The Fluidity of an Individual's Core-Periphery Position in Digital Knowledge Fields

Aaron Schecter
University of Georgia
aschecter@uga.edu

Sabine Brunswicker
Purdue University
sbrunswi@purdue.edu

Abstract

The literature on digital knowledge fields suggests that knowledge coproducers are embedded in a core-periphery social structure. This structure engenders an individual-level tension: be in the core where there is support for successful knowledge integration or be in the periphery where one can work outside of peer pressure. In this paper, we focus on the fluidity of core-periphery structures. We study the case of nanoHUB, a digital knowledge field of nanoscience and engineering. We analyze 17,821 contributions made by 251 knowledge producers who coproduce 609 scientific simulation tools over a nearly ten-year period, encompassing over six million lines of code. We find that knowledge producers seek to resolve the core-periphery tension by moving towards and then away from the temporal core. Additionally, we find that proximity to the temporal core at the point of the knowledge production has a curvilinear relationship with code produced.

1. Introduction

Digital technologies empower individuals to *coproduce* new knowledge in a variety of knowledge fields - from software development (e.g. Sorgeforce.net), statistics (R software), nanoscience & engineering (nanoHUB) or music (e.g. ccMixter.org) [1]–[3]. Such digitally-mediated knowledge coproduction can take place at a distance, outside of formal employment relationships, and through asynchronous and artifact-mediated communication [1], [4], [5]. These digital knowledge fields constitute a social structure woven from interactive coproduction processes among individual knowledge producers across multiple projects and artifacts (e.g. software projects). In these fields, knowledge producers simultaneously strive to produce and contribute new knowledge to the field while sharing common interests in advancing the knowledge in the field as a whole and using the same digital tools and resources available via a common digital platform [4]–[6].

The emerging literature on digital *knowledge fields* suggests that despite the unique characteristics of digitally-enabled coproduction, the digital knowledge field's social structure follows a core-periphery logic well known from the pre-digital era [4], [6]: The field constitutes a cohesive subgroup of knowledge producers in the core surrounded by a set of peripheral actors only loosely connected to the core [7]–[9]. There is also evidence that an individual's position within this core-periphery structure has implications for her or his relative ability to produce and contribute new knowledge (e.g., writing new lines of code), or productivity for short. In addition, prior literature points to the well-known core-periphery tension that each individual knowledge producer in the digital knowledge field is exposed to: Both the core and the periphery have positive implications for an individual's ability to produce and contribute knowledge to the field. Prior literature on knowledge fields provides a variety of explanations of how individuals may resolve this tension: e.g. by seeking an intermediate position between the core and the periphery, by mutually occupying a core position within the social structure and a marginal position within the semantic knowledge structure of the same field, or by complementing a core position within the field with many linkages to other external knowledge fields [4], [6].

However, existing research on the core-periphery tension in digital knowledge fields fails to account for the fluidity and the dynamics inherent to digitally-mediated knowledge coproduction [1], [5]. In digital knowledge fields, knowledge coproduction happens in the absence of stable structural social mechanisms. Instead, knowledge coproduction implies that knowledge producers dynamically respond to emerging tensions over time [1], [10], [11]. The dynamics in digital knowledge fields leads to two assumptions that have not been sufficiently incorporated in the existing theories on the core-periphery tension in digital knowledge fields. First, in digital knowledge fields a core group of producers within a knowledge field is constituted by a temporal social structure; a cohesive core of producers forms "in the moment" as a result of a common passion or in response to emerging needs to

coordinate coproduction efforts within the knowledge field [1]. As a result, this temporal core is fluid and moves around based on where the focus of co-work is at any given time. Second, a dynamic view assumes that knowledge producers' proximity to the temporal core will fluctuate over time as the network itself shifts [1], [10].

To incorporate such a dynamic view into the theoretical discourse on an individual's core-periphery tension in digital knowledge fields, this paper asks the following two research questions: *How does knowledge producers' proximity to the temporal core change over time? What are the productivity implications of a producer's proximity to the core at a particular point of time?* To answer these questions, we apply theories of social capital from the network science literature, and study the case of nanoHUB, a community of knowledge producers in the interdisciplinary knowledge field of nanoscience & engineering. We analyze 17,821 contributions made by 251 scientific tool producers to 609 scientific software projects over a nearly ten-year period, encompassing over six million lines of code contributed.

2. Theoretical Development

Digital *knowledge fields* emerge from digitally-mediated interactions among loosely connected social actors at a distance, outside of formal employment relationships and organizational hierarchy [6], [8]. They constitute a relational social structure that emerges from artifact-mediated coproduction processes among multiple knowledge producers who create digital artifacts using the same digital platform and coproduction tools [12]. When coproducing digital artifacts, producers engage with each other's knowledge contributions (e.g. code for a software development task or comments to an idea submission) via digitally-mediated interactions [13]. Coproduction implies a dynamic exchange of information and knowledge resources between two or more producers working on a common task or project [14]. Thus, when integrating their different knowledge resources in order to jointly produce common artifacts [15], two knowledge producers form a social relationship that embeds experiences and social practices of coproduction [12], [15]. Prior literature suggests that such artifact-mediated coproduction follows a core-periphery structure [7]: A densely connected and cohesive subset of producers, the so-called the core, which is surrounded by peripheral knowledge producers.

The pre-digital view of knowledge fields assumes that such core-periphery structures remain stable once they have converged. However, there is evidence that in digital knowledge fields, social structures are more fluid

[6], [10], [12], [16]. In other words, the notion of a stable core group does not exist [10, p. 3044]. Instead, there is only temporal convergence. A temporal "core" of densely connected coproducers forms based on emerging tasks and resource needs as well as individual preferences [1]. For example, a public event, such as the death of a popular singer, may unify a subset of Wikipedia contributors with interest in music to converge temporarily and align their actions to create new or update existing Wikipedia pages [1], [17], [18]. These coproducers synchronize their coproduction efforts for a short period of time. Once the task is completed, the cohesive relationships among the coproducers weaken or dissolve completely. In a temporal core, cohesively connected coproducers are engaging in 'situated' digitally mediated social interactions in order to integrate each other's knowledge using temporary practices of coproduction [19]. Such patterns have also been found in other digital coproduction settings such as video game design, software development, or crowdsourced innovation challenges [20]–[23].

Following theories of social capital and network embeddedness, we argue that a position at a particular point of time in both the core and the periphery of a knowledge field signals positive consequences to the individual [6]. On the one hand, being proximate to the core signals an information advantage [1], [12]. Individuals who are in or very proximal to the core may be better able to determine what types of contributions are of focal interest at a particular point of time. A core position may facilitate an individual in crafting contributions that matter to the field now. Further, proximity to the core is assumed to provide access to support from other coproducers who share a similar passion in the topics and task at hand. Thus, a core position can help producers with successfully integrating their knowledge within the field [4], [6].

A peripheral position allows for interactions only loosely related to the focal point of activity of the knowledge field. Thus, it offers knowledge producers the ability to work at their own pace, guided by their own interests without much need to coordinate with others [24]. Further, it gives producers access to information and knowledge somewhat unrelated to the central point of attention within the field. As a result, a producer may also see benefits in seeking a position *in the periphery.*

We argue that individuals may seek to resolve this tension by dynamically adjusting their position over time [1], [10]. At some points in time, producers might want to move closer to the core in order to work on topics that currently matter (or are popular) and easily integrate their knowledge within the field. However, at some point a producer may want to recess and move to

the periphery to work on unrelated topics and ideas currently not of focal interest within the knowledge field. In either case, rather than remaining in one position, they adapt their proximity to the core over time. Thus, we advance the following hypothesis:

H1: A knowledge producer's proximity to the temporal core is a non-linear function of time.

If a knowledge producer is in the core at the time of knowledge production, she/he can expect support from others in creating and implementing the knowledge into the field. Knowledge producers in the core share high levels of collective passion and "devoted enthusiasm" for advancing the field with the tasks and work at hand [1]. Further, even though a temporal core implies only temporary convergence, the temporal cohesion among the core members offers coordination benefits. Core coproducers are somewhat in sync with each other because they are densely connected and interact with each other in a highly frequent way. Thus, they have greater opportunity to fix any unresolved interdependencies [15]. This proximity to the core allows members to solve problems faster and potentially more efficiently.

However, being close to the core can create cognitive overload, peer pressure and anxiety since it requires full attention to the actions of each other core producer without much delay [1], [22], [25]. Since practices in the core are only temporary, being in the temporal core also requires a producer to invest time and effort into the development of so called "situated" production practices for the particular tasks at hand [1]. Intensive debates and discussions may create cognitive overload and anxiety, which may eventually hinder the individual knowledge producers in generating new and useful knowledge at the time of being in the core.

In a peripheral position, on the other hand, individuals can escape such challenges and pressures [4], [6], [26], [27]. The periphery offers an individual the opportunity to explore topics without being constrained by the interests and practices in the temporal core [24]. Such freedom to explore unrelated knowledge has positive consequences for a producer's ability to produce knowledge and ideas that create value for the knowledge field in the long run. However, it may not translate into greater productivity at the time of the knowledge production. In the periphery, the knowledge producers lack the support and assistance from other densely connected producers.

Since neither a position in the core nor a position in the periphery seem to be optimal from a productivity point of view, an intermediate position at a particular point of time might correlate with the highest productivity level. Essentially, for a knowledge producer to best mobilize the benefits of both the core

as well as the periphery, it seems critical for them to occupy an intermediate position at the time of their knowledge contribution. An intermediate position allows an individual to receive assistance from the densely connected and temporarily converged core, without being too entangled in conflict resolution and the development of situated practices and norms. At the same time, it also offers sufficient opportunities to get inspired by unrelated topics without being completely disconnected from the interests and the resources of the core at a particular point of time.

Thus, we suspect that in a digital knowledge field that exhibits dynamics and fluidity, there is also a U-shaped relationship between proximity the core and an individual's productivity. However, unlike in the pre-digital era, this U-shaped relationship unfolds at a particular point of time. Essentially, for a knowledge producer to best mobilize the benefits of both the core as well as the periphery, it seems critical for them to occupy an intermediate position *at the time of their knowledge contribution*. We therefore advance the following hypothesis:

H2: A knowledge producer's proximity to the core at the time of her/his contribution has an inverse U-shaped relationship with productivity.

3. Methods

3.1. Case Setting and Data

To test our hypotheses, we chose nanoHUB, a digital platform in the interdisciplinary field of nanoscience and nanotechnology. NanoHUB was launched as part of the NSF-funded Network of Computational Nanotechnology at Purdue University [28]. The tools on the platform are web-based software programs (web applets) which allow users to conduct simulations or complex calculations for research but also educational purposes.

In this study we focused on a subset of nanoHUB's members: the community of knowledge producers (scientists and engineers), who coproduce interactive simulation software tools using the platform's standardized programing interfaces, development toolkits, and software version control during period of April 2005 to September 2014. Our dataset is comprised of 251 knowledge producers interacting in 609 tools. We only selected individuals who contributed more than once and registered a non-zero distance from the core at some point. We used a versioning and revision control system, Lean Apache Subversion (SVN), to capture how individuals iteratively contribute to a certain tool. The SVN log provides access to commit trace data, which includes the exact time stamp when a knowledge producer makes a unique commit to a tool.

In total, we observe 17,821 distinct commits and 61,665,012 lines of code written across an approximately ten-year period. On average, knowledge producers initiated 71 commits with a standard error of 164.68, and on average contributed to 7.45 tools with a standard error of 36.99. For tools, the average number of received commits was 29.26 with a standard error of 369.14, and these changes were made by an average of 3.07 knowledge producers with a standard error of 3.51. Using this data, we conduct two separate sets of regressions, which we proceed to describe.

3.2. Measuring Proximity to the Core

Using Borgatti and Everett's algorithm [7], we are able to assign each individual a score on the core-periphery continuum. We use this continuous measure of proximity to the core because we are interested in relative position along the continuum [4], [9]. Larger values of proximity indicate that an individual is closer to the core, while smaller values indicate that an individual is closer to the periphery. We first find the principal eigenvector of the adjacency matrix Y_t for the network at time t. Let $\lambda(Y_t)$ be this $N \times 1$ principal eigenvector. Then, for any producer i, the value $\lambda(Y_t)_i$ is equal to their proximity to the network core. For any knowledge producers that are inactive during the time period, their corresponding entry in the eigenvector will be 0. This vector is normalized so that values can be compared across time periods. Values of proximity to the core can range in value from 0 to 1; in our dataset, this measure varied between 0 and 0.71, with a mean of 0.21 and standard deviation of 0.25.

3.3. Analysis of the Fluidity of an Individual's Proximity to the Core

Our first dependent variable is the proximity of a knowledge producer to the core during a given period. This proximity is calculated as described in the previous section. We delineate periods by considering the window of time leading up to, but not including, the instance of an individual contributing. Proximity to the core is calculated using all events that have transpired within that window; we use a time window of 7 days. Our main independent variable is the time a knowledge producer has been active. We identify the first instance at which the knowledge producer contributes, and for each subsequent activity, we calculate the time elapsed (in weeks) since that original contribution. In order to capture non-linear time trends, we also include a square and cubic term.

To test our first hypothesis, we estimate the temporal dynamics of proximity to the core using the following OLS regression equation:
$$\lambda_{it} = \beta_0 + \beta_1 \tau_i + \beta_2 \tau_i^2 + \beta_3 \tau_i^3 + \alpha_i + \varepsilon_{it}.$$
In this equation, λ_{it} is develop $i's$ proximity to the core at time t, τ_i is the time elapsed since i was first active, and α_i are knowledge producer fixed effects. We use a fixed effects model to account for unobserved heterogeneity in behavior [29].

3.4. Analysis of Productivity Implications of an Individual's Proximity to the Core

For the second stage in our analysis, the dependent variable of interest is productivity, which we measure as the lines of code (LOC) changed in a single edit. We use LOC because of its relationship with software functionality and complexity [30], [31]. In our dataset, the lines of code modified ranges from 1 to 356,171, with a median of 118, a mean of 3460.24, and a standard deviation of 19,093.33 lines. Given the skewed nature of the data, we take the natural logarithm of each observation. Our main independent variable is a knowledge producer's proximity to the core at the time they made their contribution. Following prior work, we also include the square of proximity to the core to capture any curvilinear effects [9], [32].

We include several control variables in our study. First, we include two attribute covariates, knowledge producer tenure and project age. Tenure measures the time since the individual was first active (i.e., their first contribution of code). Project age is a measure of how much time has elapsed since the project was created. Second, we include three structural controls which account for prior behavior [33]–[35]. The first is inertia, which we operationalize as the total amount of code knowledge producer i has contributed to project j in the past. The second is activity, which is equal to the total lines of code produced by knowledge producer i in the past. Finally, project activity is a measure of how many lines of code have been contributed to tool j in the past. Because of the skewed nature of the data, we take the log of the lines of code contributed for each of these three measures.

We test our second hypothesis by estimating the following fixed effects linear regression equation:
$$\log c_{ijt} = \theta_0 + \theta_1 \lambda_{it} + \theta_2 \lambda_{it}^2 + \gamma X_{ijt} + \eta_i + \varepsilon_{ijt}.$$
In the above expression, c_{ijt} is the lines of code contributed by i to j at time t. The value λ_{it} is knowledge producer i's proximity to the core at time t, and X_{ijt} is a vector of control variables as defined previously. Finally, we include knowledge producer fixed effects η_i to account for unobserved heterogeneity [29].

4. Results

We conducted our analysis in four steps: first, we provide descriptive statistics for our key study variables; second, we analyze the time trends of proximity to the core; third, we analyze the relationship between knowledge producer position and productivity; finally, we test the robustness of our results. In Table 1, we present descriptive statistics and correlations between our study variables.

Table 1. Descriptive statistics and inter-correlations for key study variables

Variable	1.	2.	3.	4.	
1. LOC (log)					
2. Prior Contribution	0.06				
3. Prior Activity	0.13	0.02			
4. Prior Project Activity	0.04	0.42	0.27		
5. Proximity to the core	0.06	0.25	0.54	0.55	-
Mean	4.84	0.86	0.09	0.33	0.21
SD	2.54	0.28	0.09	0.31	0.25
VIF		1.23	1.44	1.63	1.89

Note. N = 17,821. Pearson correlation coefficients are reported.

We note that some of the correlations, particularly those between proximity to the core and controls, are somewhat high. To ensure that multicollinearity does not negatively impact our analyses, we calculate the variance inflation factor (VIF) for each independent variable. All VIFs are below two, which is well below the common threshold of five [36]. We therefore conclude that multicollinearity does not significantly affect our models.

4.1. Main Results

We test our first hypothesis by examining the fluidity of an individual's proximity to the core. To do so, we estimate a cubic regression model with knowledge producer fixed effects. Our results are presented in Table 2. All standard errors presented are robust and grouped by individual. We find that there is a linear time trend ($\beta = 0.015$, $p < 0.001$), negative quadratic trend ($\beta = -0.024$, $p < 0.001$), and positive cubic trend ($\beta = 0.003$, $p < 0.01$). These significant coefficients for time indicate that knowledge producers do indeed change their position in the digital knowledge field over time and follow a non-linear pattern. Thus, we find support for Hypothesis 1. We next test our second hypothesis by determining the relationship between productivity – i.e., lines of code contributed – and a knowledge producer's proximity to the core at the time of contribution. Our results are presented in Table 3.

Table 2. Regression of Time on Core Proximity

Variable	Model 1	Model 2
Time	0.001	0.015
	(0.002)	(0.003)***
Time^2		-0.024
		(0.003)***
Time^3		0.003
		(0.001)**
Knowledge Producer FE	Yes	Yes
R^2	0.746	0.748
Deviance	488.14	484.17
Num. Producers	251	251
Num. Events	17,821	17,821

*Significance codes: * p < 0.05, ** p < 0.01, *** p < 0.001. Robust standard errors in parentheses.*

Table 3. Regression of Core Position on log(LOC)

Variable	Model 3	Model 4	Model 5
Prior Contribution	0.709	0.706	0.774
	(0.094)***	(0.094)***	(0.109)***
Prior Activity	2.903	2.681	3.431
	(0.282)***	(0.317)***	(0.341)***
Prior Project Activity	-0.074	-0.092	-0.367
	(0.122)	(0.122)	(0.131)**
Proximity to the Core		0.199	2.955
		(0.129)	(0.487)***
Core Proximity^2			-4.268
			(0.727)***
Knowledge Producer FE	Yes	Yes	Yes
Pseudo R^2	0.801	0.801	0.804
Deviance	104,430	104,416	104,212
Num. Producers	251	251	251
Num. Events	17,821	17,821	17,821

*Significance codes: * p < 0.05, ** p < 0.01, *** p < 0.001. Robust standard errors in parentheses.*

We find that proximity to the core has a positive linear ($\theta = 2.955$, $p < 0.001$) and negative quadratic ($\theta = -4.268$, $p < 0.001$) relationship with the log of lines of code. Put another way, there is positive but decreasing relationship between proximity to the core and productivity. However, when we conduct a two-line test for the U-shape [37], we find that the relationship is not a true "U." Instead, it is characterized by an initial increase, and then a leveling off effect. We conclude that a knowledge producer will have higher productivity – i.e., contribute more code – when they occupy an intermediary position in the digital knowledge field at the time of contribution, but do not gain or lose by moving more towards the core. Thus, we find partial support for Hypothesis 2.

4.2. Robustness Checks

For our analysis of the fluidity of a knowledge producer's proximity to the core, we tested our models on subsets of the data. We excluded knowledge producers in the bottom tenth percentile of contributions, the top tenth percentile, and both the bottom and top tenth percentiles. Additionally, we tested two alternative models, the Tobit regression model and a random effects model. A Tobit regression model allows us to explicitly account for limitations on the dependent variable; because proximity to the core is restricted to the range [0, 1], this model is appropriate. We also estimated a random effects model, which treats the producer effects as random variables. We find that the overall trend is consistent with the main findings. Finally, we recomputed proximity to the core for each knowledge producer such that the time window stretched from seven days to thirty and ninety. We again find consistent results with longer time intervals, indicating that a knowledge producer's position is fluid, even when we "zoom out" our view of their behavior.

Like our analyses of dynamics, we reran our models on subsets of the data, excluding the bottom and top ten percent of contributors based upon activity and volume. Next, we tried alternative model specifications. Because our productivity measure is a count variable that is right skewed, we ran a fixed effects negative binomial regression model using lines of code as the dependent variable. We also ran a random effects regression model instead of a fixed effects model and found consistent results. Additionally, we considered an alternative means of calculating the control variables. For our alternative model, we calculated the same statistics, but weighted them by the lines of code contributed at each event. The results of all these tests were consistent with our main findings. Finally, we recomputed proximity to the core for each knowledge producer over longer time intervals, ranging from two weeks up to three months. Interestingly, we find that the relationship between proximity to the core and performance is consistent for a time interval up to three weeks – albeit with a smaller effect size – but disappears for one month and longer. This finding augments our main conclusion: a knowledge producer's productivity is a function of their *recent position only*.

Finally, we considered the possibility that our measure of time is what influences productivity, not proximity to the core. In other words, we need to rule out the possibility that the variance explained by proximity to the core is the same variance explained by time; this would be a key endogeneity problem. First, we regressed the cubic time trend on lines of code. Second, we regressed the time trend as well as all of the original independent variables on our productivity measure. The results of these models are presented in Table 4. We find that tenure (time since joining the community) has no significant impact on productivity, measured by lines of code contributed (see Model 6). Further, the inclusion of the time trend does not affect the relationship between proximity to the core and productivity (see Model 7). Thus, tenure only influences productivity through our measure of proximity to the core.

Table 4. Regression of Proximity to the Core and Time on log(LOC)

Variable	Model 6	Model 7
Time	0.026	-0.009
	(0.042)	(0.044)
Time^2	-0.012	-0.011
	(0.047)	(0.047)
Time^3	-0.000	0.006
	(0.015)	(0.015)
Prior Contribution		0.775
		(0.095)***
Prior Activity		3.428
		(0.343)***
Prior Project Activity		-0.361
		(0.138)**
Proximity to the Core		2.965
		(0.488)***
Proximity to the Core^2		-4.282
		(0.728)***
R^2	0.801	0.804
Deviance	105,409	104,211
Num. Producers	251	251
Num. Events	17,821	17,821
*Significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*		

5. Discussion

In this paper, we investigate the fluidity in digital knowledge fields in which knowledge producers coproduce artifacts. We specifically focus on "proximity to the core," which is the distance of a knowledge producer from the core of the field network [7]. While existing literature suggests that individuals may resolve the core/periphery tension by seeking an intermediate proximity to the core [4], [9], this work fails to account for an important characteristic of core/periphery structures in digital knowledge fields: They exhibit only temporary convergence. In other words, prior empirical studies often assume a stable core, but we challenge that assumption and argue that the core/periphery structure is continually changing. Such a challenge follows Faraj et al.'s (2011) logic that an online community or digital knowledge field "is not likely to be in equilibrium, nor should an equilibrium

state even be desirable. Rather, the tensions are likely to ebb and flow, with each flux providing equally viable opportunities for knowledge collaboration" (p. 1230).

In this paper, we account for such temporal convergence and fluidity to disentangle how an individuals' position within the core/periphery structure as a temporary property that changes over time. Rather than modeling coproduction interactions at the aggregate level, we make use of digital trace data to unpack the fine-grained temporal changes taking place in the context of short-term coproduction relationships. By doing so, we make an important finding: there is fluidity in an individual's proximity to the core, suggesting an alternative way of how individuals seek to resolve the core/periphery position. Further, our study also suggests that this fluidity has implications for an individual's productivity throughout her/his trajectory in participating in a digital knowledge field.

5.1. Theoretical Contributions

We first contribute to the debate on the core periphery tension in knowledge fields and how it manifests in knowledge production. We bring to light the differences in the expression of this tension when we focus on knowledge fields in which temporary convergence is possible, such as those composed of digital artifacts. This fluid character contrasts with other stable fields such as art or science c.f. [9]. In those contexts, being central or core brings with it status, power, and recognition [5]. Over a longer period, coproduction between any two individuals may happen at very different times; thus, the field's core/periphery structure represents knowledge, practices, and values that the field has converged upon. However, it does not necessarily represent the knowledge and practices that emerge "in the moment" when multiple coproducers work on a task that requires social interaction and knowledge integration rather than independent work. Thus, in a digital knowledge field, ties and the concept of a "core" take on a different meaning. A temporarily convergent core provides a vehicle for articulating the efforts of multiple knowledge coproducers, while the core members maintain a high degree of flexibility. Indeed, the composition of the core will evolve as passion and energy shifts across projects, with individuals converging on shared tasks. Rather than maintaining a stable intermediate position [4], c.f. [9], individuals may find themselves in both core and periphery positions as the overall network structure fluctuates. Much like Burt and Merluzzi's theory of oscillation [38], we argue that individuals may actually benefit (or at least satisfy tensions) by leveraging their alternating positions between the core and periphery.

The nature of knowledge fields mediated by digital artifacts, such as nanoHUB, also enables the type of movement we observe. Without formal organizational boundaries or physical constraints, individuals are able to engage with a multitude of projects [6]. Because knowledge producers can freely move from one tool to another, they are able to rapidly change their position within the network of coproduction. Further, the lack of a hierarchy within certain knowledge fields allows any individual to attain a central position, even without status, recognition, or authority [5], [39]. As a result, the network can be rapidly reconfigured as a result of individual actions. Given that different positions within a core/periphery structure of a dynamic knowledge field can yield productivity benefits, it is possible that motivated members of the knowledge field will try to move through these positions. Of course, this behavior assumes a degree of agency beyond what we model in this paper.

Second, this study also contributes to the literature on fluidity in online communities. We build on the notion of fluidity by showing that social structures such as the network core can form, evolve, or dissolve over time. Our findings provide some empirical evidence for the tensions described by Faraj et al. (2011) with a focus on the individual-level core/periphery tension. For instance, we argue that members of the knowledge field we studied faced time-related tensions, in that individuals do not maintain a specific position over long periods. Further, the tension of temporary convergence around a problem space provides an interesting lens through which to view our results. Because of its benefits, individuals may be drawn to these periods of convergence, which we observe as some individuals become closer to the core. However, temporary convergence subsequently drives individuals away from the core. Thus, it is possible that this tension drives the movement we observe.

Relatedly, our work builds on the notion of turbulent stability and emergent roles [16]. Our focus on the actions of individuals demonstrates that by and large their behaviors are not constant over time. Knowledge producers will, at times, engage in coproduction with a closely connected group of others; at other times, they may work independently or with relatively few others. In other words, people tend to ebb and flow in and out of deeply embedded social network positions based on their current. While prior work emphasizes the dynamism inherent in forms of contribution c.f. [10], we add to these findings and suggest that there can emergent structural roles as well. Though the overall network's structure may not change, individuals are likely continuously updating themselves.

Third, our study also adds to the growing body of literature on social network dynamics and the structure

of organizing in two ways. First, there are a number of studies that attempt to disentangle the effects of short-term versus long-term patterns of organizing [12]. Rather than compare network structures across temporal dimensions, we instead focus on the short-term explicitly. We argue that in the context of a digital knowledge field, "zooming in" matters because it changes the nature of the interactions. Indeed, our analyses suggest that over longer time windows the productivity implications disappear. One potential explanation is the observability of digital artifacts and the fluidity of the knowledge fields they constitute [40]. Because a member of the knowledge field can always refer to old code, their connections to projects are less relevant than who or what they are working on at the moment. Accordingly, social networks should be defined considering the underlying processes they are meant to represent. Second, our study emphasizes the dualistic nature of network structure and behaviors e.g., [41]. Network theorists posit that networks shape and constrain individual behaviors, while the micro-level actions of agentic individuals can subsequently change the network [42]. Thus, we argue that acknowledging these dual forces is critical to understanding the growth and evolution of digital knowledge fields.

5.2. Practical Implications

Our study also has important practical implications. We find that knowledge producers become closer to the core early in their tenure but then recede into the periphery, which presents unique challenges for the long-term health of the field. Though individuals can move freely into and out of the core, it is important that when knowledge producers move out, they do not exit the knowledge field entirely. If members eventually leave, rather than reengaging with others at some stage, it may become "costly" to take advantage of fluidity and dynamics for the creation of new knowledge in the field [1]. The constant flux of membership reduces social integration, social learning, and development of norms, potentially leading to even greater overall turnover [11]. Thus, we suggest that there are two potential avenues that platform owners can pursue. First, they could offer incentives and technical affordances for knowledge producers who move to the periphery in order to encourage movement back towards the core. For example, prior work has demonstrated that increasing recognition and prestige of contributors is an effective means of encouraging further contributions [43]. Second, they could try to increase the number of newcomers who join the digital knowledge field to replace members who drop out. For instance, platform owners can make it easier to join and contribute by

lowering boundaries to entry [44], or encourage socialization of new members [45].

Our findings also suggest that for knowledge producers to produce at a high level, they need to be able to access the core and periphery of the digital knowledge field at different times. Owners of digital knowledge fields should first ensure that the platform implements tools and features that support easy access to knowledge across independent artifacts. Second, owners of digital knowledge fields should consider governance strategies that allow for greater decision-making rights and autonomy among knowledge producers [46]. Of course, this last recommendation should be balanced with each knowledge field's unique needs in terms of desired output, standardization, or rate of change.

5.3. Limitations

There are several limitations to our study that are worth noting. First, our study primarily focuses on measures constructed from digital traces and does not incorporate any psychological measures. Future work could strengthen our arguments by surveying individuals or designing controlled experiments that elicit choices to assess their intentions.

A second limitation stems from the bimodal nature of our network data. Namely, the relationships in the network exist between developers and projects, not between knowledge producers. We assume that the technical architecture of software itself reflects knowledge producer interactions and embeds not only explicit but also implicit knowledge and norms of how work is done. However, it is possible that the knowledge producers are coordinating through some unobserved channels, such as mailing lists or social media.

Finally, our data was collected from software tools on the scientific platform nanoHUB. Though we expect this community to have many characteristics of other digital platforms, there are nevertheless aspects of nanoHUB which make it a unique context. nanoHUB attracts only a certain type of individuals: those with specialized skills and also career interests in science and engineering. Thus, tools on nanoHU represent a very distinct form of digital innovation, a scientific digital innovation. Our study could be strengthened by replicating our findings in other digital innovation settings, such as developer communities of platforms for mobile apps, music or gaming.

6. Conclusion

The emerging literature on digital knowledge fields suggests that knowledge producers are embedded in a social structure that mirrors the core-periphery logic well known from the pre-digital era [4], [6]. However,

we argue that there is greater fluidity in these social structures. We hypothesize that knowledge producers experience movement both towards and away from the temporal core, leading to periods of temporary convergence among subsets of individuals. Further, we argue that an individual's productivity in a digital knowledge field is predicated on their position in the network at the time of their contribution. We test and confirm our hypotheses using digital trace data from a community creating technology simulation tools over a ten-year period. Our findings underscore the inherent tensions faced by knowledge producers in a digital field and highlight the emergent and variable nature of structural roles. There is much to learn about the temporality of core-periphery structures and their role in knowledge production; we hope this work contributes to this conversation and spurs additional exploration.

7. References

[1]   S. Faraj, S. L. Jarvenpaa, and A. Majchrzak, "Knowledge Collaboration in Online Communities," *Organization Science*, vol. 22, no. 5, pp. 1224–1239, Feb. 2011, doi: 10.1287/orsc.1100.0614.

[2]   S. Nambisan, "Digital Entrepreneurship: Toward a Digital Technology Perspective of Entrepreneurship," *Entrepreneurship Theory and Practice*, vol. 41, no. 6, pp. 1029–1055, Nov. 2017, doi: 10.1111/etap.12254.

[3]   Y. Yoo, R. J. Boland, K. Lyytinen, and A. Majchrzak, "Organizing for Innovation in the Digitized World," *Organization Science*, vol. 23, no. 5, pp. 1398–1408, 2012.

[4]   L. Dahlander and L. Frederiksen, "The Core and Cosmopolitans: A Relational View of Innovation in User Communities," *Organization Science*, vol. 23, no. 4, pp. 988–1007, Jul. 2011, doi: 10.1287/orsc.1110.0673.

[5]   N. Levina and M. Arriaga, "Distinction and Status Production on User-Generated Content Platforms: Using Bourdieu's Theory of Cultural Production to Understand Social Dynamics in Online Fields," *Information Systems Research*, vol. 25, no. 3, pp. 468–488, Sep. 2014, doi: 10.1287/isre.2014.0535.

[6]   H. Safadi, S. L. Johnson, and S. Faraj, "Core-Periphery Tension in Online Innovation Communities," *Organization Science*, p. 70, 2020.

[7]   S. P. Borgatti and M. G. Everett, "Models of core/periphery structures," *Social Networks*, vol. 21, no. 4, pp. 375–395, 1999, doi: 10.1016/S0378-8733(99)00019-2.

[8]   P. Bourdieu, *The Field of Cultural Production: Essays on Art and Literature*. Columbia University Press, 1993.

[9]   G. Cattani and S. Ferriani, "A Core/Periphery Perspective on Individual Creative Performance: Social Networks and Cinematic Achievements in the Hollywood Film Industry," *Organization Science*, vol. 19, no. 6, pp. 824–844, Dec. 2008, doi: 10.1287/orsc.1070.0350.

[10]   G. C. Kane, J. Johnson, and A. Majchrzak, "Emergent Life Cycle: The Tension Between Knowledge Change and Knowledge Retention in Open Online Coproduction Communities," *Management Science*, vol. 60, no. 12, pp. 3026–3048, Apr. 2014, doi: 10.1287/mnsc.2013.1855.

[11]   S. Ransbotham and G. C. (Jerry) Kane, "Membership Turnover and Collaboration Success in Online Communities: Explaining Rises and Falls from Grace in Wikipedia," *MIS Quarterly*, vol. 35, no. 3, pp. 613–627, 2011, doi: 10.2307/23042799.

[12]   S. Brunswicker and A. Schecter, "Coherence or flexibility? The paradox of change for developers' digital innovation trajectory on open platforms," *Research Policy*, Apr. 2019, doi: 10.1016/j.respol.2019.03.016.

[13]   L. P. Robert, A. R. Dennis, and M. K. Ahuja, "Social Capital and Knowledge Integration in Digitally Enabled Teams," *Information Systems Research*, vol. 19, no. 3, pp. 314–334, Sep. 2008, doi: 10.1287/isre.1080.0177.

[14]   G. Mangalaraj, S. Nerur, R. Mahapatra, and K. H. Price, "Distributed Cognition in Software Design: An Experimental Investigation of the Role of Design Patterns and Collaboration," *MIS Quarterly*, vol. 38, no. 1, pp. 249–274, Jan. 2014, doi: 10.25300/MISQ/2014/38.1.12.

[15]   A. Lindberg, N. Berente, J. Gaskin, and K. Lyytinen, "Coordinating Interdependencies in Online Communities: A Study of an Open Source Software Project," *Information Systems Research*, vol. 27, no. 4, pp. 751–772, Dec. 2016, doi: 10.1287/isre.2016.0673.

[16]   O. Arazy, J. Daxenberger, H. Lifshitz-Assaf, O. Nov, and I. Gurevych, "Turbulent Stability of Emergent Roles: The Dualistic Nature of Self-Organizing Knowledge Coproduction," *Information Systems Research*, vol. 27, no. 4, pp. 792–812, Sep. 2016, doi: 10.1287/isre.2016.0647.

[17]   B. Keegan, D. Gergle, and N. Contractor, "Do Editors or Articles Drive Collaboration? Multilevel Statistical Network Analysis of Wikipedia Coauthorship," p. 10, 2012.

[18]   B. Keegan, D. Gergle, and N. Contractor, "Hot Off the Wiki: Structures and Dynamics of Wikipedia's Coverage of Breaking News Events," *American Behavioral Scientist*, vol. 57, no. 5, pp. 595–622, May 2013, doi: 10.1177/0002764212469367.

[19]   E. Quintane, P. E. Pattison, G. L. Robins, and J. M. Mol, "Short-and long-term stability in organizational networks: Temporal structures of project teams," *Social Networks*, vol. 35, no. 4, pp. 528–540, 2013.

[20]   A. Majchrzak and A. Malhotra, "Effect of Knowledge-Sharing Trajectories on Innovative Outcomes in Temporary Online Crowds," *Information Systems Research*, vol. 27, no. 4, pp. 685–703, Nov. 2016, doi: 10.1287/isre.2016.0669.

[21]   C. Riedl and A. W. Woolley, "Teams vs. Crowds: A Field Test of the Relative Contribution of Incentives, Member Ability, and Emergent Collaboration to Crowd-Based Problem Solving Performance," *AMD*,

vol. 3, no. 4, pp. 382–403, Dec. 2017, doi: 10.5465/amd.2015.0097.

[22] M. Shaikh and E. Vaast, "Folding and Unfolding: Balancing Openness and Transparency in Open Source Communities," *Information Systems Research*, vol. 27, no. 4, pp. 813–833, Sep. 2016, doi: 10.1287/isre.2016.0646.

[23] B. Vedres and D. Stark, "Structural Folds: Generative Disruption in Overlapping Groups," *American Journal of Sociology*, vol. 115, no. 4, pp. 1150–1190, 2010, doi: 10.1086/649497.

[24] J. Howison and K. Crowston, "Collaboration Through Open Superposition: A Theory of the Open Source Way," *MIS Quarterly*, vol. 38, no. 1, pp. 29-A9, Mar. 2014.

[25] A. L. J. Ter Wal, O. Alexy, J. Block, and P. G. Sandner, "The Best of Both Worlds: The Benefits of Open-specialized and Closed-diverse Syndication Networks for New Ventures' Success," *Administrative Science Quarterly*, vol. 61, no. 3, pp. 393–432, Sep. 2016, doi: 10.1177/0001839216637849.

[26] F. Rullani and S. Haefliger, "The periphery on stage: The intra-organizational dynamics in online communities of creation," *Research Policy*, vol. 42, no. 4, pp. 941–953, May 2013, doi: 10.1016/j.respol.2012.10.008.

[27] P. Setia, B. Rajagopalan, V. Sambamurthy, and R. Calantone, "How Peripheral Developers Contribute to Open-Source Software Development," *Information Systems Research*, vol. 23, no. 1, pp. 144–163, Mar. 2012, doi: 10.1287/isre.1100.0311.

[28] L. Zentner, M. Zentner, V. Farnsworth, M. McLennan, K. Madhavan, and G. Klimeck, "nanoHUB.org: Experiences and Challenges in Software Sustainability for a Large Scientific Community," *arXiv:1309.1805 [cs]*, Sep. 2013, Accessed: Jul. 20, 2018. [Online]. Available: http://arxiv.org/abs/1309.1805

[29] J. M. Wooldridge, *Introductory econometrics: a modern approach*, 4th ed. Mason, OH: South Western, Cengage Learning, 2009.

[30] A. J. Albrecht and J. E. Gaffney, "Software Function, Source Lines of Code, and Development Effort Prediction: A Software Science Validation," *IEEE Transactions on Software Engineering*, vol. SE-9, no. 6, pp. 639–648, Nov. 1983, doi: 10.1109/TSE.1983.235271.

[31] J. Rosenberg, "Some misconceptions about lines of code," in *Proceedings Fourth International Software Metrics Symposium*, Nov. 1997, pp. 137–142. doi: 10.1109/METRIC.1997.637174.

[32] L. Dahlander and S. O'Mahony, "Progressing to the Center: Coordinating Project Work," *Organization Science*, vol. 22, no. 4, pp. 961–979, Sep. 2010, doi: 10.1287/orsc.1100.0571.

[33] G. Conaldi, A. Lomi, and M. Tonellato, "Dynamic models of affiliation and the network structure of problem solving in an open source software project," *Organizational Research Methods*, vol. 15, no. 3, pp. 385–412, 2012.

[34] G. Conaldi and A. Lomi, "The dual network structure of organizational problem solving: A case study on open source software development," *Social Networks*, vol. 35, no. 2, pp. 237–250, 2013.

[35] E. Quintane, G. Conaldi, M. Tonellato, and A. Lomi, "Modeling Relational Events A Case Study on an Open Source Software Project," *Organizational Research Methods*, vol. 17, no. 1, pp. 23–50, 2014.

[36] M. H. Kutner, C. J. Nachtsheim, J. Neter, and W. Li, *Applied linear statistical models*, vol. 5. New York: McGraw-Hill Irwin, 2005.

[37] U. Simonsohn, "Two Lines: A Valid Alternative to the Invalid Testing of U-Shaped Relationships With Quadratic Regressions," *Advances in Methods and Practices in Psychological Science*, vol. 1, no. 4, pp. 538–555, Dec. 2018, doi: 10.1177/2515245918805755.

[38] R. S. Burt and J. Merluzzi, "Network Oscillation," *Academy of Management Discoveries*, vol. 2, no. 4, pp. 368–391, 2016.

[39] G. Cattani, S. Ferriani, and P. D. Allison, "Insiders, Outsiders, and the Struggle for Consecration in Cultural Fields: A Core-Periphery Perspective," *American Sociological Review*, vol. 79, no. 2, pp. 258–281, Apr. 2014, doi: 10.1177/0003122414520960.

[40] R. F. Zammuto, T. L. Griffith, A. Majchrzak, D. J. Dougherty, and S. Faraj, "Information Technology and the Changing Fabric of Organization," *Organization Science*, vol. 18, no. 5, pp. 749–762, Oct. 2007, doi: 10.1287/orsc.1070.0307.

[41] P. Bhattacharya, T. Q. Phan, X. Bai, and E. M. Airoldi, "A Coevolution Model of Network Structure and User Behavior: The Case of Content Generation in Online Social Networks," *Information Systems Research*, Jan. 2019, doi: 10.1287/isre.2018.0790.

[42] G. Ahuja, G. Soda, and A. Zaheer, "The genesis and dynamics of organizational networks," *Organization Science*, vol. 23, no. 2, pp. 434–448, 2012.

[43] K. J. Boudreau and L. B. Jeppesen, "Unpaid crowd complementors: The platform network effect mirage," *Strategic Management Journal*, vol. 36, no. 12, pp. 1761–1777, 2015, doi: 10.1002/smj.2324.

[44] K. J. Boudreau, "Let a Thousand Flowers Bloom? An Early Look at Large Numbers of Software App Developers and Patterns of Innovation," *Organization Science*, vol. 23, no. 5, pp. 1409–1427, 2012, doi: 10.1287/orsc.1110.0678.

[45] Y. Fang and D. Neufeld, "Understanding Sustained Participation in Open Source Software Projects," *Journal of Management Information Systems*, vol. 25, no. 4, pp. 9–50, 2009.

[46] A. Tiwana, B. Konsynski, and A. A. Bush, "Research Commentary—Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental Dynamics," *Information Systems Research*, vol. 21, no. 4, pp. 675–687, Nov. 2010, doi: 10.1287/isre.1100.0323.