rank Wald F-statistic.

Table C.14: Export Shocks - IV = All Countries (Levels, First Stage)

	Dependent Variable: Lagged Advantage X Exports	
	(CES)	(DGKP)
Lagged Advantage X Exports IV	0.768***	0.799***
	(0.004)	(0.003)
Firm-Time FE	Yes	Yes
Firm-Product FE	Yes	Yes
Prod-Time FE	Yes	Yes
# Obs.	13782	13782
R^2	1.00	1.00

Results from OLS estimation of the first stage of (28). Clustered standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%. Exports $\text{IV}_{\nu t} = \text{Log} \sum_{k \neq DNK} \text{Exports}_{k\nu t}$.

Table C.15: Export Shocks - IV = All Countries (IHS, First Stage)

	Dependent Variable: Lagged Advantage X Exports	
	(CES)	(DGKP)
Lagged Advantage X Exports IV	0.782***	0.792***
	(0.003)	(0.003)
Firm-Time FE	Yes	Yes
Firm-Product FE	Yes	Yes
Prod-Time FE	Yes	Yes
# Obs.	13782	13782
R^2	1.00	1.00

Results from OLS estimation of the first stage of (28). Clustered standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%. Exports $\text{IV}_{\nu t} = \text{Log} \sum_{k \neq DNK} \text{Exports}_{k\nu t}$.

D Larger Cluster Analysis

In this section, we replicate the baseline results presented in the main text, using a more aggregate definition of sectors. Specifically, we consider six sectors: Animals/Vegetables/Food (CN 2-digit 01-24), Minerals/Chemicals/Plastics (25-40), Textiles/Footwear (41-43, 50-67), Stone/Metals (68-83), Machinery/Transportation (84-89), and Miscellaneous (44-49, 90-97). This replication uses the same distance transformation as the baseline. Tables D.1 and D.2 present the descriptive statistics for the cluster characteristics. Table D.3 provides summary statistics for the distances measured from product-to-product and firm-to-product. Notice that relative to our baseline results, the average distances increase by little, with the larger increases concentrated in the higher percentiles.

Tables D.4, D.5, and D.6 detail the correlation between the likelihood of product introduction and the rankings based on the firm's product distance. Tables D.7, D.8, and D.9 explore the relationship between the probability of introducing a product and its lagged distance to the firm. Table D.10 examines the impact of export demand shocks on product introduction, with the first stage regression detailed in Table D.11.

Table D.1: Cluster Descriptive Statistics

	Number of Products					Number of Firms				
Year	Avg.	Std.	Med.	25P.	75P.	Avg.	Std.	Med.	25P.	75P.
2000	207	141	191	84	260	134	64	111	90	196
2001	211	126	189	113	258	150	66	129	115	190
2002	208	141	174	91	260	143	72	129	83	190
2003	200	137	171	85	247	135	68	121	77	183
2004	205	123	175	118	250	144	56	133	128	182
2005	205	124	166	117	249	141	63	127	120	194
2006	211	128	174	124	258	136	60	130	105	192
2007	160	102	128	95	203	87	33	88	78	106
2008	163	107	130	111	208	84	34	88	79	93
2009	163	110	136	108	209	86	37	88	78	100
2010	173	97	143	123	203	100	41	111	87	131
2011	190	106	156	133	210	108	48	129	91	140
2012	195	103	162	135	225	100	40	112	82	133
2013	205	104	169	154	239	104	43	118	85	129
2014	212	108	167	166	267	105	44	117	94	129
2015	214	114	171	156	286	106	43	115	95	139
2016	208	120	160	154	288	100	46	109	80	140
2017	217	119	171	150	290	106	45	115	107	137
2018	219	123	158	151	292	107	45	118	100	140
Average	**198**	**118**	**162**	**125**	**247**	**114**	**50**	**115**	**93**	**150**

In each year, we compute average (Avg.), standard deviation (Std.), median (Med.), and 25th and 75th percentiles (25P. and 75P.) of the number of products (first five columns) and number of firms (last five columns) across clusters. In each year, there are 6 clusters (for 6 sectors defined as groups of CN 2-digit codes). The last row (Average) reports the average of the statistics across years.

Table D.2: Cluster Descriptive Statistics

	Number of Products		Number of Firms	
Cluster	Avg.	Std.	Avg.	Std.
Animals/Vegetables/Food	407	36	158	25
Machinery/Transportation	145	17	109	12
Minerals/Chemicals/Plastics	247	31	107	19
Miscellaneous	84	20	158	66
Stone/Metals	127	29	116	24
Textiles/Footwear	178	33	39	22

In each cluster, we compute average (Avg.) and standard deviation (Std.) of the number of products (first two columns) and number of firms (last two columns) across years.

Table D.3: Product Space Summary Statistics (Baseline Transformation, Year 2000)

	Product-to-Product		**Firm-to-Product**	
	CES	DGKP	CES	DGKP
Average	0.39	0.82	0.25	0.65
Std. Dev.	0.20	0.38	0.09	0.23
5th Perc.	0.12	0.32	0.11	0.29
10th Perc.	0.16	0.41	0.14	0.37
25th Perc.	0.24	0.59	0.18	0.50
50th Perc.	0.35	0.75	0.24	0.62
75th Perc.	0.48	0.99	0.29	0.77
90th Perc.	0.69	1.23	0.35	0.94
95th Perc.	0.74	1.42	0.40	1.05

The table reports the distribution of product-to-product and firm-to-product distances in the year 2000. The sample comprises all product-to-product and all firm-to-product distances across all clusters. The product space is estimated using marginal costs in logs. For confidentiality reasons, we divide distances data in 100 bins, compute the average distance within bin, and report these values in the rows with percentiles.

Table D.4: Product Introduction and Product Rankings

	Dependent Variable: Dummy=1 for Product Introduction			
	(CES)	(DGKP)	(CN)	(RCA)
Lagged Rank (=1 closest)	-0.004***	-0.004***	-0.004***	-0.006***
	(0.000)	(0.000)	(0.000)	(0.000)
Product-Time FE	Yes	Yes	Yes	Yes
Firm-Time FE	Yes	Yes	Yes	Yes
R^2	0.05	0.05	0.05	0.05
# Obs.	1332970	1332970	1332970	1332970

Results from OLS estimation of (C.1). Clustered standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%.

Table D.5: Product Introduction and Product Rankings

	Dependent Variable: Dummy=1 for Product Introduction		
	(CES)	(DGKP)	(RCA)
Lagged Rank (=1 closest)	-0.002***	-0.002***	-0.004***
	(0.000)	(0.000)	(0.000)
Product-Time FE	Yes	Yes	Yes
Firm-Time FE	Yes	Yes	Yes
CN Distance FE	Yes	Yes	Yes
R^2	0.06	0.06	0.06
# Obs.	1332970	1332970	1332970

Results from OLS estimation of (C.1). Clustered standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%.

Table D.6: Product Introduction and Product Rankings

	Dependent Variable: Dummy=1 for Product Introduction		
	(CES)	(DGKP)	(CN)
Lagged Rank (=1 closest)	-0.001***	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)
Product-Time FE	Yes	Yes	Yes
Firm-Time FE	Yes	Yes	Yes
Lagged Relatedness =	0	0	0
R^2	0.06	0.06	0.06
# Obs.	1173802	1173802	1173802

Results from OLS estimation of (C.1). Clustered standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%.

Table D.7: Product Introduction and Advantage

	Dependent Variable: Dummy=1 for Product Introduction			
	(CES)	(DGKP)	(CN)	(RCA)
Lagged Advantage	-6.166***	-2.510***	-2.555***	-23.188***
	(0.468)	(0.195)	(0.103)	(1.457)
Product-Time FE	Yes	Yes	Yes	Yes
Firm-Time FE	Yes	Yes	Yes	Yes
R^2	0.05	0.05	0.05	0.05
# Obs.	1332970	1332970	1332970	1332970

Results from OLS estimation of (27). Clustered standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%.

Table D.8: Product Introduction and Advantage

	Dependent Variable: Dummy=1 for Product Introduction		
	(CES)	(DGKP)	(RCA)
Lagged Advantage	-2.886***	-1.196***	-15.649***
	(0.453)	(0.192)	(1.358)
Product-Time FE	Yes	Yes	Yes
Firm-Time FE	Yes	Yes	Yes
CN Distance FE	Yes	Yes	Yes
R^2	0.06	0.06	0.06
# Obs.	1332970	1332970	1332970

Results from OLS estimation of (27). Clustered standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%.

Table D.9: Product Introduction and Advantage

	Dependent Variable: Dummy=1 for Product Introduction		
	(CES)	(DGKP)	(CN)
Lagged Advantage	-1.700***	-0.650***	-1.712***
	(0.343)	(0.150)	(0.107)
Product-Time FE	Yes	Yes	Yes
Firm-Time FE	Yes	Yes	Yes
Relatedness =	0	0	0
R^2	0.06	0.06	0.06
# Obs.	1173802	1173802	1173802

Results from OLS estimation of (27). Clustered standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%.

Table D.10: Export Shocks and Product Adoption

	Dependent Variable: Dummy=1 for Product Introduction							
	(CES)	(CES)	(DGKP)	(DGKP)	(CN)	(CN)	(RCA)	(RCA)
Lagged Advantage	-53.495	430.422	48.417*	125.458**	6.846	17.724	-112.107***	-26.161
	(114.169)	(266.203)	(29.045)	(55.870)	(8.115)	(16.099)	(6.531)	(145.616)
Lagged Advantage	-6.754	-48.850**	-5.755**	-12.347***	-1.022	-1.963	-5.474***	-173.642
X Exports	(9.801)	(23.082)	(2.455)	(4.759)	(0.687)	(1.387)	(1.741)	(284.484)
Firm-Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm-Product FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prod-Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# Obs.	16266	16266	16266	16266	16266	16266	16266	16266
F-Stat		1882.60		3624.34		2964.54		0.68

Results from OLS and IV estimation of (28). Clustered standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%. We instrument $\text{Exports}_{\nu t} \times \ln\left(1 + \tilde{a}_{\omega,\nu,t-1}\right)$ with $\text{Exports IV}_{\nu t} \times \ln\left(1 + \tilde{a}_{\omega,\nu,t-1}\right)$ where $\text{Exports IV}_{\nu t} = \text{Log} \sum_{k \neq DNK} \text{Exports}_{k\nu t}$ is the total exports of all countries except Denmark. F-Stat denotes the Kleibergen-Paap (KP) rank Wald F-statistic.

Table D.11: Export Shocks - IV = All Countries (First Stage)

	Dependent Variable: Lagged Advantage X Exports			
	(CES)	(DGKP)	(CN)	(RCA)
Lagged Advantage X Exports IV	0.788***	0.797***	0.777***	0.035***
	(0.003)	(0.003)	(0.004)	(0.002)
Firm-Time FE	Yes	Yes	Yes	Yes
Firm-Product FE	Yes	Yes	Yes	Yes
Prod-Time FE	Yes	Yes	Yes	Yes
# Obs.	16266	16266	16266	16266
R^2	1.00	1.00	1.00	0.92

Results from OLS estimation of the first stage of (28). Clustered standard errors in parenthesis. Cluster: product-year. ***: significant at 99%, ** at 95%, * at 90%. $\text{Exports IV}_{\nu t} = \text{Log} \sum_{k \neq DNK} \text{Exports}_{k\nu t}$.

E Statistical Summary of Fixed Cost Classifications

Table E.1: Fixed-cost thresholds used in the counterfactuals

	Aggregate		Simple average across clusters	
Fixed-cost threshold	Sales Share	Product Share	Sales Share	Product Share
Panel A: CES				
Minimum profit	0.213	0.488	0.181	0.482
Producer correction	0.373	0.127	0.371	0.174
Cluster correction	0.414	0.385	0.449	0.344
Panel B: DGKP				
Minimum profit	0.095	0.129	0.143	0.142
Producer correction	0.091	0.036	0.140	0.075
Cluster correction	0.814	0.835	0.718	0.783

Notes: The table reports the share of sales and the share of products for which the final fixed-cost threshold is determined by each step of the procedure. The first threshold is the minimum operating profit for each product. The second threshold raises the product fixed cost using the largest firm-level minimum operating profit among producers of the product when counterfactual participation violations remain. The third threshold applies the final cluster correction using firm-level minimum operating profits from potential entrants in the product-space cluster. Aggregate columns weight products by sales or product counts. The simple-average columns report the mean across product-space clusters.

F Advantage Normalization Validation

In Section 6.1, we recover counterfactual advantages in levels by normalizing the relative counterfactual advantage of each firm using an observed product in its portfolio. For each firm we select the product produced by the largest number of firms and choose $\widehat{\varphi}_{\omega,t}$ so that the counterfactual advantage for this reference product equals its observed advantage. We choose a widely produced product to help stabilize the normalization.

Here we provide a validation of this normalization. For firm-product observations for which advantage is observed, we compare observed advantage $a_{\omega,\nu,t}$ with the counterfactual advantage obtained using the same normalization as in Section 6.1. We estimate

$$\ln a_{\omega,\nu,t} = \beta \ln \widehat{a}^{cf}_{\omega,\nu,t} + \rho_{c,t} + \epsilon_{\omega,\nu,t}, \tag{F.1}$$

where $\rho_{c,t}$ denotes cluster-year fixed effects. A positive estimate of β implies that firm-products predicted to have higher counterfactual advantage indices also have higher observed advantage indices. Panel A of Table F.1 reports this relationship for observed firm-product pairs. Using the baseline log product space, the estimates are significant at 0.255 under CES and 0.407 under DGKP. The relationship is similar when constructing the product space using advantages in levels or the inverse hyperbolic sine transformation.

Panel B provides a validation based on product introductions. We restrict the sample to firm-products that are subsequently introduced and relate their observed advantage when produced to the counterfactual advantage predicted before introduction. Specifically, we replace the contemporaneous counterfactual advantage in Equation (F.1) with its lag. Under the baseline log product space, the estimated coefficient is 0.094 under CES and 0.084 under DGKP. Both are statistically significant at the 1% level. Similar results obtain using the alternative transformations. Thus, the normalization used to recover counterfactual advantages in levels is positively associated with observed advantage both for contemporaneously observed firm-products and for introduced products.

Table F.1: Observed and Counterfactual Advantage: Reference-Product Normalization

	Dependent Variable: Observed Absolute Advantage					
	Log		Linear		IHS	
	(CES)	(DGKP)	(CES)	(DGKP)	(CES)	(DGKP)
Panel A: Contemporaneous Counterfactual Advantage						
Counterfactual Advantage	0.255***	0.407***	0.284***	0.388***	0.278***	0.402***
	(0.009)	(0.005)	(0.008)	(0.005)	(0.009)	(0.005)
Cluster-Time FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.21	0.50	0.27	0.49	0.24	0.50
# Obs.	52995	52995	52995	52995	52995	52995
Panel B: Product Introductions						
Lagged Advantage	0.094***	0.084***	0.089***	0.065***	0.091***	0.077***
	(0.018)	(0.021)	(0.019)	(0.020)	(0.019)	(0.021)
Cluster-Time FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.08	0.07	0.08	0.07	0.08	0.07
# Obs.	3706	3706	3706	3706	3706	3706

Notes: The table reports log–log regressions of observed absolute advantage on counterfactual absolute advantage, corresponding to the normalization used in Section 6.1. Panel A uses contemporaneous advantage. Panel B restricts the sample to product introductions and uses lagged advantage. Log, Linear, and IHS denote the transformation used to construct the product space. All regressions include cluster-year fixed effects. Standard errors clustered at the firm level are reported in parentheses. *** significant at 99%, ** at 95%, * at 90%.