International Competitiveness and Migration: Diversity, Networks or Knowledge Diffusion?

Gianluca Orefice[1], Hillel Rapoport[1,2] Gianluca Santoni[1]

[1]CEPII [2]Paris School of Economics

VIII Meeting on International Economics
Univ. Jaume I - September 26^{th}-27^{th} 2019

Background & Motivation

Trade and Migration, three consolidated channels:

1. **Transaction cost**: migrants help firms in overcoming cultural barriers to trade and create international business relationships (Head & Ries, 1998; Rauch, 2001; Rauch & Trindade 2002; Parsons and Vezina 2018)

2. **Knowledge diffusion**: immigrants as vehicles of knowledge from origin to destination $\Rightarrow$ increased productivity (Bahar & Rapoport 2018; Kerr 2008)
   - Nijaz Hastor founded the Bosnias largest seat covers company after having returned from Wolfsburg in Germany (Volkswagens home town)
   - Italian immigrants producing and exporting shoes from Argentina to RoW

3. **Diversity**: migrants bring at destination a diverse set of skills, expertise and problem-solving capabilities $\Rightarrow$ multi-cultural teams perform better than homogeneous teams (Lazear 1999; Hong & Page 2001):
   - Economic growth (Ortega & Peri 2014; Alesina et al. 2016; Docquier et al. 2018)
   - Firm level productivity (Brunow et al. 2015)
   - Exports of creative goods (Orefice & Santoni 2018)

What do we do

- Main contributions

  1. Test in a *unified empirical framework* the three channels through which immigrants may affect the international competitiveness of countries: diversity, transaction costs and knowledge diffusion.
  2. Dissecting the mechanism through which diversity affects the international competitiveness of sectors (heterogeneity across sectors' characteristics)
  3. Methodological contribution: improvement on the shift share IV for immigrants (based on the recent works by Jaeger et el. 2018; Goldsmith-Pinkham et al. 2018).
  4. Byproduct: Provide a new database on revealed comparative advantages.

Structure of the talk

1. Theoretical Justification
2. Descriptive evidence
3. Empirical Strategy: three channels one specification
   - Data and identification strategy
   - Capturing the mechanism
   - Endogeneity
4. Results
   - Baseline (OLS and 2SLS)
   - Rob Check: alternative IVs
   - Extension: the role of skills
   - Rob Check: alternative diversity index (polarization)
   - Rob Check: country-sector aggregated results
5. Placebo test
6. Concluding remarks

Theoretical Justification (1)

Based on Grossman and Maggi (2000, *AER*) and Bombardini et al. (2014, *AER*):

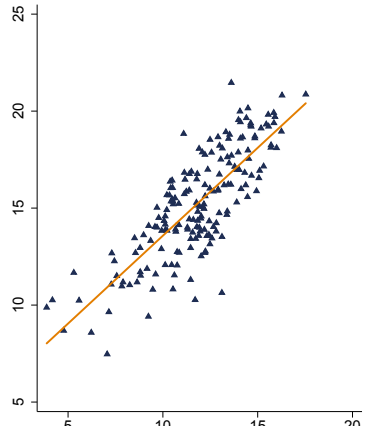
- Two countries with different workforce ability/skill distributions
- Each country produces two types of goods:
  - *Super-modular* production function good: performing better one task raises the marginal value of a better performance in other tasks (car industry) ▸ Super
  - *Sub-modular* production function goods: performing better one task mitigates the need for better performances in the other tasks. Industries requiring creativity and problem solving (fashion, design and cultural goods) ▸ Sub
- The country with a more (less) dispersed ability/skill distribution has comparative advantage in the sector characterized by sub-modular (super-modular) technology

Theoretical Justification (2)

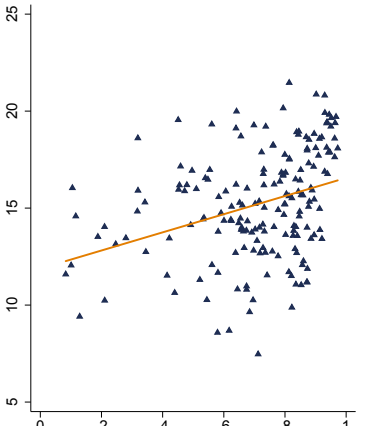
- Immigrants coming from different origins are imperfectly substitute in production (Ottaviano and Peri 2012)
- Immigrants are widely heterogeneous in their skills (positive/negative selection of migrants based on their origin composition - Borjas (1987)
- Workers from different origins are different factors of production (Ortega and Peri 2014).
- Host countries with high birthplace diversity have highly disperse distribution of worker abilities/skills (workers are *horizontally* differentiated).

$\Rightarrow$ **Conjecture**: *Birthplace Diversity is expected to improve the export performance of the host countries with a magnified effect in sectors requiring cognitive and problem solving capabilities.*

Descriptive Evidence



Beta: 0.91*** , N=181



Beta: 4.68***, N=181

Note: Scatter plot between aggregate country specific exports and immigrant stocks are reported on the left side. Birthplace diversity and country export on the right side.

Empirical Strategy

Baseline specification:

$$y_{ikjt} = \beta_1 Mig_{ijt} + \beta_2 KD_{ikt} + \beta_3 BD_{it} + \mathbf{X}_{ijkt} + \theta_{ij} + \theta_{jkt} + \theta_{rckt} + \varepsilon_{ijkt} \quad (1)$$

Dependent variables:

- Export Participation (*extensive margin*)
- Volume of Exports (*intensive margin*, Ln(export))

Main explanatory variable:

- $Mig_{i,j,t}$: stock of immigrants in destination i from origin j (transaction cost)
- $KD_{i,k,t}$: share of immigrants in i coming from origins $o \neq j$ with comparative advantage in sector k in 1995 (knowledge diffusion)
- $BD_{i,t} = 1 - \sum_j^D s_{ijt}^2$ birthplace diversity index $(1 - HH)$ (diversity) ▸ BD

Empirical Strategy

Fixed Effects

- θ_{jkt}: importing country multilateral resistance term
- θ_{rckt}: exporter country region-income-sector-year fixed effect (macro-region and income levels from World Bank classification)
- θ_{ij}: distance, contiguity, language, colonial ties, RTA

Set of controls $\mathbf{X}_{ijkt}$

- $log(1+\tau_{ijkt})$: tariff faced by country i in exporting to j in sector k
- $log(emigrants_{i->j,t})$: home bias in consumption pattern
- Importer MRT: θ_{jkt}
- Exporter MRT: θ_{rckt} ; $ln(Remote)_{it}$; export bins (ikt exports towards RoW excluding j and j's macro-region)

Estimation Sample: $\approx$ 195 countries, $\approx$ 142 SITC sectors, 5-year intervals (1995-2015)

Empirical Strategy: the mechanism

Augmented specification to test the mechanism:

$$y_{ikjt} = \beta_1 Mig_{ijt} + \beta_2 KD_{ikt} + \beta_3 (BD_{it} \times Abstract_k) + \mathbf{X}_{ijkt} + \theta_{ij} + \theta_{jkt} + \theta_{it} + \varepsilon_{ijkt} \quad (2)$$

with:

- $BD_{it} \times Abstract_k$, interaction between birthplace diversity an the job intensity in abstract tasks.
  - Abstract intensity (Autor and Dorn 2013) is a dummy variable indicating whether the sector k is intensive in complex and abstract tasks.
  - Rob checks: Team-Work (O*NET, Bombardini et al 2012), Complexity (Costinot 2009), Skills, Technology (UNCTAD), Differentiated (Rauch 1999).
- With a focus on the ikt specific interaction, country-year fixed effects θ_{it} can be included to reduce omitted variable concern.
- Mig_{ijt}, KD_{ikt} and $\mathbf{X}_{ijkt}$ as in the baseline specification

Data

- Bilateral stocks of migrants from UN, for a 195*195 matrix of origin-destination combinations, for the years 1995, 2000, 2005, 2010, 2015.
  - Rob Check dropping country pairs with imputed stocks
- Bilateral export flows from/to 195 countries from BACI (CEPII), over the period 1995-2015, aggregated at SITC level.
- Data on the presence of Preferential Trade Agreements, distance, common language, border and colony come from CEPII databases.
- Data on tariffs from WITS
- Data on GDP per capita (used to calculate the remoteness measures), income and regional classifications are from World Bank Development Indicators data.

2SLS estimation: baseline IV

IV 1: *Predicted supply-driven* stock of immigrants to remove the labor-demand component at the base of many criticisms on Bartik Instruments (shift-share IV)

- Step 1: run gravity-type equation on bilateral migration

$$M_{ijt} = \delta_{it} + \delta_{jt} + \delta_{ij}^{ij=ji} + \beta_1 \textit{Mig Sh}_{ij,60} * ln(M)_{jt} + \beta_2 \textit{Mig Sh}_{ij,60} + \mu_{ijt} \quad (3)$$

- Step 2: keep the fit and remove the demand-driven component

$$\widehat{Adj Mig}_{ijt} = \widehat{M_{ijt}} - \widehat{\delta_{it}} \quad (4)$$

- Step 3: calculate the BD index using the predicted ($\widehat{Adj Mig}_{ijt}$) rather than the observed stock of immigrants

2SLS estimation: alternative IV

The origin mix for a given destination can be invariant over time ⇒ high degree of autocorrelation in the shift-share instrument ⇒ it captures both the short- and the long-term effect of immigration at destination country (Jaeger et al. 2018).

IV 2: Modified shift-share based IV subtracting the long-term effect

- Step 1: obtain the *predicted supply-driven* stock of immigrants in each country i based on:

$$\begin{aligned} M_{ijt} &= \delta_{it} + \delta_{jt} + \delta_{ij}^{ij=ji} + \beta_1 \textit{Mig Sh}_{ij,60} * ln(M)_{jt} \\ &\quad + \beta_2 \textit{Mig Sh}_{ij,60} * ln(M)_{jt-5} + \beta_3 \textit{Mig Sh}_{ij,60} + \mu_{ijt} \end{aligned} \tag{5}$$

- Step 2: take the predicted value $\widehat{M_{ijt}}$ and subtract the destination-year fixed effect, the initial settlement of immigrants and the long-term component:

$$\widehat{Adj Mig_{ijt}^{short}} = \widehat{M_{ijt}} - \widehat{\delta_{it}} - \widehat{\beta_2} \textit{Mig Sh}_{ij,60} * ln(M)_{jt-5} - \widehat{\beta_3} \textit{Mig Sh}_{ij,60} \tag{6}$$

2SLS estimation: alternative IV

IV 3: Natural Disaster based

We use the time variation in $\widehat{AdjMig}_{ijt}$ only from countries that experienced exogenous shock (natural disasters), i.e. exogenous immigrants supply shocks.

- Step 1: take the predicted stock of immigrants $\widehat{AdjMig}_{ijt}$ from IV 1
- Step 2: use the time variation of $\widehat{AdjMig}_{ijt}$ from countries experienced natural disaster in the pre-treatment period (1985-1990), i.e $I_{j,85-90}$=1

$$\widehat{BD}_{it}^{ND} = 1 - \sum_{j=1}^{J} \left[\left(I_{j,85-90} \frac{\widehat{AdjMig}_{ijt}}{\sum_{j=1}^{J} \widehat{AdjMig}_{ijt}} \right)^2 + \left((1 - I_{j,85-90}) \frac{Mig_{ij,60}}{\sum_{j=1}^{J} Mig_{ij,60}} \right)^2 \right] \tag{7}$$

Baseline Results: Extensive Margin (OLS and 2SLS)

Dep var	$Dummy = 1$ if $Export_{ijkt} > 0$					
	OLS			**2SLS**		
	(1)	(2)	(3)	(4)	(5)	(6)
Mig_{ijt}	0.026***	0.026***	0.018***	0.023***	0.017***	0.017***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
KD_{ikt}	0.013***	0.012***	0.036***	0.013***	0.044***	0.044***
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.005)
BD_{it}	0.076***	0.076***		0.114***		
	(0.012)	(0.012)		(0.016)		
$BD_{it} \times Abstract_k$		0.014***	0.033***		0.041***	0.041***
		(0.003)	(0.004)		(0.005)	(0.010)
IV				$\widehat{IMMI}_{ijt}$	$\widehat{IMMI}_{ijt}$	$\widehat{IMMI}_{ijt}$
F-stat				1917	1378	290.2
FEs	ij, jkt, rckt	ij, jkt, rckt	ij, jkt, it	ij, jkt, rckt	ij, jkt, it	ij, jkt, it
Cluster	it, jt	it, jt	it, jt	it, jt	it, jt	i, j
Observations	20,156,093	20,156,093	20,156,093	20,156,093	20,156,093	20,156,093
R-squared	0.578	0.578	0.595			

Note: $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Exporters $i = 195$, Importers $j = 176$, Sectors (SIC72) $k = 142$, Year $t = 5$. All control variables included but not reported.

Baseline Results: Intensive Margin (OLS and 2SLS)

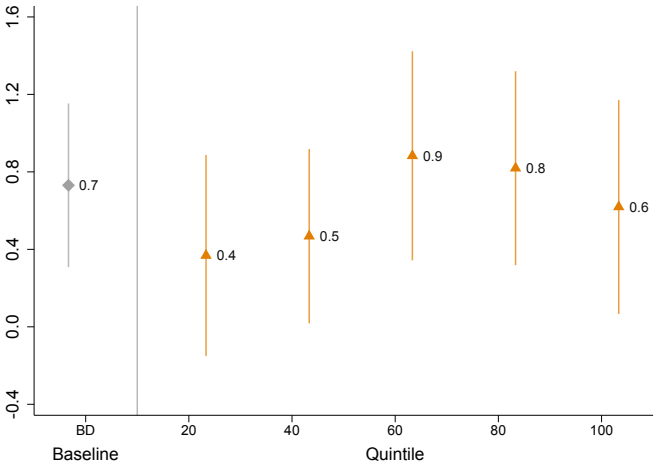
Dep var	$Ln(export_{ijkt})\|_{export(t-1)>0}$					
	OLS			**2SLS**		
	(1)	(2)	(3)	(4)	(5)	(6)
Mig_{ijt}	0.153***	0.153***	0.106***	0.191***	0.169***	0.169***
	(0.004)	(0.004)	(0.004)	(0.006)	(0.007)	(0.012)
KD_{ikt}	0.434***	0.420***	0.828***	0.538***	0.995***	0.995***
	(0.054)	(0.055)	(0.052)	(0.058)	(0.055)	(0.107)
BD_{it}	0.247**	0.247**		0.772***		
	(0.123)	(0.123)		(0.149)		
$BD_{it} \times Abstract_k$		0.168**	0.748***		0.790***	0.790***
		(0.081)	(0.080)		(0.098)	(0.192)
IV				$\widehat{IMMI}_{ijt}$	$\widehat{IMMI}_{ijt}$	$\widehat{IMMI}_{ijt}$
F-stat				386	349.5	84.7
FEs	ij, jkt, rckt	ij, jkt, rckt	ij, jkt, it	ij, jkt, rckt	ij, jkt, it	ij, jkt, it
Cluster	it, jt	it, jt	it, jt	it, jt	it, jt	i, j
Observations	4,575,395	4,575,395	4,575,395	4,575,395	4,575,395	4,575,395
R-squared	0.709	0.709	0.706			

Note: $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Exporters $i = 195$, Importers $j = 176$, Sectors (SIC72) $k = 142$, Year $t = 5$. All control variables included but not reported.

▸ First Stage

Baseline by quintile in coordination cost: Intensive Margin (2SLS)

Non-linear effect: BD_{it} coefficient by bins in the average index of language dissimilarity in i



NOTE: Quintiles based on country-year specific average index of language dissimilarity to avoid biased calculations of quintiles due to repeated observations within it (repetitions across js)

Rob Check alternative IVs

Dep var	Extensive		Intensive	
	(1)	(2)	(3)	(4)
Mig_{ijt}	0.017***	0.017***	0.169***	0.174***
	(0.002)	(0.002)	(0.012)	(0.016)
KD_{ikt}	0.044***	0.040***	0.994***	1.046***
	(0.005)	(0.008)	(0.107)	(0.132)
$\text{BD}_{it} \times Abstract_k$	0.041***	0.045***	0.791***	1.029***
	(0.010)	(0.016)	(0.192)	(0.258)
FEs	ij, jkt, it	ij, jkt, it	ij, jkt, it	ij, jkt, it
Cluster	i, j	i, j	i, j	i, j
IV	No feedback effect	Natural Disaster	No feedback effect	Natural Disaster
IV Mig_{ijt}	0.641***	0.445***	0.516***	0.321***
IV KD_{ikt}	0.934***	0.354***	0.954***	0.437***
IV $\text{BD}_{it} \times Abstract_k$	0.892***	0.588***	0.893***	0.678***
F-stat	291.2	150.8	201.9	88.1
Observations	20,156,093	20,156,093	4,575,395	4,575,395

Note: In all regressions standard errors in parentheses are double clustered at exporter and importer country level. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Exporters $i = 195$, Importers $j = 176$, Sectors (SIC72) $k = 142$, Year $t = 5$. All control variables included but not reported.

Extension by education level of migrants. OECD, DIOC-E database

Dep var	Extensive					Intensive
	(1)	(2)	(3)	(4)	(5)	(6)
Mig_{ijt}	0.032***	0.032***	0.031***	0.031***	0.031***	0.177***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.024)
KD_{ikt}	0.041***	0.039***	0.040***	0.039***	0.039***	0.275
	(0.013)	(0.012)	(0.013)	(0.013)	(0.014)	(0.192)
$BD_{it}^{Primary}$	0.047			0.078	0.078	-0.032
	(0.035)			(0.063)	(0.064)	(1.054)
$BD_{it}^{Secondary}$		0.022		-0.245*	-0.245	-2.395*
		(0.047)		(0.132)	(0.155)	(1.383)
$BD_{it}^{Tertiary}$			0.101**	0.278**	0.278**	3.885***
			(0.049)	(0.109)	(0.121)	(1.310)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
FEs	jk, rck	jk, rck	jk, rck	jk, rck	jk, rck	jk, rck
Cluster	it , jt	it , jt	it , jt	it , jt	i , j	i , j
Observations	2,001,680	2,001,680	2,001,680	2,001,680	2,001,680	1,036,479
R-squared	0.528	0.528	0.529	0.529	0.529	0.663

Note: $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Exporters $i = 114$, Importers $j = 174$, Sectors (SIC72) $k = 142$, Year $t = 2$ (i.e. 2000, 2010). All control variables included but not reported.

Country-sector aggregate regression

1. Birthplace Diversity is country-year specific
2. Birthplace Diversity is expected to have a heterogeneous effect across sectors based on problem solving capabilities needed in each sector k ($Abstract_k$)

(1)+(2)⇒ Country-sector aggregated regressions for a clear-cut identification of the effect of Birthplace Diversity on the international competitiveness of countries

Country-sector aggregate specification

$$y_{ikt} = \beta_1(BD_{i,t} \times Abstract_k) + \beta KD_{ikt} + \theta_{it} + \theta_{kt} + \epsilon_{i,k,t} \quad (8)$$

Country-sector aggregate regression

Dependent variables:

- Ex-ante measure of comparative advantage (Costinot et al. 2012) purged by transaction cost channel Comp. Adv.
- Total exports of country i in sector k time t
- Number of destinations reached by country j in a given sector-year

Explanatory variables:

- Birthplace Diversity (BD_{it}) interacted with Abstract intensity ($Abstract_k$)
- θ_{it} controls for any country-year specific factor and (partially) for the transaction cost channel
- We also include KD_{ikt} to control for the knowledge diffusion channel

Country-sector aggregate regression. Results for Abstract intensive jobs

Dep var	RCA_{ikt}			$Export_{ikt}$	$\#\ Dest_{ikt}$
	(1)	(2)	(3)	(4)	(5)
BD_{it}	0.438***	0.440***			
	(0.064)	(0.066)			
$BD_{it} * Abstract_k$		0.200***	0.208***	0.400***	0.034**
		(0.044)	(0.045)	(0.047)	(0.017)
KD_{ikt}	0.474***	0.466***	0.461***	1.168***	0.030
	(0.040)	(0.040)	(0.040)	(0.050)	(0.024)
Mig_{it}	0.038*	0.039*			
	(0.020)	(0.020)			
$FTA^{\#}_{i,t-5}$	0.006***	0.006***			
	(0.001)	(0.001)			
$ln(GDP)_{i,t-5}$	0.307***	0.307***			
	(0.023)	(0.023)			
Observations	114,262	114,262	114,262	114,262	113,745
R-squared	0.742	0.742	0.763	0.848	0.971
Country FE	Yes	Yes	No	No	No
Country-Year FE	No	No	Yes	Yes	Yes
Sector-Year FE	Yes	Yes	Yes	Yes	Yes

Note: In all regressions standard errors in parentheses are double clustered at country-sector and sector-year. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Exporters $i = 186$, Sectors (SIC72) $k = 142$, Year $t = 5$.

Placebo test on aggregate estimations

Diversity channel is expected to work only in sectors intensive in abstract NOT routine tasks

Dep var	RCA_{ikt}	$Export_{ikt}$	$\#\ Dest_{ikt}$
	(1)	(2)	(3)
$BD_{it} * Manual_k$	-0.121**	-0.003	-0.029
	(0.052)	(0.054)	(0.019)
KD_{ikt}	0.469***	1.185***	0.034
	(0.040)	(0.050)	(0.024)
Observations	114,262	114,262	113,745
R-squared	0.763	0.848	0.971
Country-Year FE	Yes	Yes	Yes
Sector-Year FE	Yes	Yes	Yes

Note: In all regressions standard errors in parentheses are double clustered at country-sector and sector-year. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Exporters $i = 186$, Sectors (SIC72) $k = 142$, Year $t = 5$.

Robustness Checks

Bilateral Regression:

- Team-Work Table
- Polarization index Table

Country-sector aggregate:

- Alternative proxies of problem solving: Table
  - Team-Work (O*NET, Bombardini et al 2012)
  - Complexity (Costinot 2009)
  - Skill intensity in production (from UNCTAD, classification SITC 3-digit)
  - Technology intensity (from UNCTAD, classification SITC 3-digit)
  - Differentiated vs homogeneous goods (Rauch classification SITC 3-digit)
- Sector Specific Regression Table

Concluding remarks

- Diversity in the birthplace of immigrants helps the international competitiveness of the country $\rightarrow$ Migration is more than a simple positive labor supply shock
- The diversity channel works on top of the transaction cost and knowledge diffusion channel
- Birthplace diversity has bigger (positive) effect on problem solving intensive sectors

Bilateral Regression

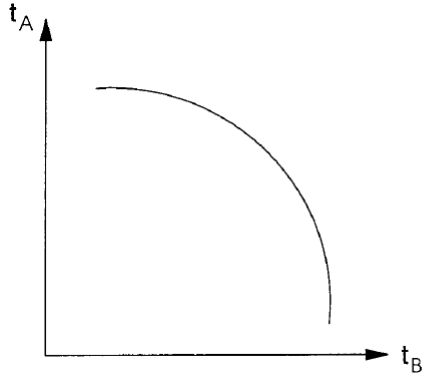
Main explanatory variable: birthplace diversity index as defined in Alesina et al. (2016) and Ortega Peri (2014), i.e. one minus the Herfindahl-Hirschman (HH) concentration index applied to the population of immigrants:

$$BD_{i,t} = 1 - \sum_{j=1}^{J} s_{ijt}^2 \tag{9}$$

- s_{ijt} is the share of immigrants originating from j in the total population of immigrants residing in country i at time t.
- BD increases with the diversity of migrants' birthplaces in the country (equal to 0 if there is only one country of origin of immigrants).

Back

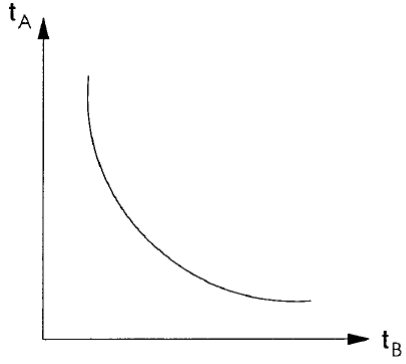
Submodularity



Back

- Increasing the talent of one worker reduces the marginal value of increasing the talent of the other $\Rightarrow$ The two tasks are *substitutable* in producing the output
- A reduction in the talent of one worker is compensated by a *decreasing* rise in the talent of the other worker
- $F(t_A, t_B) + F(t'_A, t'_B) \geq F[min(t_A, t'_A), min(t_B, t'_B)] + F[max(t_A, t'_A), max(t_B, t'_B)]$

Supermodularity



Back

- Increasing the talent of one worker increases the marginal value of increasing the talent of the other $\Rightarrow$ The two tasks are *complementary* in producing the output
- A reduction in the talent of one worker is compensated by an *increasing* rise in the talent of the other worker
- $F(t_A, t_B) + F(t'_A, t'_B) \leq F[min(t_A, t'_A), min(t_B, t'_B)] + F[max(t_A, t'_A), max(t_B, t'_B)]$

Ex-ante measure of Comparative Advantage

Step One: run a bilateral export regression using country-sector-year ($\gamma_{i,k,t}; \gamma_{j,k,t}$), country-pair ($\gamma_{i,k,j}$) fixed effects, and control for $Mig_{i,j,t}$

$$Export_{i,k,j,t} = \gamma_{i,k,t} + \gamma_{j,k,t} + \gamma_{i,k,j} + Mig_{i,j,t} + \mu_{i,j,t}, \tag{10}$$

Step Two: recover the exporter-sector-year fixed effects ($\widehat{\gamma_{i,k,t}}$) capturing all the country-sector-year specific determinants of international competitiveness (Costinot et al. 2012 ReSTUD)

Back

First Stage Results of baseline equation using IV 1

Dep var	$Dummy = 1$ if $Export_{ijkt} > 0$			Ln(export)$\|_{export(t-1)>0}$		
	Mig$_{ijt}$	KD$_{ikt}$	BD$_{it} * Abstract_k$	Mig$_{ijt}$	KD$_{ikt}$	BD$_{it} * Abstract_k$
	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{Mig_{ijt}}$	0.640***	-0.0001***	0.0000	0.514***	-0.0001*	0.0000
	(0.022)	(0.0000)	(0.0000)	(0.021)	(0.0001)	(0.0000)
$\widehat{KD_{ikt}}$	-0.017***	0.933***	0.019*	-0.081***	0.953***	0.023
	(0.004)	(0.016)	(0.011)	(0.010)	(0.014)	(0.017)
$\widehat{BD_{it}} * Abstract_k$	0.001	-0.001	0.893***	0.042***	-0.0003	0.894***
	(0.001)	(0.001)	(0.051)	(0.01)	(0.007)	(0.048)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
FEs	ij, jkt, it	ij, jkt, it	ij, jkt, it	ij, jkt, it	ij, jkt, it	ij, jkt, it
Cluster	i, j	i, j	i, j	i, j	i, j	i, j
Observations	20,156,093	20,156,093	20,156,093	4,575,395	4,575,395	4,575,395
F-test	314.2	2433.5	108.4	239.2	1811.2	207.6

Note: In all regressions standard errors in parentheses are double clustered at exporter and importer country level. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Exporters $i = 195$, Importers $j = 176$, Sectors (SIC72) $k = 142$, Year $t = 5$.

Back

Sector Specific Regression

French data. District-sector cross section approach on 1990 Census data and 2015 export data:

$$Exp_{dk} = \alpha + BD_{dk} + Ln(Mig)_{dk} + \theta_d + \theta_k + \epsilon_d$$

Dep var	$Export_{dk}$			$\#\ Dest_{dk}$	RCA_{dk}	
	(1)	(2)	(3)	(4)	(5)	(6)
BD_{dk}	1.141***	1.426***	1.555***	14.144**	1.427***	
	(0.368)	(0.382)	(0.395)	(5.534)	(0.370)	
Mig_{dk}	0.650***	0.637***	0.619***	5.746***	0.606***	
	(0.052)	(0.042)	(0.050)	(0.716)	(0.046)	
BD_d						0.634
						(1.269)
Mig_d						0.952***
						(0.073)
Observations	1,271	1,269	1,266	1,288	1,266	1,266
R-squared	0.795	0.786	0.782	0.871	0.633	0.498
Trade Year	2013	2015	2017	2017	2017	2017
Region FE	No	No	No	No	No	Yes
District FE	Yes	Yes	Yes	Yes	Yes	No
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: In all regressions standard errors in parentheses are clustered at the Region level (Region $r = 22$). $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Districts $i = 95$, Sectors $k = 15$.

Back

Rob Check: alternative measures for problem solving intensity

Dep var	RCA_{ikt} (ln)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
BD_{it}	0.438***	0.440***					
	(0.064)	(0.065)					
BD_{it}*Complexity$_k$		0.190***	0.200***				
		(0.040)	(0.040)				
BD_{it}*Team Work$_k$				0.162***			
				(0.039)			
BD_{it}*Differentiated$_k$					0.203		
					(0.143)		
BD_{it}*Skill Intensive$_k$						0.613***	
						(0.094)	
BD_{it}*High Tech$_k$							0.836***
							(0.143)
Observations	114,262	114,262	114,262	111,440	193,465	196,858	196,858
R-squared	0.742	0.742	0.763	0.766	0.857	0.607	0.607
Country FE	Yes	Yes	No	No	No	No	No
Country-Year FE	No	No	Yes	Yes	Yes	Yes	Yes
Sector-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: In all regressions standard errors in parentheses are double clustered at country-sector and sector-year. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Exporters $i = 186$, Sectors (SITC) $k = 264$, Year $t = 5$. RTA and GDP included in column (1) and (2) when country-year fixed effects are not included.

Back

Rob Check: Polarization index as in Montalvo and Reynal-Querol (2004)

Dep var	Extensive			Intensive		
	(1)	(2)	(3)	(4)	(5)	(6)
Mig_{ijt}	0.022***	0.017***	0.017***	0.193***	0.169***	0.169***
	(0.001)	(0.001)	(0.002)	(0.006)	(0.007)	(0.012)
KD_{ikt}	0.013***	0.044***	0.044***	0.540***	0.969***	0.969***
	(0.003)	(0.003)	(0.005)	(0.057)	(0.056)	(0.110)
RQ_{it}	-0.160***			-0.572***		
	(0.019)			(0.198)		
$\text{RQ}_{it} * Abstract_k$		-0.068***	-0.068***		-1.019***	-1.019***
		(0.005)	(0.010)		(0.075)	(0.154)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
FEs	ij, jkt, rckt	ij, jkt, it	ij, jkt, it	ij, jkt, rckt	ij, jkt, it	ij, jkt, it
Cluster	it, jt	it, jt	i, j	it, jt	it, jt	i, j
IV Mig_{ijt}	0.711***	0.640***	0.640***	0.589***	0.514***	0.514***
IV KD_{ikt}	0.916***	0.933***	0.933***	0.941***	0.955***	0.955***
IV RQ_{it}	0.867***			0.902***		
IV $\text{RQ}_{it} * Abstract_k$		0.893***	0.893***		0.918***	0.918***
F-stat	1555	1378	290.2	947.2	787.7	204.4
Observations	20,156,093	20,156,093	20,156,093	4,575,395	4,575,395	4,575,395

Note: In all regressions standard errors in parentheses are double clustered at exporter and importer country level. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Exporters $i = 195$, Importers $j = 176$, Sectors (SIC72) $k = 142$, Year $t = 5$. RQ_{it} is instrumented using the predicted supply from the PPML estimation using the distribution of foreign born residents in year 1960, as in the baseline specification.

Back

Rob Check: Team Work intensity as proxy for problem solving

Dep var	$Dummy = 1$ if $Export_{ijkt} > 0$			Ln(export)$\vert_{export(t-1)>0}$		
	(1)	(2)	(3)	(4)	(5)	(6)
Mig_{ijt}	0.026***	0.018***	0.017***	0.153***	0.106***	0.169***
	(0.001)	(0.001)	(0.002)	(0.004)	(0.004)	(0.012)
KD_{ikt}	0.014***	0.037***	0.044***	0.422***	0.827***	0.994***
	(0.003)	(0.002)	(0.005)	(0.055)	(0.052)	(0.108)
BD_{it}	0.077***			0.259**		
	(0.012)			(0.124)		
$BD_{it} * TeamWork_k$	0.005	0.122***	0.141***	0.979*	4.152***	4.417***
	(0.020)	(0.023)	(0.052)	(0.509)	(0.456)	(1.099)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
FEs	ij, jkt, rckt	ij, jkt, it	ij, jkt, it	ij, jkt, rckt	ij, jkt, it	ij, jkt, it
Cluster	it, jt	it, jt	i, j	it, jt	it, jt	i, j
IV Mig_{ijt}			0.640***			0.515***
IV KD_{ikt}			0.933***			0.954***
IV $RQ_{it} * TeamWork_k$			0.895***			0.898***
F-stat			290.3			204.8
Observations	19,591,612	19,591,612	19,591,612	4,501,275	4,501,275	4,501,275

Note: In all regressions standard errors in parentheses are double clustered at exporter and importer country level. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Exporters $i = 195$, Importers $j = 176$, Sectors (SIC72) $k = 138$, Year $t = 5$. BD_{it} is instrumented using the predicted supply from the PPML estimation using the distribution of foreign born residents in year 1960, as in the baseline specification.

Back