Lena Abou El-Komboz and *Moritz Goldbeck*

International Collaboration in Digital Knowledge Work: A Production-Side Assessment of Europe's Digital Single Market*

Europe's digital economy is lagging behind global competition, especially the US, despite comparable market size and human capital availability. The largest US tech companies – the "magnificent seven," i.e., Alphabet (Google), Amazon, Apple, Meta, Microsoft, Nvidia, and Tesla – generate around ten times more revenue than the EU's largest seven (Foucart 2024). At the same time, the digital economy is becoming increasingly indispensable for productivity and economic growth. Fostering market integration to further deepen the digital single market is crucial to advancing Europe's competitiveness as an attractive location for firms pursuing digital business models.

Digital business models typically require high upfront fixed-cost investments in product development, large parts of which are personnel costs for knowledge workers, before scaling on low (sometimes even close to zero) marginal cost becomes possible. With such production technologies, a market environment that allows firms to efficiently organize and coordinate available human capital is paramount. Existing evidence emphasizes the crucial role of collaboration in knowledge worker teams for productivity (Jones 2009; Wuchty 2007). With a geographically disproportionately distributed workforce in Europe (Wachs et al. 2022), increasing international collaboration is an important lever to facilitate digital business.

However, border effects (i.e., reductions of economic exchange across the EU's many national borders) could constitute a significant barrier to international collaboration. Border effects are one of the most robust and consistent empirical findings in international economics (McCallum 1995; Anderson and van Wincoop 2003). Yet, in the digital economy, many of the traditional explanations for border effects do not apply, such as transportation costs (Blum and Goldfarb 2006). Still, other frictions to economic exchange along intra-European national borders, e.g., cultural or language differences, are potentially significant barriers to international collaboration.

Hence, a production-side assessment of digital markets with a focus on collaboration is crucial to inform digital single market policymaking. In this article, which is based on Abou El-Komboz and Goldbeck (2024b), we contribute to this question by asking

KEY MESSAGES

- Low market integration is a major barrier to European competitiveness in the digital economy

- International collaboration of knowledge workers could help to harness market size advantage

- Data from the largest coding platform reveals digital collaboration in Europe drops by 16.4 percent at borders

- This border effect is 73 percent larger than between US states

- Organizations and cultural proximity facilitate international collaboration

- European digital single market policymaking should address production-side barriers to support the digital economy

if there is a border effect in virtual collaboration of knowledge workers and explore potential remaining frictions along national borders within Europe.[1]

DATA

We investigate software developers as a prime example of highly digitalized knowledge work with the potential for fully virtual collaboration (Emanuel et al. 2023). To measure interregional collaboration, we tap data from *GitHub*, by far the largest online code repository platform for collaborative software development. We extract the activity of around 144,000 European and 191,000 US software developers in 10,735,071 public repositories (projects) from *GHTorrent*, a project that mirrors the content available through the official GitHub REST API and structures it in a relational database (Gousios 2013). Users' self-reported location allows us to elicit spatial collaboration patterns via a geocoding procedure. We measure interregional collaboration by the number of user pairs with joint projects in each NUTS2 region pair of 34 countries from 2015 to

---

[1] In this article, we use Europe and the EU interchangeably, as our results apply mostly to EU policymaking, but our data extends to non-EU member states in Europe, e. g., Switzerland.
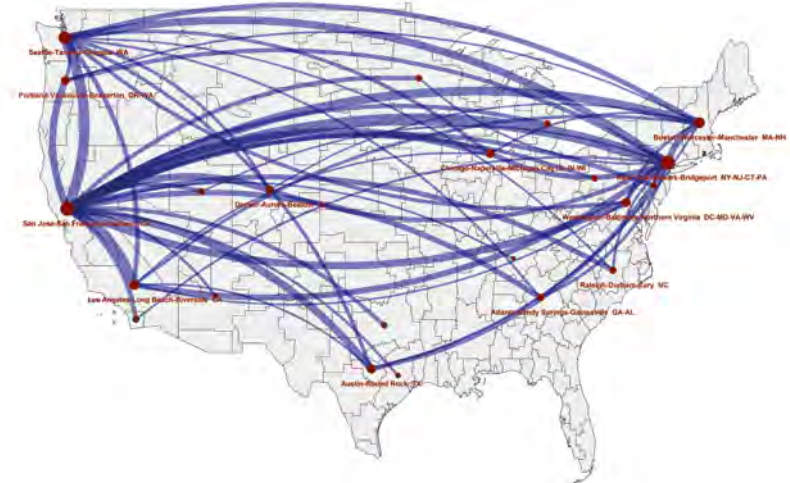
* Corresponding author: goldbeck@ifo.de.

Figure 1

Inter-regional Software Developer Collaboration, 2015–2021

Panel A: Europe

Panel B: United States

Note: Panels A and B map the structure of the inter-regional software developer collaboration network for Europe and the US, respectively. Regions in Europe are NUTS2 and BEA Economic Areas in the United States. Important edges of the network, defined as inter-regional links above 25,000 connections, are shown in blue and their width is scaled logarithmically. Regions are shown in gray. Bold (thin) black lines represent national (regional) borders. Regions with important edges feature their centroids as nodes in red, scaled by their logarithmic sum of inter-regional connections. Ireland, Hawaii, and Alaska are not shown.
Source: GHTorrent, Bureau of Economic Analysis, World Cities Database, authors' compilation.

© ifo Institute

2021. To this end, we define joint projects of two users as repositories that receive at least one commit (code contribution) by each of the users in the observation period. Figure 1 maps the most important nodes and edges of the interregional collaboration network and contrasts the pattern in Europe and the US.

We complement the data for Europe by combining it with data on cultural proximity and other cross-country differences. Importantly, we tap the measure of cultural proximity of Obradovich et al. (2022), who use data on online behavior from the Facebook marketing API to compute a bottom-up, data-driven, and granular assessment of interest overlap between populations. The authors show that this measure aligns with traditional measures of cultural differences while improving on granularity and representativeness. We further add data on genetic distance from Spolaore and Wacziarg (2009), a well-established proxy for cultural factors associated with ethnicity, as well as common language, shared colonial or same-country history, and religious distance from the *CEPII Gravity Database* (Conte et al. 2022). Note that differing availability of these metrics for European countries impacts the sample.

EMPIRICAL APPROACH

We estimate a parsimonious gravity equation following the canonical model by McCallum (1995) to identify border effects:

$$\ln(y_{i,j}) = \beta_0 + \beta_1 \text{crossborder}_{i,j} + \beta_2 \text{coloc}_{i,j} + \beta_3 \ln(\text{dist}_{i,j}) + \delta_i + \delta_j + \varepsilon_{i,j}$$

where $y_{i,j}$ represents the number of bilateral collaborations between regions i and j, including domestic collaborations $i=j$. The dummy variable $\text{crossborder}_{i,j}$ indicates if region i is located in a different country than region j, and $\text{dist}_{i,j}$ denotes the geographic distance between the regions. We further add a colocation indicator, $\text{coloc}_{i,j}$, to account for strong colocation effects in collaboration (Goldbeck 2023). Origin and destination fixed effect δ_i and δ_j account for unobserved regional determinants of collaboration common across all partner regions. The coefficient β_2 captures the elasticity of collaboration with respect to geographic distance, which we expect to be negative from theory. The border effect is given by our coefficient of interest β_1.

DIGITAL BORDER EFFECT

Table 1 reports our main regression results. The first two columns present baseline estimates of the border effect at European and US state borders, respectively. The border effect in Europe is −16.4 percent, i. e., software developers collaborate, on average, 16.4 percent less with developers located in other countries as compared to national developers. Importantly, this is after controlling for geographic distance, collaboration potential, and other unobserved regional factors through origin and destination fixed effects. Comparing this result to the border effect found in similar models for international trade (Havranek and Irsova 2017; Santamaría et al. 2023a and 2023b), we find that the digital border effect is about five to six times smaller. This reflects the generally lower barriers for interregional collaboration in the digital economy compared to brick-and-mortar industries.

In contrast to the EU, the border effect in the US, which naturally features higher market integration and lower cultural and language barriers, is only

−9.5 percent. This is 6.9 percentage points or 73 percent lower compared to the border effect in Europe, a sizable difference. Notably, colocation in same region is much less relevant for collaboration in Europe compared to the US while geographic distance has a stronger effect. The more equal spatial distribution of Europe's population is likely the main reason for this result. Additionally, most interregional collaboration in the US is happening between the large cities on the east and west coasts, resulting in a higher share of long-distance relative to short-distance collaboration.

PRODUCTION-SIDE BARRIERS TO COLLABORATION

To investigate what explains the higher border effect in Europe, we account for factors commonly associated with being a barrier to economic exchange in columns three and four of Table 1. In column three, we add cross-country cultural and ethnic differences. Cultural differences are strongly significantly and consistently related to collaboration, while ethnic differences are economically insignificant but still statistically significantly related to international collaboration. In column four, we include further cross-country differences related to specific potential barriers. Religious differences and colonial history are not associated statistically significantly with collaboration. In contrast, a common language is related to 8.2 percent higher collaboration, significant at the five percent level. A shared history as same country is negatively related to collaboration, reflecting disrupted relations due to the history of the former Yugoslavia and Austria-Hungary. Importantly, once these factors are included in the model, the estimate of the border effect becomes statistically insignificant with point estimates close to zero. This implies that the digital border effect in Europe can be entirely explained by cultural proximity and language barriers.

A decomposition of the measure of cultural proximity by Obradovich et al. (2022) into subindices of interest overlap yields the largest association for the category "non-local business." This result suggests organizations or at least shared professional interests enable international collaboration. This is in line with evidence in literature suggesting that organizations often facilitate exchange and are therefore well-suited to overcoming external barriers to collaboration. Also using the data from GitHub, Goldbeck (2023) shows, for example, that developers from the same, large firm engage disproportionately in remote collaboration.

Our results in Abou El-Komboz and Goldbeck (2024b) further demonstrate that the border effect is systematically related to the number of countrywide users. We show the border effect roughly doubles when a small country is involved, defined as hosting an below-median

Table 1

Digital Border Effect

Collaboration	EU	USA	EU	EU
	(1)	(2)	(3)	(4)
Cross-border	−0.180***	−0.100***	−0.009	−0.014
	(0.014)	(0.033)	(0.037)	(0.037)
Colocation	0.862***	2.191***	1.485***	1.476***
	(0.068)	(0.073)	(0.069)	(0.070)
Geographic distance [log]	−0.129***	−0.060***	−0.016**	−0.018**
	(0.007)	(0.011)	(0.008)	(0.008)
Cultural distance			−0.097***	−0.081***
			(0.016)	(0.017)
Genetic distance			−0.001**	−0.001*
			(0.000)	(0.000)
Common language				0.082**
				(0.034)
Religious distance				−0.005
				(0.020)
Same country history				−0.071**
				(0.028)
Colonial history				0.011
				(0.016)
Origin fixed effects	x	x	x	x
Destination fixed effects	x	x	x	x
Observations	84,100	32,041	55,169	55,169
Adj. R^2	0.922	0.922	0.947	0.947

Notes: The outcome variable is the natural logarithm of collaborations between region-pairs plus one. Colocation indicates collaboration between users in the same region. Robust standard errors are reported in parenthesis. *** p < 0.01, ** p < 0.05, * p < 0.1.

Source: GHTorrent, Bureau of Economic Analysis, World Cities Database, CEPII, Obradovich et al. (2022), Spolaore and Wacziarg (2009), author's compilation.

number of users. The effect does not differ depending on whether both countries are small or just one, implying there is a smaller border effect among large countries. This points to substantial difficulties for developer communities in small countries to make connections to the hubs, which are mostly located in large countries.
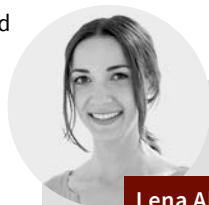
Lena Abou El-Komboz

is a Junior Economist at the ifo Institute and a PhD student at the University of Munich. Her research fields are the economics of platforms, digitalization, and regional policy.

Moritz Goldbeck

is an Associate at Analysis Group in San Francisco.[*] Previously, he completed his PhD at the University of Munich and was a Junior Economist at the ifo Institute. His research interests include innovation, labor economics, and competition policy.

[*] The views expressed here are those of the author and do not necessarily reflect those of Analysis Group or its clients.

POLICY CONCLUSIONS

The EU market features several disadvantages for digital business relative to its closest global competitors, the US and China. Scalability is generally lower not only due to administrative and bureaucratic barriers along national borders, but also because of soft factors like differences in language, preferences, or culture. Existing digital single market policymaking focuses on the consumer side of digital markets and is predominantly targeted at ensuring equal market access for consumers, e. g., by harmonizing VAT regimes and data protection law. While some business needs, like improving investment through a deepening of the capital market or interoperability and data access, are increasingly recognized, the production side is currently not sufficiently considered.

Especially the need to address inherent production-side challenges of the EU market due to a geographically distributed and multinational workforce is largely overlooked. A lower geographic concentration of knowledge workers is beneficial for regional cohesion, but at the same time requires more interregional collaboration to exploit the size of the EU's labor market. Our results show that international production in the digital economy is generally easier than in brick-and-mortar industries. However, a significant border effect of on average −16.4 percent still exists. For international collaboration of knowledge workers, soft factors like cultural and language differences are relatively more important. Since countries hosting only small knowledge worker communities exhibit higher border effects, a desirable side effect of policies fostering international collaboration is a potential decrease in regional inequality.

For the software industry, a well-known blind spot of studies based on patent data (see, e. g., the discussion in Abou El-Komboz et al. 2024), our results highlight the embeddedness of the industry in open-source communities. As we observe public activity of developers, our findings indicate significant positive effects on production in open innovation ecosystems through professional connections formed in and facilitated by organizations. As open innovation ecosystems produce valuable public goods (Korkmaz et al. 2024; Abou El-Komboz and Goldbeck 2024a), such interconnections potentially induce significant positive spillovers. In addition, the focus of the data on public activity, together with organizations playing a crucial role in connecting developers across national borders, suggests the effectiveness of measures to increase international collaboration and reduce the border effect might be even higher for non-public activity of private organizations.

Multinational production in the digital economy is not only a challenge but offers opportunities as well. International teams are likely better positioned to develop digital products that serve the diverse consumer base of the European market (Bahar et al. 2023). This advantage might extend to scaling digital business models effectively beyond any national border. Our findings suggest that international collaboration might be best facilitated through organizations and shared professional interests that connect people across borders. Thus, European policy aimed at a lasting increase in international collaboration among knowledge workers should consider targeting organizations rather than individuals.

REFERENCES

Abou El-Komboz, L. and M. Goldbeck (2024a), "Career Concerns as Public Good: The Role of Signaling for Open Source Software Development", *ifo Working Paper* 405 available at: https://www.ifo.de/en/publications/2024/working-paper/career-concerns-public-good-role-signaling-open-source-software-0.

Abou El-Komboz, L. and M. Goldbeck (2024b), "Virtually Borderless? Cultural Proximity and International Collaboration of Developers", *Economics Letters* 244, 111951.

Abou El-Komboz, L., T. Fackler and M. Goldbeck (2024), "Productivity Spillovers among Knowledge Workers in Agglomerations: Evidence from GitHub", *CESifo Working Paper* 11277 available at: https://www.cesifo.org/en/publications/2024/working-paper/productivity-spillovers-among-knowledge-workers-agglomerations.

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

Bahar, D., N. Carlson and E. Hernandez (2023), "Global Palette: The Impact of Immigrant Talent on Multinational Product Strategy", *The Wharton School Research Paper*, https://ssrn.com/abstract=4710865.

Blum, B. and A. Goldfarb (2006), "Does the Internet Defy the Law of Gravity?", *Journal of International Economics* 70, 384–405.

Conte, M., P. Cotterlaz and T. Mayer (2022), *The CEPII Gravity Database*, CEPII.

Emanuel, N., E. Harrington and A. Pallais (2023), "The Power of Proximity to Coworkers: Training for Tomorrow or Productivity Today?", *NBER Working Paper* 31880.

Foucart, R. (2024), "Europe's Tech Industry Is Lagging behind the US – But It Gives the Continent a Chance to Write the Rules of the Game", *The Deepening*, https://www.thedeeping.eu/2024/03/12/europes-tech-industry-is-lagging-behind-the-us-but-it-gives-the-continent-a-chance-to-write-the-rules-of-the-game/.

Goldbeck, M. (2023), "Bit by Bit: Colocation and the Death of Distance in Software Developer Networks", *ifo Working Paper* 386.

Gousios, G. (2013), "The GHTorent Dataset and Tool Suite", *IEEE 10th Working Conference on Mining Software Repositories (MSR)*, 233–236.

Havranek, T. and Z. Irsova (2017), "Do Borders Really Slash Trade? A Meta-analysis", *IMF Economic Review* 65, 365–396.

Jones, B. F. (2009), "The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder?", *The Review of Economic Studies* 76, 283–317.

Korkmaz, G., J. B. Santiago Calderón, B. L. Kramer, L. Guci and C. A. Robbins (2024), "From GitHub to GDP: A Framework for Measuring Open Source Software Innovation", *Research Policy* 53(3), 104954.

McCallum, J. (1995), "National Borders Matter: Canada-US Regional Trade Patterns", *American Economic Review* 85, 615–623.

Obradovich, N., Ö. Özak, I. Martín, I. Ortuño-Ortín, E. Awad, M. Cebrián, R. Cuevas, K. Desmet, I. Rahwan and Á. Cuevas (2022), "Expanding the Measurement of Culture with a Sample of Two Billion Humans", *Journal of the Royal Society Interface* 19(190), 20220085.

Santamaría, M., J. Ventura and U. Yeşilbayraktar (2023a), "Borders Within Europe", *NBER Working Paper* 28301.

Santamaría, M., J. Ventura and U. Yeşilbayraktar (2023b), "Exploring European Regional Trade", *Journal of International Economics* 146, 103747.

Spolaore, E. and R. Wacziarg (2009), "The Diffusion of Development", *The Quarterly Journal of Economics* 124, 469–529.

Wachs, J., M. Nitecki, W. Schueller and A. Polleres (2022), "The Geography of Open Source Software: Evidence from GitHub", *Technological Forecasting and Social Change* 176, 121478.

Wuchty, S., B. F. Jones and B. Uzzi (2007), "The Increasing Dominance of Teams in Production of Knowledge", *Science* 316(5827), 1036–1039.