The Future of Estimating Gravity Models?

Tom Zylkin
University of Richmond

ASSA Panel on the Future of Gravity Models

January 3rd, 2025

This talk

Gravity model estimation - core method for:

- Parameterizing quantitative spatial models (for trade, migration...)
- Quantifying the economic effects of bilateral policies (trade agreements, tariffs)
- Disentangling direct partial effects from general equilibrium effects

What I'm going to talk about

- Survey results: what 90+ researchers in the field expect to change / stay the same
- Emerging developments
- Other directions that (in my opinion!) deserve more emphasis

This talk

Gravity model estimation - core method for:

- Parameterizing quantitative spatial models (for trade, migration...)
- Quantifying the economic effects of bilateral policies (trade agreements, tariffs)
- Disentangling direct partial effects from general equilibrium effects

What I'm going to talk about

- Survey results: what 90+ researchers in the field expect to change / stay the same
- Emerging developments
- Other directions that (in my opinion!) deserve more emphasis

Some themes: Heterogeneity and aggregation; disaggregated data as a frontier; how to incorporate deviations from gravity; need more convergence with panel data econometrics

Survey results: What do researchers in the field think?

Survey details

Shared via email and Twitter in mid-December. Responses were anonymous and will be shared publicly after the conference.

92 responses total:

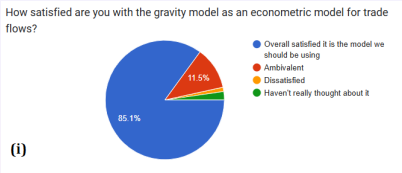
- 44.6% senior, 35.9% mid-career, 19.5% junior
- 48.9% classify themselves as "expert in the field"
- 58.2% say the estimation of gravity models plays a primary role in their work; 36.3% say secondary role
- About 78% apply the gravity model to trade primarily

Definition of "gravity model"

For the purposes of the survey, a "gravity model" is defined as an econometric model for spatial flows data with origin and destination fixed effects. That is:

$$y_{ij} = A_i \times B_j \times f(x_{ij}) \times \varepsilon_{ij} \qquad \text{with} \qquad E(\varepsilon_{ij}|A_i, B_j, x_{ij}) = 1.$$

Survey results: What do researchers in the field think?



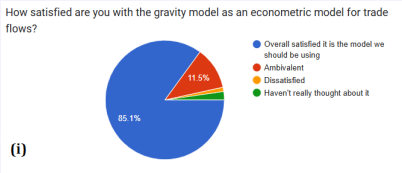
Most respondents agree that:

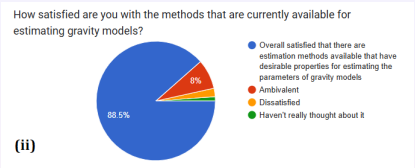
(i) gravity is the "right" econometric model for trade flows,

(ii) current methods have satisfactory properties, and

(iii) gravity estimation will continue to be common in the future.

Survey results: What do researchers in the field think?





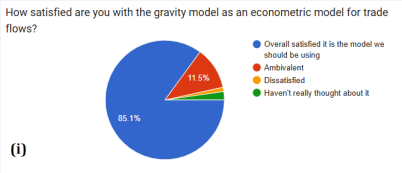
Most respondents agree that:

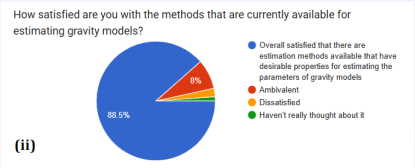
(i) gravity is the "right" econometric model for trade flows,

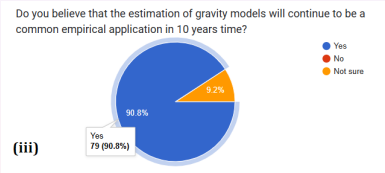
(ii) current methods have satisfactory properties, and

(iii) gravity estimation will continue to be common in the future.

Survey results: What do researchers in the field think?







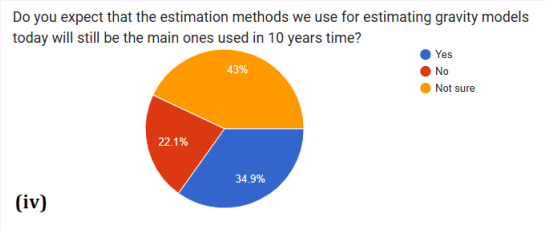
Most respondents agree that:

(i) gravity is the "right" econometric model for trade flows,

(ii) current methods have satisfactory properties, and

(iii) gravity estimation will continue to be common in the future.

Survey results: What do researchers in the field think?



However, only 35% expect current methods for estimation will continue to be widely used in 10 years time.

How might methods for estimating gravity models be different in 10 years time?

How might methods for estimating gravity models be different in 10 years time?

Some survey responses

"more advanced treatment of heterogeneous effects"

"better handling of zeroes"

"With more granular data, I expect it will go more structural. If history repeats, this will initially be complex, before someone comes along and simplifies the procedure for all."

"Availability of big (micro)data will challenge conventional approaches"

"Estimation approaches that account for staggered treatments will become more commonplace"

"More progress on product-level estimation and how to deal with heterogeneous effects"

"[A unifying] framework, which would nest conventional constant elasticity gravity and alternative formulations such as Translog gravity. More rigorous results on dynamic gravity [...] Heterogeneous elasticities estimators"

"I think that we will be spending more time on models featuring complex firms using oligopoly/oligopsony/bargaining models. In the short run, the new datasets with intranational and international sales for disaggregated products will put tariffs front and center in gravity modeling."

"using big data and machine learning techniques in this context."

"better ways of accounting for the dependency of the observations"

How might methods for estimating gravity models be different in 10 years time?

To gain some additional perspective, I want to ask the following questions:

1. Looking *back* 10 years, what has changed since 2015? (and why?)
2. What has happened in the last 5 or so years that we may not have fully digested yet?

How might methods for estimating gravity models be different in 10 years time?

Looking *back* 10 years, what has changed since 2015? (and why?)

How might methods for estimating gravity models be different in 10 years time?

Looking *back* 10 years, what has changed since 2015? (and why?)

Some key developments:

- Major contributions from the *previous* 10-15 years were thoroughly digested
  - Eaton and Kortum (2002), Anderson and van Wincoop (2003), Santos Silva and Tenreyro (2006), Baier and Bergstrand (2007), Baldwin and Taglioni (2006), ACR (2012)
- Best practices were synthesized and democratized
  - Monographs by Head and Mayer (2014) and Yotov et al. (2016), Yoto Yotov's gravity "mini-course" put everyone on the same page
  - Software packages (!) for Stata and R were developed and became widely used.
- Important technical details were worked out
  - Contributions by Paolo Guimãraes (in particular!), myself, Sergio Correia, Amrei Stammann, and Laurent Bergé made the estimation of nonlinear models with fixed effects much more feasible
  - New econometric results demonstrated validity and deepened understanding: Fernández-Val and Weidner (2016), Jochmans (2017), Weidner and Zylkin (2021), Pfaffermayer (2019, 2021), Stammann (2023)

How might methods for estimating gravity models be different in 10 years time?

Looking *back* 10 years, what has changed since 2015? (and why?)

To summarize: if the model for our outcome variable looks like either of these models

$$y_{ij} = \exp(\alpha_i + \gamma_j + x_{ij}'\beta) \times \varepsilon_{ij} \qquad \textbf{(cross-sectional gravity)}$$

$$y_{ijt} = \exp(\alpha_{it} + \gamma_{jt} + x_{ijt}'\beta) \times \varepsilon_{ijt} \qquad \textbf{(panel data gravity)}$$

we by now have a mature understanding of how to obtain estimates and interpret the results

However:

- The last few years have also shown trade differs in important ways from how these models are written
- It does not seem as though we have fully digested the implications for estimation or interpretation

How might methods for estimating gravity models be different in 10 years time?

What has happened in the last 5 or so years that we may not have fully digested yet?

How might methods for estimating gravity models be different in 10 years time?

What has happened in the last 5 or so years that we may not have fully digested yet?

Answers:

- New results for how PPML aggregates heterogeneous effects: Breinlich, Novy, and Santos Silva (2024), Tyazhelnikov, Zhou, and Shi (2024), Santos Silva and Winkelmann (2024); Nagengast and Yotov (2025)
- Why the information in disaggregated trade data matters:
  - (***biases from aggregation***) Imbs and Mejean (2015), French (2016, 2019), Redding and Weinstein (2019), Breinlich, Novy, and Santos Silva (2024), French and Zylkin (2024)
  - (***bringing comparative advantage back to the forefront***) Romalis (2004), Chor (2010), Costinot, Donaldson, and Komunjer (2012), Nunn and Trefler (2014), Levchenko and Zhang (2016), French (2017), Hanson, Lind, and Muendler (2015), Atkin, Costinot, and Fukui (2021), Shapiro (2024), Bartelme, Lan, and Levchenko (2024)
- Applications for machine learning techniques: (***trade agreement provisions***) Breinlich et al. (2023), Regmi and Baier (2023), Bergstrand and Paniagua (2024); (***flexible elasticity estimation***) Egger and Erhardt (2024)
  - **Still much under-explored here!** causal machine learning, out-of-sample validation, modeling nuisance functions...
- New data applications: migration, commuting, patents, financial flows...
- Deviations from "CES" gravity:
  - (***oligopoly***) Eaton, Kortum, and Sotelo (2013), Breinlich, Fadinger, Nocke, and Schutz (2020); Heid and Stahler (2024); Eaton, Kortum, and Kramarz (2023)
  - (***flexible demand***) Fieler (2011), Novy (2013), Carrère, Mrázová, and Neary (2020), Chen and Novy (2022)
  - (***unrestricted firm heterogeneity***) Adão, Arkolakis, and Ganapati (2024), Bas, Mayer, and Thoenig (2017)

How might methods for estimating gravity models be different in 10 years time?

What has happened in the last 5 or so years that we may not have fully digested yet?

Answers:

- New results for how PPML aggregates heterogeneous effects: Breinlich, Novy, and Santos Silva (2024), Tyazhelnikov, Zhou, and Shi (2024), Santos Silva and Winkelmann (2024); Nagengast and Yotov (2025)
- Why the information in disaggregated trade data matters:
  - (***biases from aggregation***) Imbs and Mejean (2015), French (2016, 2019), Redding and Weinstein (2019), Breinlich, Novy, and Santos Silva (2024), French and Zylkin (2024)
  - (***bringing comparative advantage back to the forefront***) Romalis (2004), Chor (2010), Costinot, Donaldson, and Komunjer (2012), Nunn and Trefler (2014), Levchenko and Zhang (2016), French (2017), Hanson, Lind, and Muendler (2015), Atkin, Costinot, and Fukui (2021), Shapiro (2024), Bartelme, Lan, and Levchenko (2024)
- Applications for machine learning techniques: (***trade agreement provisions***) Breinlich et al. (2023), Regmi and Baier (2023), Bergstrand and Paniagua (2024); (***flexible elasticity estimation***) Egger and Erhardt (2024)
  - **Still much under-explored here!** causal machine learning, out-of-sample validation, modeling nuisance functions...
- New data applications: migration, commuting, patents, financial flows...
- Deviations from "CES" gravity:
  - (***oligopoly***) Eaton, Kortum, and Sotelo (2013), Breinlich, Fadinger, Nocke, and Schutz (2020); Heid and Stahler (2024); Eaton, Kortum, and Kramarz (2023)
  - (***flexible demand***) Fieler (2011), Novy (2013), Carrère, Mrázová, and Neary (2020), Chen and Novy (2022)
  - (***unrestricted firm heterogeneity***) Adão, Arkolakis, and Ganapati (2024), Bas, Mayer, and Thoenig (2017)

How might methods for estimating gravity models be different in 10 years time?

What has happened in the last 5 or so years that we may not have fully digested yet?

Answers:

- New results for how PPML aggregates heterogeneous effects: Breinlich, Novy, and Santos Silva (2024), Tyazhelnikov, Zhou, and Shi (2024), Santos Silva and Winkelmann (2024); Nagengast and Yotov (2025)
- Why the information in disaggregated trade data matters:
  - (***biases from aggregation***) Imbs and Mejean (2015), French (2016, 2019), Redding and Weinstein (2019), Breinlich, Novy, and Santos Silva (2024), French and Zylkin (2024)
  - (***bringing comparative advantage back to the forefront***) Romalis (2004), Chor (2010), Costinot, Donaldson, and Komunjer (2012), Nunn and Trefler (2014), Levchenko and Zhang (2016), French (2017), Hanson, Lind, and Muendler (2015), Atkin, Costinot, and Fukui (2021), Shapiro (2024), Bartelme, Lan, and Levchenko (2024)
- Applications for machine learning techniques: (***trade agreement provisions***) Breinlich et al. (2023), Regmi and Baier (2023), Bergstrand and Paniagua (2024); (***flexible elasticity estimation***) Egger and Erhardt (2024)
  - **Still much under-explored here!** causal machine learning, out-of-sample validation, modeling nuisance functions...
- New data applications: migration, commuting, patents, financial flows...
- Deviations from "CES" gravity:
  - (***oligopoly***) Eaton, Kortum, and Sotelo (2013), Breinlich, Fadinger, Nocke, and Schutz (2020); Heid and Stahler (2024); Eaton, Kortum, and Kramarz (2023)
  - (***flexible demand***) Fieler (2011), Novy (2013), Carrère, Mrázová, and Neary (2020), Chen and Novy (2022)
  - (***unrestricted firm heterogeneity***) Adão, Arkolakis, and Ganapati (2024), Bas, Mayer, and Thoenig (2017)

How might methods for estimating gravity models be different in 10 years time?

What has happened in the last 5 or so years that we may not have fully digested yet?

Answers:

- New results for how PPML aggregates heterogeneous effects: Breinlich, Novy, and Santos Silva (2024), Tyazhelnikov, Zhou, and Shi (2024), Santos Silva and Winkelmann (2024); Nagengast and Yotov (2025)
- Why the information in disaggregated trade data matters:
  - (***biases from aggregation***) Imbs and Mejean (2015), French (2016, 2019), Redding and Weinstein (2019), Breinlich, Novy, and Santos Silva (2024), French and Zylkin (2024)
  - (***bringing comparative advantage back to the forefront***) Romalis (2004), Chor (2010), Costinot, Donaldson, and Komunjer (2012), Nunn and Trefler (2014), Levchenko and Zhang (2016), French (2017), Hanson, Lind, and Muendler (2015), Atkin, Costinot, and Fukui (2021), Shapiro (2024), Bartelme, Lan, and Levchenko (2024)
- Applications for machine learning techniques: (***trade agreement provisions***) Breinlich et al. (2023), Regmi and Baier (2023), Bergstrand and Paniagua (2024); (***flexible elasticity estimation***) Egger and Erhardt (2024)
  - **Still much under-explored here!** causal machine learning, out-of-sample validation, modeling nuisance functions...
- New data applications: migration, commuting, patents, financial flows...
- Deviations from "CES" gravity:
  - (***oligopoly***) Eaton, Kortum, and Sotelo (2013), Breinlich, Fadinger, Nocke, and Schutz (2020); Heid and Stahler (2024); Eaton, Kortum, and Kramarz (2023)
  - (***flexible demand***) Fieler (2011), Novy (2013), Carrère, Mrázová, and Neary (2020), Chen and Novy (2022)
  - (***unrestricted firm heterogeneity***) Adão, Arkolakis, and Ganapati (2024), Bas, Mayer, and Thoenig (2017)

How might methods for estimating gravity models be different in 10 years time?

What has happened in the last 5 or so years that we may not have fully digested yet?

Answers:

- New results for how PPML aggregates heterogeneous effects: Breinlich, Novy, and Santos Silva (2024), Tyazhelnikov, Zhou, and Shi (2024), Santos Silva and Winkelmann (2024); Nagengast and Yotov (2025)
- Why the information in disaggregated trade data matters:
  - (***biases from aggregation***) Imbs and Mejean (2015), French (2016, 2019), Redding and Weinstein (2019), Breinlich, Novy, and Santos Silva (2024), French and Zylkin (2024)
  - (***bringing comparative advantage back to the forefront***) Romalis (2004), Chor (2010), Costinot, Donaldson, and Komunjer (2012), Nunn and Trefler (2014), Levchenko and Zhang (2016), French (2017), Hanson, Lind, and Muendler (2015), Atkin, Costinot, and Fukui (2021), Shapiro (2024), Bartelme, Lan, and Levchenko (2024)
- Applications for machine learning techniques: (***trade agreement provisions***) Breinlich et al. (2023), Regmi and Baier (2023), Bergstrand and Paniagua (2024); (***flexible elasticity estimation***) Egger and Erhardt (2024)
  - **Still much under-explored here!** causal machine learning, out-of-sample validation, modeling nuisance functions...
- New data applications: migration, commuting, patents, financial flows...
- Deviations from "CES" gravity:
  - (***oligopoly***) Eaton, Kortum, and Sotelo (2013), Breinlich, Fadinger, Nocke, and Schutz (2020); Heid and Stahler (2024); Eaton, Kortum, and Kramarz (2023)
  - (***flexible demand***) Fieler (2011), Novy (2013), Carrère, Mrázová, and Neary (2020), Chen and Novy (2022)
  - (***unrestricted firm heterogeneity***) Adão, Arkolakis, and Ganapati (2024), Bas, Mayer, and Thoenig (2017)

How might methods for estimating gravity models be different in 10 years time?

What has happened in the last 5 or so years that we may not have fully digested yet?

Answers:

- New results for how PPML aggregates heterogeneous effects: Breinlich, Novy, and Santos Silva (2024), Tyazhelnikov, Zhou, and Shi (2024), Santos Silva and Winkelmann (2024); Nagengast and Yotov (2025)
- Why the information in disaggregated trade data matters:
  - (***biases from aggregation***) Imbs and Mejean (2015), French (2016, 2019), Redding and Weinstein (2019), Breinlich, Novy, and Santos Silva (2024), French and Zylkin (2024)
  - (***bringing comparative advantage back to the forefront***) Romalis (2004), Chor (2010), Costinot, Donaldson, and Komunjer (2012), Nunn and Trefler (2014), Levchenko and Zhang (2016), French (2017), Hanson, Lind, and Muendler (2015), Atkin, Costinot, and Fukui (2021), Shapiro (2024), Bartelme, Lan, and Levchenko (2024)
- Applications for machine learning techniques: (***trade agreement provisions***) Breinlich et al (2023), Regmi and Baier (2023), Bergstrand and Paniagua (2024); (***flexible elasticity estimation***) Egger and Erhardt (2024)
  - **Still much under-explored here!** causal machine learning, out-of-sample validation, modeling nuisance functions...
- New data applications: migration, commuting, patents, financial flows...
- Deviations from "CES" gravity:
  - (***oligopoly***) Eaton, Kortum, and Sotelo (2013), Breinlich, Fadinger, Nocke, and Schutz (2020); Heid and Stahler (2024); Eaton, Kortum, and Kramarz (2023)
  - (***flexible demand***) Fieler (2011), Novy (2013), Carrère, Mrázová, and Neary (2020), Chen and Novy (2022)
  - (***unrestricted firm heterogeneity***) Adão, Arkolakis, and Ganapati (2024), Bas, Mayer, and Thoenig (2017)

Theme #1: Weighting and heterogeneity

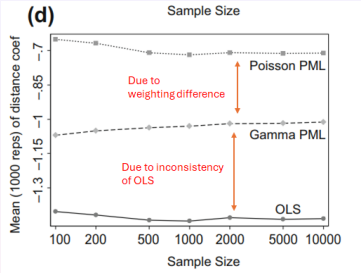
- As noted, methods for estimating gravity models with homogeneous coefficients are by now relatively mature
- But what if the true coefficients exhibit unmodeled heterogeneity (as they almost certainly do)?
- Gravity model with heterogeneous coefficients:

$$y_{ij} = e^{\alpha_i + \gamma_j + x_{ij}\beta_{ij}} \times \varepsilon_{ij}$$

What do we get if we estimate a single coefficient for β using PPML (or some other estimator)?

- how are different observations weighted?
- how does this affect interpretation in general?

Theme #1: Weighting and heterogeneity



From simulations with both heterogeneous distance elasticities and heteroskedasticity

"*More generally, major divergence in large samples between Poisson and Gamma PML—as exhibited in Figure 3.3 (c) and (d)—can signal model mis-specification.*"

from Head and Mayer (2014), Section 5.1.

- Interestingly, Head and Mayer already had weighting and heterogeneity in mind in 2014!
  - Heuristically, PPML places a higher weight on larger observations than Gamma PML
  - Not much attention given to what the "weights" are exactly until recently

What does PPML estimate when coefficients are heterogeneous?

The results of Breinlich, Novy, and Santos Silva (2024) and Tyazhelnikov, Zhou, and Shi (2024) both imply that:

If $E(y_{ij}|z_{ij}) = \mu_{ij} = e^{z'_{ij}\theta_{ij}}$*, then the uniform PPML coefficient estimate* $\widehat{\theta}$ *estimates*

$$\theta^* = \left[\sum_{ij} H_{ij}^{(m)}\right]^{-1} \sum_{ij} H_{ij}^{(m)} \theta_{ij},$$

where

- $H_{ij}^{(m)} = \exp(z'_{ij}\theta_{ij}^{(m)}) z_{ij} z'_{ij}$ is the PPML Hessian evaluated at $\theta_{ij}^{(m)}$
- $\theta_{ij}^{(m)}$ is an intermediate point between θ_{ij} and θ^*
- Higher weighting on larger trade flows comes from the dependence of the Hessian on the predicted trade value $\exp(z'_{ij}\theta_{ij}^{(m)})$.

Theme #1: Weighting and heterogeneity

- Breinlich, Novy, and Santos Silva (2024) and Tyazhelnikov, Zhou, and Shi (2024): PPML places a higher weight on larger trade flows through its Hessian
- Why does this matter?

Theme #1: Weighting and heterogeneity

- Breinlich, Novy, and Santos Silva (2024) and Tyazhelnikov, Zhou, and Shi (2024): PPML places a higher weight on larger trade flows through its Hessian
- Why does this matter?

  A few reasons:

  1. Many practitioners either do not think about this issue or are uncertain about how to think about it
  2. Matters for interpretation
  3. There are lots of ways to aggregate heterogeneity, so is this the one we want?
  4. Suggests we may want to compare PPML estimates with those of other estimators (Gamma PML, Multinomial PML), as Head and Mayer originally suggested.
  5. Weights can be negative, as in Staggered Diff-in-Diff settings
     - de Chaisemartin and D'Haultfœuille, Callaway and Sant'Anna, Nagengast and Yotov
  6. When we know heterogeneity is present, we should think seriously about how to model it
     - Baier, Bergstrand, and Clance 2018, Baier, Yotov, and Zylkin 2019, Chen and Novy 2022, Egger and Erhardt 2024

An illustration of how PPML weights larger trade flows more

Data: CEPII gravity database for year 2015 (Trade data is from IMF DOTS)

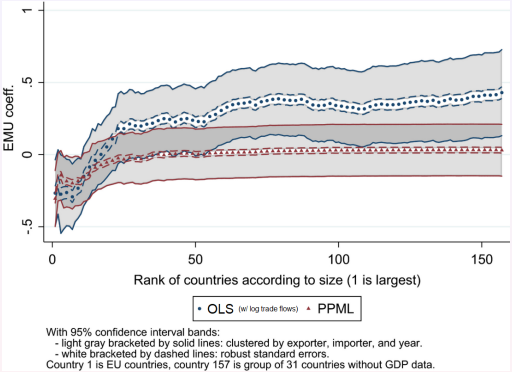
	Dependent variable: Total Bilateral Imports (from IMF DOTS)					
	(1)	(2)	(3)	(4)	(5)	(6)
Log Distance	-0.816***	-0.817***	-0.816***	-0.812***	-0.788***	-0.768***
	(0.030)	(0.030)	(0.030)	(0.029)	(0.030)	(0.031)
Historical colonial relationship	0.304	0.300	0.291	0.258	0.176	0.074
	(0.192)	(0.192)	(0.190)	(0.185)	(0.176)	(0.175)
Common official language	0.021	0.019	0.014	0.012	0.013	0.033
	(0.064)	(0.064)	(0.064)	(0.064)	(0.065)	(0.070)
Common legal origin	0.128***	0.128***	0.132***	0.130***	0.130***	0.108**
	(0.040)	(0.040)	(0.040)	(0.040)	(0.041)	(0.044)
FTA	0.441***	0.441***	0.439***	0.422***	0.371***	0.332***
	(0.056)	(0.056)	(0.056)	(0.055)	(0.056)	(0.058)
Sample	Full	No zeroes	Largest 50% of trade flows	Largest 25%	Largest 10%	Largest 5%
Observations	33,486	26,943	17,858	9,174	3,692	1,838

Aggregate bilateral trade flows for year 2015. All estimations use PPML with exporter and importer fixed effects. Standard errors, which appear in parentheses, are clustered by pair. * $p < 0.10$, ** $p < .05$, *** $p < .01$.

Due to the skewness of trade data, 75% of all trade flows receive small weight and play almost no role in the result

- Is this consistent with what we think we are estimating?

Another example: Did the European Monetary Union increase trade?



From Larch, Wanner, Yotov, and Zylkin (2019)

- Shows how adding more non-EMU countries to the sample affects the estimated EMU effect
- Here, it's probably **good** that PPML downweights all the smaller trade partners

Theme #2: Using the information in disaggregated trade data

What has happened in the last 5 or so years that we may not have fully digested yet?

Answers:

- New results for how PPML aggregates heterogeneous effects: Breinlich, Novy, and Santos Silva (2024), Tyazhelnikov, Zhou, and Shi (2024), Santos Silva and Winkelmann (2024); Nagengast and Yotov (2025)
- Why the information in disaggregated trade data matters:
  - (***biases from aggregation***) Imbs and Mejean (2015), French (2016, 2019), Redding and Weinstein (2019), Breinlich, Novy, and Santos Silva (2024), French and Zylkin (2024)
  - (***bringing comparative advantage back to the forefront***) Romalis (2004), Chor (2010), Costinot, Donaldson, and Komunjer (2012), Nunn and Trefler (2014), Levchenko and Zhang (2016), French (2017), Hanson, Lind, and Muendler (2015), Atkin, Costinot, and Fukui (2021), Shapiro (2024), Bartelme, Lan, and Levchenko (2024)
- Applications for machine learning techniques: (***trade agreement provisions***) Breinlich et al (2023), Regmi and Baier (2023), Bergstrand and Paniagua (2024); (***flexible elasticity estimation***) Egger and Erhardt (2024)
  - **Still much under-explored here!** causal machine learning, out-of-sample validation, modeling nuisance functions...
- New data applications: migration, commuting, patents, financial flows...
- Deviations from "CES" gravity:
  - (***oligopoly***) Eaton, Kortum, and Sotelo (2012), Breinlich, Fadinger, Nocke, and Schutz (2020); Heid and Stahler (2024); Eaton, Kortum, and Kramarz (2023)
  - (***flexible demand***) Fieler (2011), Novy (2013), Carrère, Mrázová, and Neary (2020), Chen and Novy (2022)
  - (***unrestricted firm heterogeneity***) Adão, Arkolakis, and Ganapati (2024), Bas, Mayer, and Thoenig (2017)

Theme #2: Using the information in disaggregated trade data

Another context where PPML's weighting is useful is weighting different products or sectors.

Consider the following pooled, product-level gravity model:

$$y_{ijkt} = \exp\left(\delta_{ikt} + \gamma_{jkt} + \eta_{ijk} + x_{ijt}\beta_k\right) \times \varepsilon_{ijkt}.$$

Theme #2: Using the information in disaggregated trade data

Another context where PPML's weighting is useful is weighting different products or sectors.

Consider the following pooled, product-level gravity model:

$$y_{ijkt} = \exp\left(\delta_{ikt} + \gamma_{jkt} + \eta_{ijk} + x_{ijt}\beta_k\right) \times \varepsilon_{ijkt}.$$

When we estimate this model with PPML, we get the following:

- Aggregate-level estimates are a **special case** where all coefficients are homogeneous across products
  - i.e. let $\delta_{ikt} = \delta_{it}, \gamma_{jkt} = \gamma_{jt}$, $\eta_{ijk} = \gamma_{ij}$, $\beta_k = \beta$
  - true for PPML and **only** PPML (Amhrein and Flowerdew 1992, French 2019)
- This means we can use pooled PPML to relax model restrictions that we implicitly make when we use aggregate trade data (French and Zylkin 2024.)
- It also means we should be cautious how to interpret pooled estimates using other estimators
  - could be driven by products with small amounts of trade

Theme #2: Using the information in disaggregated trade data

What happens at the disaggregated product level is important for aggregate estimates!

From French and Zylkin (2024):

	Using *aggregate* trade flows	Using (SITC 5 digit) *product-level* data		
	(1)	(2)	(3)	(4)
FTA	0.113***	0.113***	0.054***	-0.104***
	(0.034)	(0.034)	(0.020)	(0.023)
FTA × Least Traded Product				0.741***
				(0.038)
it and *jt* FEs	x	x		
ij FEs	x	x		
ikt and *jkt* FEs			x	x
ijk FEs			x	x
Observations	60,614	42,721,982	39,663,541	39,663,541

PPML estimates for pooled sample of 5 digit SITC3 trade flows between 109 countries over the period 1991-2015, every 4 years. Aggregate trade data is the sum of trade across products. Standard errors are clustered by pair. * $p < 0.10$, ** $p < .05$, *** $p < .01$.

- Column 3 is telling us a significant part of the 0.113 FTA estimate in column (1) is **not** due to changes in bilateral frictions but rather due to changes in product-country-time factors
- Product-level comparative advantage plays a much more important role in driving trade than aggregate-level gravity estimates imply.

  (also see French 2016, 2019.)

Theme #2: Using the information in disaggregated trade data

Challenges and opportunities with disaggregated gravity

- More computational advances needed? adding a sectoral dimension means bigger data sets, implies many more fixed effects
- Data quality: Oberhofer and Wang (2023) raise concerns about standard data sources. Discrepancies at the disaggregated level could be large:
  - World Trade Flows data and UN COMTRADE give very different gravity estimates
  - Different versions of BACI also give different estimates
  - Lukaszuk and Torun (2022) propose a way to "harmonize the harmonized system"
- Firm-level gravity: How to incorporate the additional detail from firm-level data within the general equilibrium structure implied by gravity models?
  - Aytun, Hinz, Ozguzel, and Wanner (2024) propose combining firm-level data with product-level trade data
  - Bas, Mayer, and Thoenig (2017) use firm-level exports from multiple countries to obtain "firm-level gravity"
- What level of disaggregation matters? French and Zylkin (2024) found similar results for 2 digit SITC data (63 industries) as for 5 digit SITC (2,771 industries)
- Unpacking comparative advantage: What you export matters!
  - Hausmann, Hwang, and Rodrik (2007), Blanchard and Olney (2017), Bartelme, Lan, and Levchenko (2024)
  - What explains the large fluctuations in comparative advantage over time documented in Hanson, Lind, and Muendler (2015)?

Theme #2: Using the information in disaggregated trade data

Challenges and opportunities with disaggregated gravity (cont'd)

- Tariffs will become more focal:
  - (***tariff elasticity estimates at disaggregated level***) Fontagne, Guimbard, and Orefice (2022)
  - (***improved tariff data will challenge existing results***) Teti (2024), Caliendo, Feenstra, Romalis, and Taylor (2023)
- How does market structure at the micro-level matter for aggregation? Some products may be better described by a gravity model than others.

Theme #3: Expanding the toolkit

What has happened in the last 5 or so years that we may not have fully digested yet?

Answers:

- New results for how PPML aggregates heterogeneous effects: Breinlich, Novy, and Santos Silva (2024), Tyazhelnikov, Zhou, and Shi (2024), Santos Silva and Winkelmann (2024); Nagengast and Yotov (2025)
- Why the information in disaggregated trade data matters:
  - (***biases from aggregation***) Imbs and Mejean (2015), French (2016, 2019), Redding and Weinstein (2019), Breinlich, Novy, and Santos Silva (2022), French and Zylkin (2024)
  - (***bringing comparative advantage back to the forefront***) Romalis (2004), Chor (2010), Costinot, Donaldson, and Komunjer (2012), Nunn and Trefler (2014), Levchenko and Zhang (2016), French (2017), Hanson, Lind, and Muendler (2015), Atkin, Costinot, and Fukui (2021), Shapiro (2024), Bartelme, Lan, and Levchenko (2024)
- Applications for machine learning techniques: (***trade agreement provisions***) Breinlich et al (2023), Regmi and Baier (2023), Bergstrand and Paniagua (2024); (***flexible elasticity estimation***) Egger and Erhardt (2024)
  - **Still much under-explored here!** causal machine learning, out-of-sample validation, modeling nuisance functions...
- New data applications: migration, commuting, patents, financial flows...
- Deviations from "CES" gravity:
  - (***oligopoly***) Eaton, Kortum, and Sotelo (2012), Breinlich, Fadinger, Nocke, and Schutz (2020); Heid and Stahler (2024); Eaton, Kortum, and Kramarz (2023)
  - (***flexible demand***) Fieler (2011), Novy (2013), Carrère, Mrázová, and Neary (2020), Chen and Novy (2022)
  - (***unrestricted firm heterogeneity***) Adão, Arkolakis, and Ganapati (2024), Bas, Mayer, and Thoenig (2017)

Theme #3: Expanding the toolkit (machine learning)

Some possible applications for machine learning techniques

Out-of-sample validation

- Setting a baseline model for performing counterfactuals (Dingel and Tintelnot 2023)
- Testing how well deviations from "CES" gravity improve predictions
- Determining how much heterogeneity to allow for in the coefficients (to avoid overfitting)
- Assessing data quality of competing data sources (firm-level data vs. regular trade data, BACI vs. COMTRADE)
- Determining what level of disaggregation to use

Building richer econometric models

- Dealing with high-dimensional covariates (Breinlich et al. 2023, Regmi and Baier 2023, Bergstand and Paniagua 2024)
- Allowing for arbitrary nonlinearity:
  - in trade cost elasticities (Egger and Erhardt 2024)
  - as a time-varying confounder (Baier and Standaert (2024))
- Causal machine learning: Synthetic controls, matrix completion, ...
- Using LLMs to create text-based indicators of trade sentiment (Caldara and Iacoviello 2022)

Theme #3: Expanding the toolkit (econometrics in general)

Many other modern econometric approaches have not been fully absorbed by the gravity literature

These considerations are often straightforward to address using linear panel data models, less so using nonlinear models with gravity fixed effects:

- "Random growth" models with unit-specific time trends (Baier, Bergstrand, and Feng 2014)
- Dynamic panel data models that allow for trade persistence
- Implications of staggered treatment + heterogeneous treatment effects (Nagengast and Yotov 2025)
- Spatial dynamics across markets (Morales, Sheu, and Zahler 2019)
- Synthetic control-based inference methods
- Instrumental variables (Jochmans and Verardi 2022)

Summarizing

- In the past 10 years, we have become very proficient at estimating the gravity model as it is typically written
- Nonetheless, there is reason to believe the estimates we get from current approaches may paint an incomplete picture of what we really want to know.
- I hope (and expect) that future work will shed more light on the following issues:
  - How to model heterogeneity
  - Weighting of heterogeneous coefficients when there is unmodeled heterogeneity
  - The information contained in disaggregated trade data
  - How much deviations from "CES" matter quantitatively (and when)
  - The role that multinationals play
  - Dynamics

These were not the only issues raised in the survey results; will share the full results soon!

Summarizing

Thank you!

References I

Adão, R., C. Arkolakis, and S. Ganapati (2024): "From Heterogeneous Firms to Heterogeneous Trade Elasticities: The Aggregate Implications of Firm Export Decisions," Working paper.

Anderson, J. E. and E. van Wincoop (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93, 170–192.

Arkolakis, C., A. Costinot, and A. Rodríguez-Clare (2012): "New Trade Models, Same Old Gains?" *The American Economic Review*, 102, 94–130.

Atkin, D., A. Costinot, and M. Fukui (2021): "Globalization and the Ladder of Development: Pushed to the Top or Held at the Bottom?" Working paper, National Bureau of Economic Research.

Aytun, U., J. Hinz, C. Ozgüzel, and J. Wanner (2024): "Firm-level Gravity with Fixed Effects," Working paper.

Baier, S. and S. Standaert (2024): "Gravity, globalization and time-varying heterogeneity," *European Economic Review*, 163, 104679.

Baier, S. L. and J. H. Bergstrand (2007): "Do Free Trade Agreements actually Increase Members' International Trade?" *Journal of International Economics*, 71, 72–95.

Baier, S. L., J. H. Bergstrand, and M. W. Clance (2018): "Heterogeneous effects of economic integration agreements," *Journal of Development Economics*, 135, 587–608.

References II

Baier, S. L., J. H. Bergstrand, and M. Feng (2014): "Economic integration agreements and the margins of international trade," *Journal of International Economics*, 93, 339–350.

Baier, S. L. and N. R. Regmi (2023): "Using machine learning to capture heterogeneity in trade agreements," *Open Economies Review*, 34, 863–894.

Baier, S. L., Y. V. Yotov, and T. Zylkin (2019): "On the widely differing effects of free trade agreements: Lessons from twenty years of trade integration," *Journal of International Economics*, 116, 206–226.

Baldwin, R. and D. Taglioni (2006): "Gravity for Dummies and Dummies for Gravity Equations," Working Paper 12516, National Bureau of Economic Research.

Bartelme, D., T. Lan, and A. A. Levchenko (2024): "Specialization, market access and real income," *Journal of International Economics*, 150, 103923.

Bas, M., T. Mayer, and M. Thoenig (2017): "From micro to macro: Demand, supply, and heterogeneity in the trade elasticity," *Journal of International Economics*, 108, 1–19.

Bergstrand, J. H. and J. Paniagua (2024): "Do Deep Trade Agreements' Provisions Actually Increase–or Decrease–Trade and/or FDI?" .

References III

Blanchard, E. J. and W. W. Olney (2017): "Globalization and human capital investment: Export composition drives educational attainment," *Journal of International Economics*, 106, 165–183.

Breinlich, H., V. Corradi, N. Rocha, M. Ruta, J. Santos Silva, and T. Zylkin (2022): "Machine learning in international trade research-evaluating the impact of trade agreements," CEPR discussion paper no. DP17325.

Breinlich, H., H. Fadinger, V. Nocke, and N. Schutz (2020): "Gravity with granularity," CEPR discussion paper no. DP15374.

Breinlich, H., D. Novy, and J. Santos Silva (2024): "Trade, gravity and aggregation," *Review of Economics and Statistics*, 1–29.

Caldara, D. and M. Iacoviello (2022): "Measuring geopolitical risk," *American Economic Review*, 112, 1194–1225.

Caliendo, L., R. C. Feenstra, J. Romalis, and A. M. Taylor (2023): "Tariff Reductions, Heterogeneous Firms, and Welfare: Theory and Evidence for 1990–2010," *IMF Economic Review*, 71, 817–851.

Carrère, C., M. Mrázová, and J. P. Neary (2020): "Gravity without apology: the science of elasticities, distance and trade," *The Economic Journal*, 130, 880–910.

References IV

Chen, N. and D. Novy (2022): "Gravity and heterogeneous trade cost elasticities," *The Economic Journal*, 132, 1349–1377.

Chor, D. (2010): "Unpacking sources of comparative advantage: A quantitative approach," *Journal of International Economics*, 82, 152–167.

Costinot, A., D. Donaldson, and I. Komunjer (2012): "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," *Review of Economic Studies*, 79, 581–608.

Dingel, J. and F. Tintelnot (2023): "Spatial Economics for Granular Settings," Working paper.

Eaton, J. and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779.

Eaton, J., S. Kortum, and F. Kramarz (2023): "Firm-to-Firm Trade: Imports, Exports, and the Labor Market," .

Eaton, J., S. Kortum, and S. Sotelo (2013): "International Trade: Linking Micro and Macro," in *Advances in Economics and Econometrics: Tenth World Congress*, Cambridge University Press, vol. 2, 329.

Egger, P. H. and K. Erhardt (2024): "Heterogeneous effects of tariff and nontariff trade-policy barriers in quantitative general equilibrium," *Quantitative Economics*, 15, 453–487.

Fernández-Val, I. and M. Weidner (2016): "Individual and time effects in nonlinear panel models with large N, T," *Journal of Econometrics*, 192, 291–312.

References V

Fieler, A. C. (2011): "Nonhomotheticity and Bilateral Trade: Evidence and a Quantitative Explanation," *Econometrica*, 79, 1069–1101.

Fontagné, L., H. Guimbard, and G. Orefice (2022): "Tariff-based product-level trade elasticities," *Journal of International Economics*, 137, 103593.

French, S. (2016): "The composition of trade flows and the aggregate effects of trade barriers," *Journal of International Economics*, 98, 114–137.

——— (2017): "Revealed comparative advantage: What is it good for?" *Journal of International Economics*, 106, 83–103.

——— (2019): "Comparative advantage and biased gravity," UNSW Business School Research Paper 2017-03.

French, S. and T. Zylkin (2024): "The effects of free trade agreements on product-level trade," *European Economic Review*, 104673.

Hanson, G. H., N. Lind, and M.-A. Muendler (2015): "The Dynamics of Comparative Advantage," Working Paper 21753, National Bureau of Economic Research.

Hausmann, R., J. Hwang, and D. Rodrik (2007): "What you export matters," *Journal of economic growth*, 12, 1–25.

References VI

Head, K. and T. Mayer (2014): "Gravity equations: workhorse, toolkit, and cookbook," *Handbook of International Economics*, 4, 131–196.

Heid, B. and F. Stähler (2024): "Structural gravity and the gains from trade under imperfect competition: Quantifying the effects of the European Single Market," *Economic Modelling*, 131, 106604.

Imbs, J. and I. Mejean (2015): "Elasticity optimism," *American Economic Journal: Macroeconomics*, 7, 43–83.

Jochmans, K. (2017): "Two-way models for gravity," *Review of Economics and Statistics*, 99, 478–485.

Jochmans, K. and V. Verardi (2022): "Instrumental-variable estimation of exponential-regression models with two-way fixed effects with an application to gravity equations," *Journal of Applied Econometrics*, 37, 1121–1137.

Larch, M., J. Wanner, Y. V. Yotov, and T. Zylkin (2019): "Currency unions and trade: a PPML re-assessment with high-dimensional fixed effects," *Oxford Bulletin of Economics and Statistics*, 81, 487–510.

Levchenko, A. A. and J. Zhang (2016): "The evolution of comparative advantage: Measurement and welfare implications," *Journal of Monetary Economics*, 78, 96–111.

Lukaszuk, P. and D. Torun (2022): "Harmonizing the Harmonized System," SSRN working paper 4302540.

References VII

Morales, E., G. Sheu, and A. Zahler (2019): "Extended gravity," *The Review of economic studies*, 86, 2668–2712.

Nagengast, A. J. and Y. V. Yotov (2025): "Staggered difference-in-differences in gravity settings: Revisiting the effects of trade agreements," *American Economic Journal: Applied Economics*, 17, 271–296.

Novy, D. (2013): "Gravity Redux: Measuring International Trade Costs with Panel Data," *Economic Inquiry*, 51, 101–121.

Nunn, N. and D. Trefler (2014): "Domestic institutions as a source of comparative advantage," *Handbook of international economics*, 4, 263–315.

Oberhofer, H. and Z. Wang (2023): "Everything You Always Wanted to Know About EU Membership Trade Effects But Were Afraid to Ask," .

Pfaffermayr, M. (2019): "Gravity models, PPML estimation and the bias of the robust standard errors," *Applied Economics Letters*, 1–5.

——— (2021): "Confidence intervals for the trade cost parameters of cross-section gravity models," *Economics Letters*, 201, 109787.

References VIII

Redding, S. J. and D. E. Weinstein (2019): "Aggregation and the gravity equation," in *AEA Papers and Proceedings*, American Economic Association, vol. 109, 450–455.

Romalis, J. (2004): "Factor proportions and the structure of commodity trade," *American Economic Review*, 94, 67–97.

Santos Silva, J. and R. Winkelmann (2024): "Misspecified exponential regressions: Estimation, interpretation, and average marginal effects," *Review of Economics and Statistics*, 1–25.

Santos Silva, J. M. C. and S. Tenreyro (2006): "The log of gravity," *Review of Economics and Statistics*, 88, 641–658.

Shapiro, J. S. (2023): "Institutions, Comparative Advantage, and the Environment," Working paper, National Bureau of Economic Research.

Stammann, A. (2023): "Debiased Fixed Effects Estimation of Binary Logit Models with Three-Dimensional Panel Data," *arXiv preprint arXiv:2311.04073*.

Teti, F. (2024): "Missing Tariffs," Working paper.

Tyazhelnikov, V., X. Zhou, and X. Shi (2024): "PPML, Gravity, and Heterogeneous Trade Elasticities," Working paper.

References IX

Weidner, M. and T. Zylkin (2021): "Bias and consistency in three-way gravity models," *Journal of International Economics*, 132, 103513.

Yotov, Y. V., R. Piermartini, J.-A. Monteiro, and M. Larch (2016): "An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model," *World Trade Organization, Geneva*.