Pattern Making and Pattern Breaking: Measuring Novelty in Brazilian Economics

Marcos Paulo Correia

Department of Economics

University of Brasília

and

Bernardo Mueller

Department of Economics

University of Brasília

Motivation

How do new ideas, novelty and impact arise and spread?

Novelty and impact is usually measured through number of publications or citations.

We use information theory: topic modelling and Kullback-Leibler dvergence.

Context: Annual meeting of the Brazilian National Association of Post-graduate studies – ANPEC.

1600+ papers from 2013 to 2019.

ANPEC – Brazilian Economic Association

Founded in 1973.

Members: 28 departments of economics.

Yearly meeting in early December each year.

Highly competitive selection.

Most in Portuguese, but many in English.

Divided in 13 fields of Economics.

Brazilian Economics: well-connected to the global frontier, yet basically in a follower status.

Information theory and the measurement of novelty

Topic Modelling – Latent Dirichlet Allocations.

Bag of words.

Each paper is described as a probability distribution over 30 topics.

Kullback-Leibler divergence to measure patterns in how different probability distributions diverge from each other.

Novelty: When new patterns emerge.

Transience: When patterns that existed are no longer repeated (novelty from the future to the present).

Resonance: High novelty with low transience.

Individuals, institutions, and innovation in the debates of the French Revolution

Alexander T. J. Barron[a], Jenny Huang[a,b], Rebecca L. Spang[c], and Simon DeDeo[b,d,1]

[a]School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN 47408; [b]Santa Fe Institute, Santa Fe, NM 87501; [c]Department of History, College of Arts and Sciences, Indiana University, Bloomington, IN 47405; and [d]Department of Social and Decision Sciences, Dietrich College, Carnegie Mellon University, Pittsburgh, PA 15213

The French Revolution brought principles of "liberty, equality, fraternity" to bear on the day-to-day challenges of governing what was then the largest country in Europe. Its experiments provided a model for future revolutions and democracies across the globe, but this first modern revolution had no model to follow. Using reconstructed transcripts of debates held in the Revolution's first parliament, we present a quantitative analysis of how this body managed innovation. We use information theory to track the creation, transmission, and destruction of word-use patterns across over 40,000 speeches and a thousand speakers. The parliament as a whole was biased toward the adoption of new patterns, but speakers' individual qualities could break these overall trends. Speakers on the left innovated at higher rates, while speakers on the right acted to preserve prior patterns. Key players such as Robespierre (on the left) and Abbé Maury (on the right) played information-processing roles emblematic of their politics. Newly created organizational functions—such as the Assembly president and committee chairs—had significant effects on debate outcomes, and a distinct transition appears midway through the parliament when committees, external to the debate process, gained new powers to "propose and dispose." Taken together, these quantitative results align with existing qualitative interpretations, but also reveal crucial information-processing dynamics that have hitherto been overlooked. Great orators had the public's attention, but deputies (mostly on the political left) who mastered the committee system gained new powers to shape revolutionary legislation.

parliament, the body itself had little precedent to follow. Its members therefore faced a double challenge: how to convey points in a way familiar enough to be intelligible by others, while nonetheless making claims that were in many cases substantially novel ("revolutionary," even). The NCA was a site, therefore, of both epistemic and political innovation. Conceiving it as such suggests two sets of questions. First, how did new ideas enter that parliament room; how were they adopted, adapted, or discarded by the men who heard them? Second: What institutions did the parliament evolve to manage the onslaught of novelty and reaction, optimism and grievance, philosophical argument and organizational minutiae that characterized the day-to-day tasks of governance and nation-building?

The digitization of historical archives allows us to answer these questions in a fundamentally new way. Using latent Dirichlet allocation (2) and new techniques in information theory drawn from the cognitive sciences (3), we track the emergence and persistence of word-use patterns in over 40,000 speeches made in the NCA and later reconstructed in the *Archives Parliamentaires* (AP) from detailed records kept at the time. Two critical measures—novelty (how unexpected a speech's patterns are, given past speeches) and transience (the extent to which those patterns fade or persist in future speeches)—allow us to trace both new manners of speech and the emergence of new institutions. Our mapping of the French Revolution's turbulent early days in terms of the creation, sharing, and destruction of word-use patterns complements existing studies of specific ideas

Information theory

Entropy

$$H(X) = -\sum_{i=1}^{n} p(x_i) log_2 p(x_i)$$

Kullback-Leibler divergence

$$KL(p \mid q) = \sum_{i=1}^{n} p(x_i) log_2 \frac{p(x_i)}{q(x_i)}$$

Information theory

Novelty (Bayesian surprise)

$$\mathcal{N}_w(j) = \frac{1}{w} \sum_{d=1}^{w} KL(s^j \mid s^{j-d})$$

Transience (Bayesian surprise) from future to past.

$$\mathcal{T}_z(j) = \frac{1}{w} \sum_{d=1}^{z} KL(s^j \mid s^{j+d})$$

Resonance (new pattern that persists)

$$\mathcal{R}(j) = \mathcal{N}_w(j) - \mathcal{T}_z(j)$$

Literature (Kullback-Leibler)

Barron et al. (2017): French Revolution.

Murdock et al. (2017): Darwin's notebooks.

Jing et al. (2019): Fanfiction.

Degaetano-Ortlieb & Teich (2018): Papers of the Royal Society of London.

Literature (other methods)

Uzzi et al. 2013: Web of Science.

Askin & Mauskapf, 2017: Pop music.

Foster et al., 2015; papers in chemistry and biomedicine.

Uzzi and Spiro (2005): Broadway musicals.

Mueller, 2021: Pop music and music genres.

Youn et al., 2015: Patents.

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
Trade	Innovation	Worker	Space	Shock
Export	Technological	Salary	City	Inflation
Industry	Firm	Women	Regional	Regime
Commercial	Growth	Occupation	Municipality	Cycle
Import	Technology	Man	Urban	Forecasting
Industrial	Patent	Age	Distance	Expectation
China	Industrial	Income	Center	Monetary
Exporter	Interaction	Time	Industry	IPCA
Input	Industry	Working	Worker	Hiatus
'Trade'	Innovative	Home	Density	Curve
Household	Intensity	Wages	Mobility	Structural
Export	Network	Differential	Concentration	Matrix
Sectoral	Innovation	White	Agglomeration	Breakage
Matrix	Effort	Boss	Location	Phillips
Worldwide	Innovative	Benefit	Transport	Meta
Area 7	Area 9	Area 13	Area 10	Area 4
International	Industrial	Labor	Regional &	Macro., Mon.
Economics	and Tecnology	Economics	Urban	& Finance

Topic 6	Topic 11	Topic 12	Topic 13	Topic 14
Currency	Emissions	Science	Marxs	Inequality
Active	Energy	Economist	Capitalist	Poverty
Crisis	Scenario	Institution	Money	Regional
Risk	Family	Knowledge	Merchandise	Poor
Monetary	Environmental	Concept	Profit	Northeast
Portfolio	Ethanol	Scientific	Capitalist	Regional
Return	Simulation	Veblen	Accumulation	Decomposition
Liquidity	Energy	Human	Strength	Rural
Flow	Transport	World	Category	Education
Global	Shock	Vision	Wealth	Income
Action	Fuel	Thought	Class	Southeast
Interest	Balance	Hayek	Expansion	North
Credit	Oil	Critical	Trading	Family
Title	Climate	Practice	World	Urban
Exchange	Input	Action	Infrastructure	Gini
Area 4	Area 11	Area 1	Area 2	Area 12
Macro., Mon. & Finance	Agricultural & Environment	Hist Thought Methodology	Political Economy	Social & Demog. Econ.

Selected Topic: 0 | Previous Topic | Next Topic | Clear Topic

Intertopic Distance Map (via multidimensional scaling)

Economic history

History of economic thought - methodology

Trade

Micro:
Labor, social, demographic, urban, etc.
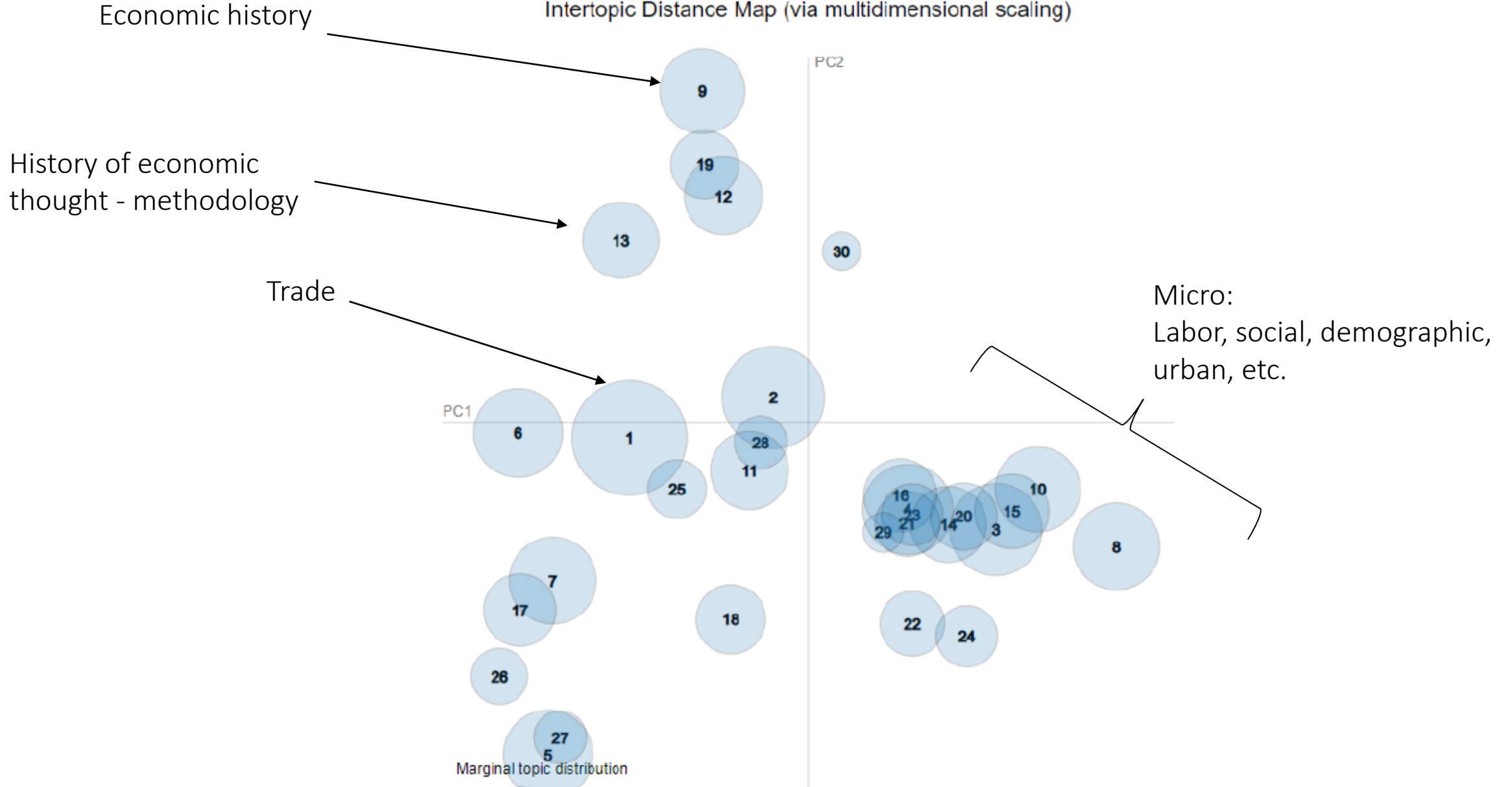
PC2

PC1

Marginal topic distribution

Figure 4: Novelty vs. Transience at ANPEC meetings
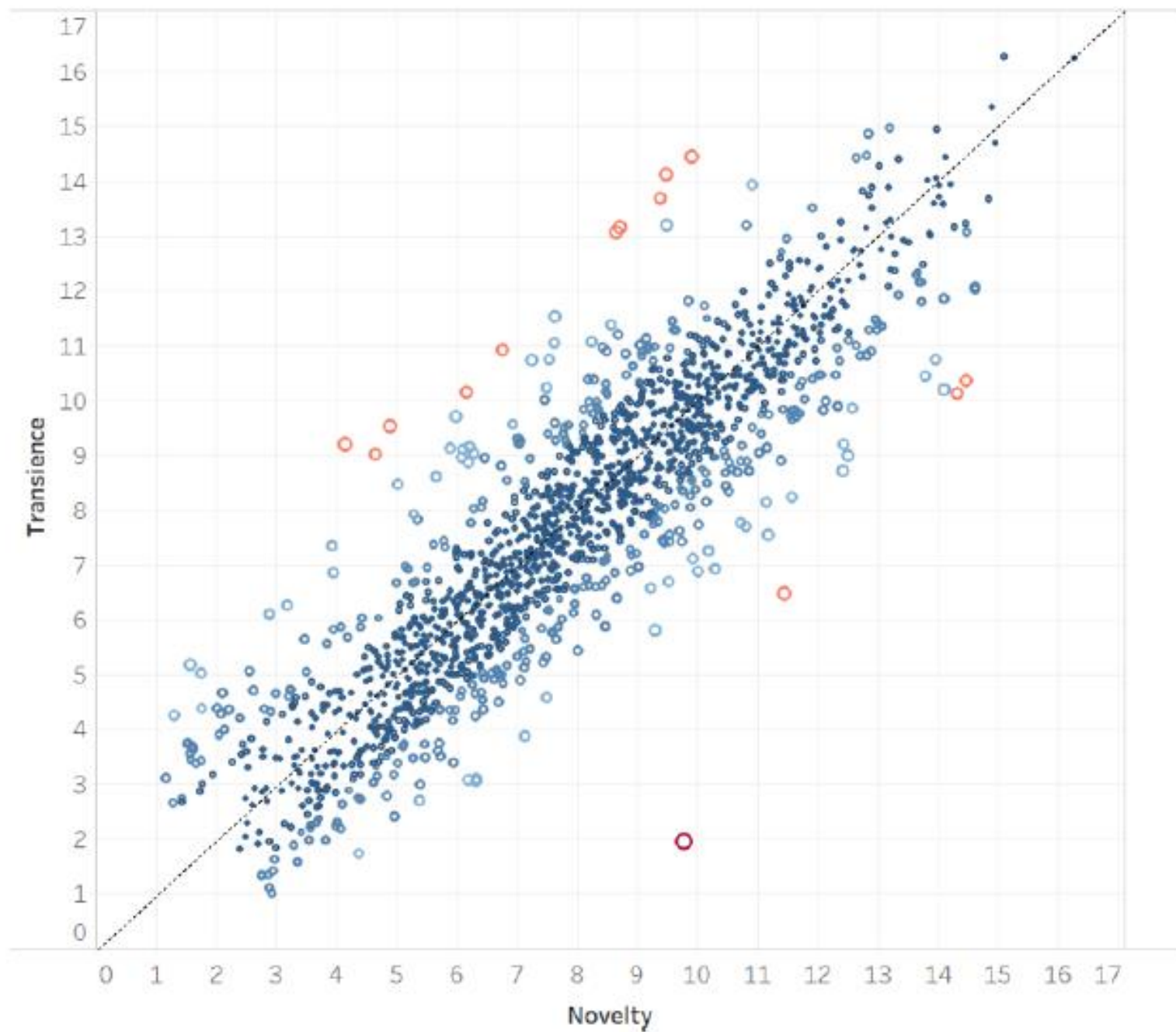
Figure 5: Novelty vs. Resonance at ANPEC meetings
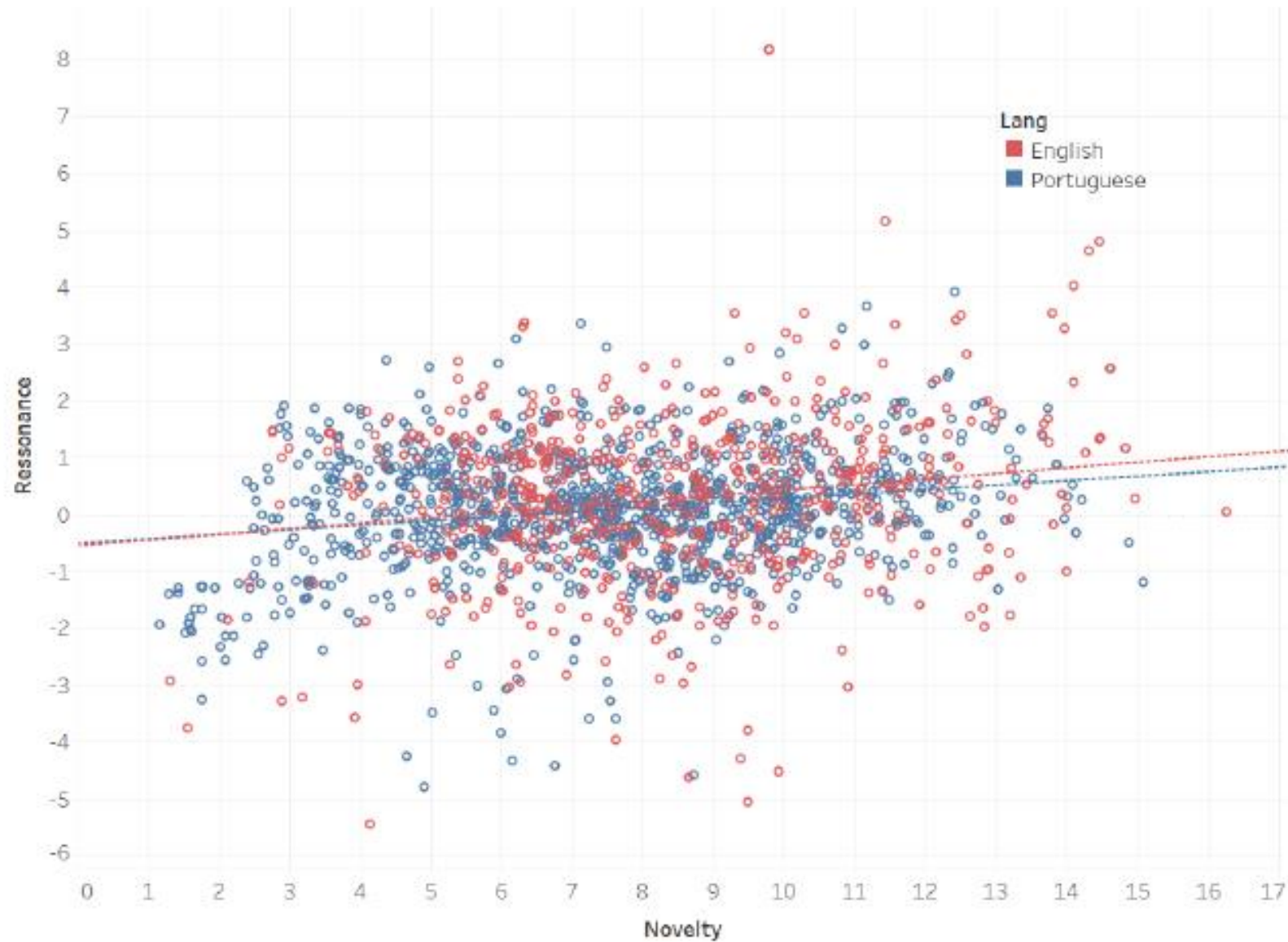
Source: Authors' elaboration using data from ANPEC http://www.anpec.org.br/ and code from Barron et al. (2018).
The slope for English is 0.98 and for Portuguese 0.78, both statistically significant at 1%

Dep. variable	I Citations	II Citations	III Citations
Novelty	-0.011		
	(-0.20)		
Transience		−0.114*	
		(1.76)	
Resonance			0.421*
			(1.73)
Brazilian journal	4.496***	4.485***	4.558***
	(8.02)	(7.99)	(8.13)
Foreign journal	8.954***	8.920***	8.949***
	(5.21)	(5.24)	(5.26)
English	0.576**	0.652**	0.509**
	(2.23)	(2.42)	(2.03)
Number of authors	0.294*	0.295*	0307*
	(1.84)	(1.85)	(1.95)
2014	-0.836	-0.846	-0.928
	(-0.90)	(-0.91)	(-1.02)
2015	−2.234***	−2.248***	−2.361***
	(-3.01)	(-3.03)	(-3.10)
2016	−3.666***	−3.647***	−3.836***
	(-5.31)	(-5.35)	(-5.27)
2017	−4.296***	−4.314***	−4.578***
	(-6.15)	(-6.20)	(-5.92)
2018	−4.501***	−4.544***	−4.759***
	(-7.20)	(-7.22)	(-6.92)
2019	−4.836***	−4.934***	−5.134***
	(-7.85)	(-7.84)	(-7.39)
Constant	3.149***	3.818***	3.122***
	(3.27)	(3.74)	(3.39)
Observations	1,676	1,675	1,675
R-squared	0.23	0.23	0.23

Notes: Ordinary least squares, robust errors. * 10%, ** 5%, *** 1%. Base year is 2013. Foreign and Brazilian publications compared to papers only published in the ANPEC annals.

Figure 6: Economics departments by novelty and resonance

Notes: Departments set according to the first author. Only departments with more that 10 papers at the meetings from 2013 to 2019 were included.

Figure 7: ANPEC areas by novelty and resonance
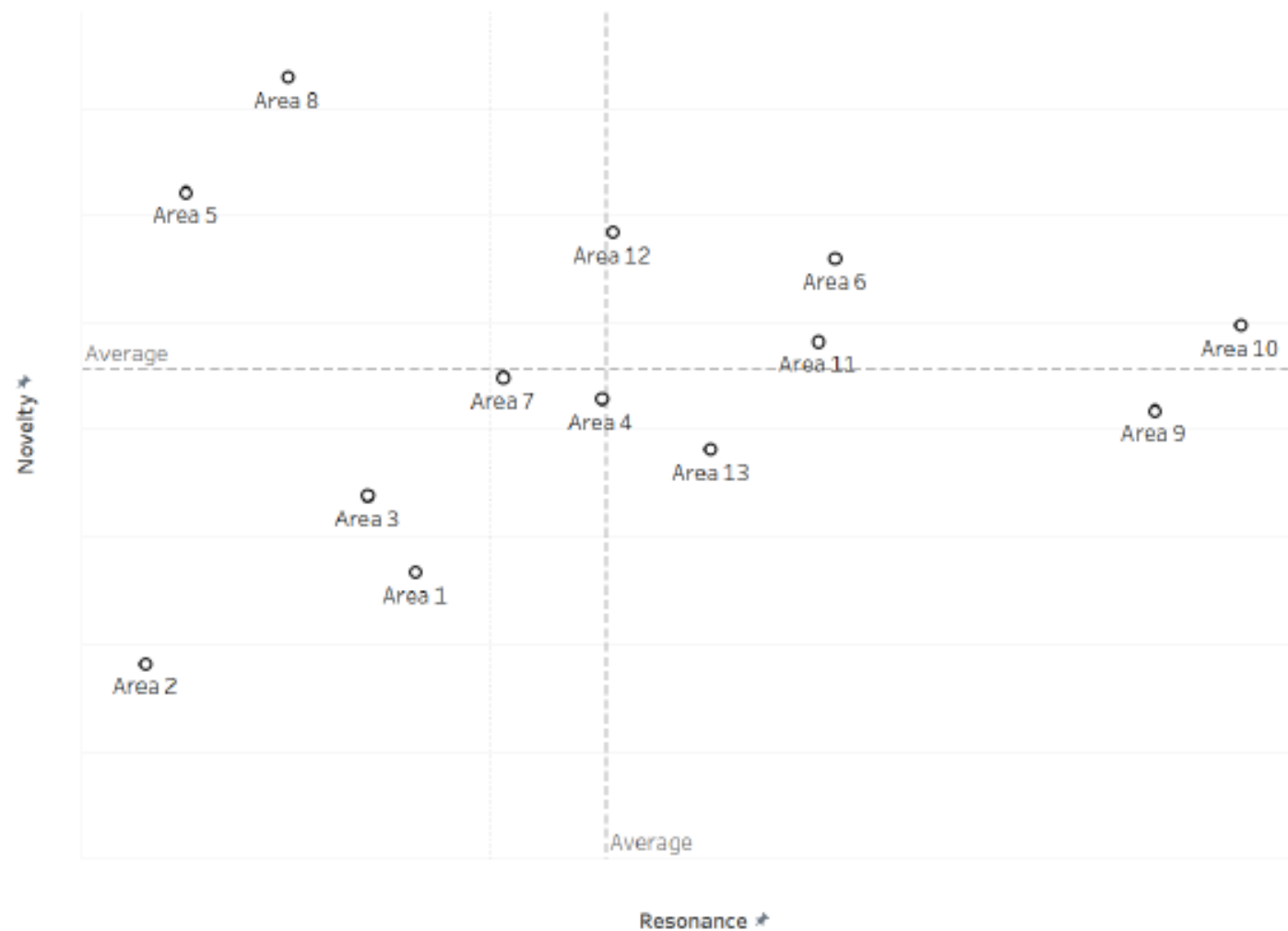
Notes: Area 1 - History of economic thought; Area 2 - Political economy; Area 3 - Economic history; Area 4 - Macro, monetary, finance; Area 5 - Public sector econ.; Area 6 - Growth, develop., institutions; Area 7 - International econ.; Area 8 - Micro. Quant. methods, finance; Area 9 - Industrial econ. & technology; Area 10 - Regional and urban; Area 11 - Agricultural & Environmental; Area 12 - Social & demographic; Area 13 - Labor economics.

Conclusions

Novelty is necessary but not sufficient for impact.

Science is conservative. New ideas face reactionary forces.

Information theory-based methods align partially with traditional citation-based methods to measure impact.