

Tilburg University

Essays on diffusion and categories

van Hugten, Joeri

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
van Hugten, J. (2015). *Essays on diffusion and categories*. CentER, Center for Economic Research.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Essays on Diffusion and Categories

Proefschrift ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. E.H.L. Aarts, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op woensdag 2 september 2015 om 14.15 uur door

Joeri Gregorius Wilfridus Judith van Hugten

geboren op 15 augustus 1988 te Ospel

Promotiecommissie

Promotor:

Prof. T. Simons

Copromotor:

Dr. J.G. Kuilman

Overige leden promotiecommissie:

Prof. A. van Witteloostuijn

Prof. O. Sorenson

Dr. Ö. Koçak

TABLE OF CONTENTS

Introductory Chapter	5
How does a concept suggest a premise?.....	6
Plausibility	8
Argument structure	11
The intersection between theory and concept	13
How do my papers fit with these ideas?	15
References	17
Chapter 2	18
Imitative or Independent Market Entry? Foreign Banks in Tokyo (1907–2002), Shanghai (1847–2002), and Hong Kong (1845–2002)	18
Abstract	18
Introduction	19
Theoretical Background.....	22
Theoretical Framework.....	24
Hypotheses	27
Empirical Setting.....	32
Data and Method	35
Results.....	39
Discussion.....	47
References	52
Chapter 3	56
Is it a Bird? Is it a Robin? Imitation when Potential Categories Vary in Level of Detail .	56
Abstract	56
Introduction	57
The Role of Categories in Imitation.....	61
Informativeness and Distinctiveness.....	65
Hypotheses	68
Data and Method	73
Results.....	78

Robustness checks	82
Discussion.....	85
Conclusion.....	92
References	92
Chapter 4	97
The Heterogeneous Diffusion of Money	97
Abstract.....	97
Introduction	97
Applying the Diffusion Model	99
Conceptualizing the phenomenon in diffusion terms	99
Crossing boundary conditions.....	102
Hypotheses	106
Data and Method.....	111
Measures.....	113
Results.....	116
Discussion.....	118
Robustness checks	119
Implications.....	123
References	126
Closing Chapter.....	129
Short summary of the results.....	129
Insights emergent from the interactions between the papers	131

Introductory Chapter

In this chapter I describe my motivation for asking the questions that I discuss and attempt to answer in the papers. I also describe a perspective on research that helps the reader read, evaluate, and appreciate the choices in the papers. I start with specifying what I think makes a contribution, which is subjective enough for sustained debate (e.g. the many editorials with each their own nuance). If I were a journal editor, what would my editorial say about which papers I would accept?

To me a contribution is adding a new (to the target theory) premise to a theory's toolbox. Premises are usually statements like 'concept A positively affects concept B', or 'A moderates a relationship between B and C'. A premise is 'added' in three steps: 1) interpreting concept A in a new way, 2) the interpretation suggests that A is associated with B, and 3) an empirical test supports a causal correlation between A and B. For example, the concept of 'firm' exists, and Coase (1937) interprets it as being an alternative to 'market' in a dimension called 'governance mode of a transaction'. Then, this interpretation suggests that 'firm' should correlate with 'market failure'. Then, researchers gather data on 'vertical integration', as a proxy for firm, and see if it correlates with 'partner-specific investments', a proxy for market failure.

A phenomenon-driven formulation of this kind of contribution is 'using a new (to the target phenomenon) metaphor'. The new metaphor then suggests a new premise. The premise 'concept A leads to concept B' may suggest multiple hypotheses if multiple variables or measures are related to one concept. Adding a new variable or measure is a weak contribution. In contrast, if previous research finds a correlation between two measures, then explaining this correlation in terms of new concepts is a contribution (see also Bacharach (1989) on new propositions vs. new

hypotheses). A new metaphor is often an application of an existing theory to a phenomenon where it had not been applied before. Otherwise, a new metaphor is the start of an entirely new theory.

Some premises are more useful to add than others. One dimension of usefulness is how much doubt there is about the accuracy of a premise. If a premise is highly doubtful, it is more useful to add (i.e. find evidence for). The less doubtful a premise is, the more likely it is that the concept would already suggest this premise, so the degree to which the premise is ‘newly added’ is small. The other important dimension of usefulness is how central a premise is to the theory. One way to gauge the centrality of a premise is the consistency with which it is suggested by concepts in the theory. To compare, premises with low centrality are often context-specific moderators. One analogy is that premises associated with a concept are like indicators of a construct and the most central premises are those with the highest Cronbach’s alpha.

How does a concept suggest a premise?

In my explanation I used the phrase ‘a concept suggests a premise’ a couple of times. This phrase emphasizes that the plausibility of premises such as ‘A leads to B’ is mostly a function of our understanding of A (or B). We find it plausible that A leads to B because B is included in A’s definition, or A’s connotations, or the connotations of words used in A’s definition or words that describe characteristics of A (Quine, 1951). I also find it helpful to think of consequences of A as those things that would have been defining characteristics of A were it not for moderating effects that separate A from its consequences under certain conditions. In other words, the line between defining characteristics of a concept and consequences of that concept is that there are no conditions under which the association between a concept and its characteristics is prevented.

Which premises a concept suggests depends on the dimension and unit of analysis that constitute its definition. A new interpretation of a concept can lead it to suggest new premises. A

new interpretation of a concept involves changing the dimension or even unit of analysis involved in the definition of a concept. This involves changing the definition and connotations of A, which in turn changes the premise that we understand to be associated with A. For example, imagine if the definition of firm is ‘a building with a name and a telephone number’. This definition specifies firms in three dimensions and the unit of analysis is a physical location (or a more abstract word for it, such as ‘site’). Alternatives to firms are ‘named things with a phone number which are not buildings’, such as people, or ‘buildings that cannot be called’, such as statues, or ‘buildings with a telephone number that do not have names’ such as houses. With this definition, the premise ‘firms are correlated with market failure’ seems implausible, or indeed completely out of the blue and unrelated to firms.

A building connotes permanence and high upfront investment. The premise ‘the likelihood that a location has a firm is correlated with long-term (rather than short-term) goals of activities at that location’ seems plausible given those connotations. That plausibility is also because of connotations of the concept of firm not in the definition, such as ‘people related to firms engage in activities related to firm goals’. In contrast, imagine we change the definition of firm to ‘a transaction’s governance mode that is based on fiat and flat incentives’. Alternatives to firms are now ‘governance modes based on bargaining and prices, and strong incentives’, such as markets. This shift in definition, inspires us to think of new premises such as ‘firms are correlated with the degree of market failure in a transaction’, and the shift simultaneously makes such premises seem plausible.

Plausibility

In the previous section, I evaluate premises in terms of their plausibility. The following dialogue between a paper presenter and a listener considers why plausibility is a useful basis for evaluation.

P: “Does anyone have an alternative explanation?”

L: “My purposefully silly alternative explanation is that instead of your mechanism, X leads to rainfall and rainfall leads to Y”

P: “That makes no sense as an alternative explanation!”

L: “Why do you think so?”

P: “My explanation is way more logical”

L: “I think both our explanations were of the type A leads to B and B leads to C therefore A leads to C. Do you see how our explanations have the same logical structure? They must be equally logical.”

P: “But X is completely unrelated to rainfall! Why would X have anything to do with rainfall.”

L: “I could say it’s because X leads to cookies and cookies lead to rainfall.

P: “The logics may be valid but the argument is just not sound because it uses false premises.”

L: “Exactly, but how do good researchers like us know which premises are false and which are true? For example, how do we know that the premises you use for your hypothesis are more true than my premise that X leads to rainfall?”

P: “My premises have reasons and yours don’t.”

L: “No, my premise that X leads to rainfall has the reason that X leads to cookies and cookies lead to rainfall. You could question that reason but then I could just continue giving more silly reasons.”

P: “Ok. Then it’s because my premises have empirical support.”

L: “Let’s think about what it means for a premise to have empirical support. For example, at the end of your paper you claim that your theory was supported by the data. On what grounds do you conclude this?”

P: “Because the data support the hypothesis that I deduced from the theory”

L: “But the data also support the hypothesis that I deduced from my rainfall theory, right?”

P: “Yes, that’s what an alternative explanation does”

L: “So, how can you say that your premises are more true than mine when the premises of both of us have empirical support? From the results of your own study even. The logic of my argument is the same as yours, and the empirical support is also the same.”

P: “Given my background knowledge, my reasoning seems more plausible”

L: “Yes, so instead of saying, there is an alternative explanation, we want to know whether there exists a *good* alternative explanation, and we find that the logic of the alternative explanation is an unfruitful way to define good enough, and the empirical support for the alternative explanation is also by its very nature not a useful way to go in order to determine whether an alternative explanation is threatening enough. Instead we rely on a notion of plausibility.”

P: “But at least you must grant that the research showed a significant correlation, so the research did tell us that the premise X increases Y is more true than the premise the X decreases Y.”

L: “Except that X may in fact decrease Y but this was masked by a positive common cause. Our ability to exclude all common causes depends again on the plausibility that we give to potential common causes. A plausibility which is unrelated to logic and evidence.”

P: “Plausibility may be unrelated to logic and evidence but it is not given arbitrarily. It is based on our background knowledge about the phenomenon.”

Surely we must decide that rainfall theory is not preferred. But on what grounds can we deny it? What does the presenter’s last reply mean? Three criteria for judging a theory are accuracy, simplicity, and generalizability. Is rainfall theory less accurate than other explanations of the result? No, it is specifically designed it to be accurate. Is rainfall theory less simple? What does simplicity mean? If it means that fewer premises are needed to justify a prediction, that means little because all arguments have infinitely many implicit premises. Can we generalize rainfall

theory to fewer alternative settings than our preferred explanation? No, everything can be generalized everywhere. It may just lead to inaccurate implications, but that inaccuracy can be fixed by adding contingencies that change the implications of the concepts in a way that fits the data until accurate generalization is achieved (a point also made by Hannan, Pólos, and Carroll (2007)). For example, in 1845, the orbit of the planet Uranus around the sun was observed. The orbit was not as hypothesized by Newton's theory of gravity. Instead of concluding the theory was false, astronomers questioned the contingency that there was no planet beyond Uranus. With this new contingency, Newton's theory could make accurate predictions about Uranus' orbit. Thus, the astronomers assumed, out of thin air, the existence of an entire planet to make their theory accurate. In 1846, Neptune was indeed discovered.

The three criteria seem unable to provide the grounds to deny rainfall theory. However, a decisive blow to rainfall theory comes from the interaction of the three criteria. Adding contingencies for accuracy surely makes rainfall theory less simple. Thus, *if we generalize rainfall theory, more contingencies are needed to make it accurate than if we were to generalize our preferred theory*. Thus, our preferred theory is preferred because its concepts have the same implications across varying contexts that we want to generalize to. In other words, the concepts' meaning is consistent across contexts (where the word 'meaning' covers a concept's definition, its connotations, its implications, and the connotations of words used in its definition or words that describe characteristics).

Those associations of a concept that are consistent across contexts are the premises with the highest centrality. Thus, a theory is more plausible if it uses the most central associations of its concepts.

Argument structure

The view that premises are suggested by concepts, and that plausibility is the criteria to evaluate explanations, has implications for the structure with which arguments are presented. Each of my papers starts with some concepts and derives hypotheses from these concepts based on which premises they suggest. Suggested premises vary in how doubtful they are, or in other words, in the degree to which they are suggested, or in other words, in the degree to which they rely on situation-specific contingencies. All non-tautological premises are somewhat doubtful. For doubtful premises, readers often ask for arguments that support the premise, or in other words, readers ask for an explicit description of the situation necessary for the premise to be true. I see three ways in which we can use arguments to support a doubtful premise. One view is that arguments guide readers step by step from independent variable to dependent variable via a chain of premises about intermediate mechanisms. In this view, good arguments are like detailed and concrete descriptions of the process by which the independent variable leads to the dependent variable. The smaller the steps, the better; which means, the more steps, the better. I use this form of argument in my discussion of a Spanish shoe retailer who imitates his competitor.

The second view uses analogy. Instead of the two concepts in the premise to be supported, analogies use an existing plausible association between two other but similar concepts (i.e. analogues). If these analogues are plausibly associated, then the premise gains plausibility to the extent of the similarity between the analogues and the premise's concepts. I use analogy only in the initial stage of the research when reinterpreting a concept. However, such a fundamental analogy carries on throughout the paper such that my hypotheses are analogous to old hypotheses. For example, my hypothesis that 'receiving money leads to spending money' is analogous to the

hypothesis in the innovation literature that ‘adoption by a related firm leads to adoption by the focal firm’.

The third view is that an argument is a set of two premises that logically justify a conclusion. Initially, the main premise to be added is the conclusion to be supported. The two premises that justify this conclusion, are then to be justified themselves. They are treated as new conclusions for which a new set of two premises must be invoked. Thus, for each additional justification, two additional yet unjustified statements are added. Therefore, this view finds that the fewer steps, the fewer unjustified statements, the better.

So, where can we draw the line to stop justifying further and further? Or indeed, where do we draw the line where justifications are needed to begin with (e.g. extreme empiricists believe that the best that scientists can do is to compile a collection of correlations)? The ideal case is to justify until you use only premises that are justified by empirical data (which is an alternative to justification via another layer of premises). A softer version of this ideal case is to stop justifying when you use only premises that are plausible. This third view is my default way of using arguments to support premises. For example, I justify the conclusion ‘there are more organizations that react to events at an average speed than there are organizations with a fast or slow reaction speed’ by using empirical findings on founding speeds as one premise and ‘an organization’s founding speed relative to other organizations is the same as its reaction speed relative to other organizations’ as the other premise.

Often, it is tempting to follow the first view and expand an argument by using more and more supporting premises that are more and more detailed and concrete. However, such premises are often arbitrary and ad-hoc in order to support a target conclusion. In addition, many other supporting premises could have been used instead with the same result. My starting point is that

the premises I use are the contribution I make. This also means I should take all premises I use as seriously as my contribution. It is misleading to add such premises to the theory's toolbox. They would be *target-conclusion-specific* associations of the concepts. Thus, if I were to add such premises to the concept, then its meaning would be inconsistent across target conclusions. Consistent concepts are preferred. In addition, they are non-central premises, so adding them is a weak contribution. Therefore, I should be hesitant in using such premises.

Without target-conclusion-specific premises, the reader is left to evaluate the plausibility of the doubtful premise as is. If those premises help such evaluation, the reader is free to imagine their preferred or non-preferred supporting premise as one of many other implicit premises. At the same time, the reader can imagine supporting premises that would prevent the supporting argument's logical validity.

A counterpoint to such hesitance is that the ideal case is to describe multiple potential premises; some that fit the target conclusion and some that do not. I also do this at times. For example, in the hypothesis section of the paper that compares common cause and imitation I discuss which premises could result in a different pattern, and how implausible they are. And in the paper on the diffusion of money, I explicitly discuss and reject some context-specific premises in prior literature about the effect of below aspiration performance. However, one drawback is that such descriptions impede readers from following the flow of the main argument. A second drawback is that such descriptions shift the critical attention of readers from the evaluation of an important premise to the evaluation of the supporting premises.

The intersection between theory and concept

I sometimes say that premises are added to theories and other times that premises are added to concepts. A theory is a collection of snapshots of one interpretation of multiple concepts. A

concept can appear in multiple theories, but defined in different dimensions and at different units of analysis. In addition, different theories sometimes use different words as labels of the same concept to emphasize the theories' difference in interpretation. A new premise can be added to a concept if a new interpretation of the concept suggests a premise that no other interpretation of that concept suggested.

The centrality of a premise is defined per concept. Central premises are those very consistent meanings of a concept. In a continuum from defining characteristic to situation-specific implication of a concept, the consistent meanings are those closest to being a defining characteristic. Another way to gauge centrality is to think of the consistency of a premise across different theories' interpretations of the concept. For example, imagine that 'ties' and 'relationships' are different theories' words for their unique interpretation the same concept. Then, the associations of the concept that are more similar between the two theories are the more central associations. Or, the two definitions of firms: 'a building with a name and a phone' and 'a transaction's governance mode that is based on fiat and flat incentives'. In both definitions, firms still connote activities by people. The frequency with which a premise is suggested across all theories is related to how close the premise is to being a defining characteristic of the concept as opposed to a situation-specific implication.

The doubtfulness of a premise is defined per theory. Because each theory may define concepts differently, the degree to which a concept suggests a premise may differ between theories. Centrality and doubtfulness are both defined by the degree to which a premise is suggested. Because doubtfulness is defined per theory and centrality per concept, it is possible to add a doubtful yet central premise to a theory by importing a premise from another theory that uses the same concept. Only in combination with the applicability of a theory, is doubtfulness, in that sense,

associated with the likelihood of finding empirical support for a premise. For example, a premise with low doubtfulness may not receive empirical support if the theory does not apply in the empirical setting.

I discussed before how a theory is plausible to the extent that the meaning of its concepts is consistent. The goal of research is finding central premises of concepts, and discovering which premises are not so central to a concept. A contribution is a contribution because it contributes to this goal. The more central premises we know of a concept, the more patterns in the world that that concept helps understand. Theories using such a concept will be simple, generalizable, and accurate, because one concept suggests many premises that are not situation-specific, and are in line with empirical evidence.

How do my papers fit with these ideas?

In my first paper I interpret the concepts of imitation and common cause on the dimension of time. This interpretation suggests new (to the imitation literature) premises, especially about how the correlation between prior entry and the likelihood of new entry varies over time. The concept of time has not been worked out in much detail at all in any literature, so it is difficult to gauge the centrality of these premises. The doubtfulness of the premises is intermediate, as indicated by some papers on imitation that already hinted at possible effects of time. The argument structure is simple. The concept of time suggests the premise that information takes time to lead to actions and the premise that information becomes obsolete. The premise of a unimodal distribution of reaction speed is an empirical finding in previous research. The concepts of imitation and common cause both suggest that information is a mechanism. These premises are arranged in arguments to justify the conclusion: a pattern of the likelihood of new entry over time. This conclusion is in line with our data.

My second paper interprets the concept of categories on a dimension called the level of detail in a taxonomy. This interpretation suggest a whole set of new premises described in the psychology literature on categories that I add to the toolbox of the organizational literature on categories. For example, ‘Detailed categories are more likely to be salient (relative to broader categories in their taxonomy) if they are atypical’. The premises to be added all have empirical support from experiments. The most doubtful premise is that those new premises apply in the context of organizations. My approach to test this premise is to translate hypotheses that have been tested in the psychology literature to hypotheses about interorganizational imitation. The argument structure includes an example about a Spanish shoe retailer. That example follows the first view on arguments in its aim to show many concrete intermediate mechanisms between taxonomy structure and the foreign entry behavior of a firm.

In my third paper I interpret the concept of money as a diffusion material, which is a shift in unit of analysis (from bank account or wallet to a flow). This interpretation adds many new premises to the concept of money. The premises suggested by the concept of diffusion are described in the literature on innovation diffusion. Like in the second paper, I translate these hypotheses to hypotheses about money. The premises underlying the hypotheses that I test in the paper have empirical support in a different context. But this support relied on many context-specific premises. As a result, translation of hypotheses was not as straightforward as in the second paper. For example, I show how the premise ‘below aspiration performance increases susceptibility to diffusion materials’ relies on context-specific supporting arguments. This also means that the paper adds new premises to the concept of diffusion: boundary conditions (that are crossed when moving to the context of money). Thus, my findings show that lower susceptibility

is not such a central implication of below aspiration performance as a reading of the literature on diffusion of innovation might suggest.

References

Bacharach, S. B. 1989. Organizational Theories: Some Criteria for Evaluation. *Academy of Management Review*, 14(4): 496–515.

Coase, R. H. 1937. The Nature of the Firm. *Economica*, 4(16): 386–405.

Hannan, M. T., Pólos, L., & Carroll, G. R. 2007. *Logics of organization theory: audiences, codes, and ecologies*. Princeton University Press.

Quine, W. V. 1951. Main Trends in Recent Philosophy: Two Dogmas of Empiricism. *The Philosophical Review*, 60(1): 20–43.

Chapter 2

Imitative or Independent Market Entry? Foreign Banks in Tokyo (1907–2002), Shanghai (1847–2002), and Hong Kong (1845– 2002)

ABSTRACT

When one firm enters a market, another may soon be likely to enter as well. There are two reasons for this. They may be independently reacting to the same market opportunity, or one may be imitating the other. In this paper, we develop ideas about how the influence of the most recent prior entry on the likelihood of new entry changes over time. If independent reaction underlies that influence, then that change over time should differ from situations where imitation underlies the influence. Then, we show how to use this insight to discriminate between the two mechanisms when observing market entries. We use a piecewise-constant hazard model to analyze more than a century of data on entries by foreign banks into the Tokyo, Shanghai and Hong Kong markets. Our findings highlight the relative importance and temporal boundaries of each mechanism. This allows research to move toward more fine-grained theory building with respect to imitation.

INTRODUCTION

Researchers and analysts typically observe that entries into new markets are followed by entries by other firms (Carroll & Hannan, 2000). Theoretical explanations abound but they usually fall into one of two categories. One is based on the assumption that each firm makes its own independent assessment of market opportunities, and when new opportunities arise in a market (e.g. a deregulation of a market – an opportunity common to all – which makes entry more attractive), firms respond irrespective of one another (Gimeno, Hoskisson, Beal, & Wan, 2005). The other explanation is imitation (Haveman, 1993), which can occur because other entrants are believed to have superior information or because of a desire to maintain competitive parity or limit rivalry (Lieberman & Asaba, 2006; Terlaak & Gong, 2008). Anecdotal evidence exists for each motivation. For instance, Vermeulen (2010) describes a case where a large multinational decided not to enter the Scandinavian market despite the fact that it was presented with systematic intelligence that entering the market would be profitable. However, when a large competitor entered the market, it decided to follow suit.

Yet the two explanations are hard to differentiate in more general theorizing about observed market entries. When we see two consecutive market entries, we are observing these as outcomes of an unobserved process and typically do not know whether imitation or an independent assessment of opportunities led to the market entry decision. However, to make further inferences we often need to focus on one mechanism as a basis for these inferences. For example, in order to predict how entering soon after a competitor affects the performance of an entry, it would be relevant to know whether entry timing relative to that competitor was caused by imitation or independent reaction. If we then focus on imitation we run the risk of making an attribution error. That is, we might well be overstating the importance of competitive imitation in the management

literature if in fact most firms just happen to respond to the same market opportunity independently of one another, and as a result flock together into a new market. After all, as the old adage says: correlation is not causation.

How to address this thorny issue of inference and develop a more fine-grained theory of market entry? In order to develop a theory that describes the relationship between a prior market entry and the likelihood of new entry over time, we build on organizational ecology's work on market entries (Hannan, Pólos, & Carroll, 2007) and the interorganizational imitation literature (Gimeno et al., 2005). Our theory is built on the idea that each mechanism's effect manifests differently over time. Specifically, we take the time elapsed since a given entry as our focal point, and study the likelihood of a new entry occurring. Our key insight is that the independent market entry mechanism predicts a particular pattern in the likelihood of new entry over the period since the immediately prior entry, while the imitation mechanism predicts a different pattern (Belderbos, Olffen, & Zou, 2011; Haunschild, 1993; Lieberman & Asaba, 2006).

First, we develop and test new theory about the various relationships that can exist between the time elapsed since a prior entry and the likelihood of a new entry. By doing so, we offer a new vantage on market entry decisions, one which explicitly considers the time at which an entry occurs (Mitchell & James, 2001). In this way, we develop theoretically, and test more rigorously, an intuition already used in prior research (e.g. Haunschild, 1993).

Second, large-scale empirical research typically cannot distinguish well between the market opportunity and imitation motivations: it can only observe that entries in one period affect entries in a subsequent period (Carroll & Hannan, 2000). Disentangling the effects of imitation and market opportunity allows us to assess the relative importance and boundary conditions of

each motivation, leading to more informed theorizing on market entry decisions and a reduced likelihood of attribution bias.

Third, studies of organizer ecologies (Kuilman, Vermeulen, & Li, 2009; Lomi, Larsen, & Freeman, 2005; Sørensen & Sorenson, 2003) have examined the lag between the perception of an opportunity and market entry, in particular how the duration of such lags affect the likelihood of actual entry, organizational survival, and – more broadly – how industries evolve (Lomi et al., 2005; Ruef, 2006). This line of research however has not yet been able to isolate any theoretical mechanism responsible for such lag structures. In this study, we develop theory that specifies such mechanisms.

A fourth contribution is methodological. A common empirical model in this literature relates the likelihood of market entry at time t to the number of entries by competitors, but it is unclear whether the competitors' entries should be measured at time t (Lu, 2002), or up to time t (Guillén, 2003), or lagged one year, or even lagged three years (Delacroix & Carroll, 1983). Constructing an empirical model requires specifying a time, but there has been no theoretical guidance as to the best specification. Our theory and results can offer such guidance. The importance of theoretical and empirical focus on time is a crucial step towards more precision and rigor in theory and empirical testing, especially if a wrong specification of lags not only leads to underestimates or overestimates of an effect, but also leads to capturing an entirely different theoretical mechanism (e.g. market opportunity effects instead of imitation) (Mitchell & James, 2001).

We thus look at two types of motives for market entry and focus on the speed at which various organizations enter a foreign market. In the next section we describe a framework for understanding why the effects of different mechanisms manifest at different times. Then, for each

mechanism, we derive what the inter-arrival time pattern over a population should look like if that mechanism were true. Then, we see if our expectations are matched by data on entries into foreign banking populations in Hong Kong (from 1845 to 2002), Shanghai (1847–2002), and Tokyo (1907–2002).

THEORETICAL BACKGROUND

A consistent finding of past research is the positive relationship between the number of past market entries and the likelihood of further entries (Gimeno et al., 2005; Haunschild & Miner, 1997; Haveman, 1993; Henisz & Delios, 2001; Lieberman & Asaba, 2006). We will first discuss the details of two mechanisms underlying that relationship. Then we discuss how each mechanism's influence varies depending on the time in between the past market entry and the potential new entry.

We distinguish two mechanisms underlying the relationship. Each mechanism suggests a different time dependence. The first is referred to as an independent market entry mechanism, which involves each aspiring entrant making an entry decision based on opportunities that exist in the market and one's own capability to generate rents from such opportunities, irrespective of the (potential) behavior of likely competitors (Dunning, 1981; Buckley and Casson, 1998). For instance, Dunning's (1981) seminal framework emphasizes the ownership, location, and internalization advantages that would be considered in a market entry decision. According to this perspective, the likelihood of a market entry therefore correlates with the emergence of seemingly profitable business opportunities. But these opportunities also increase the likelihood of market entries by other organizations. Thus, one may well observe a correlation between any entry and the likelihood of subsequent entry because both are correlated with the emergence of the same business opportunities (Baron, 2006; Manski, 1993; Sørensen & Sorenson, 2003). For example,

the end of the Second World War revived for banks the promise of entry into East-Asian markets, and indeed, a cluster of new entries occurred in the period 1945–1950.

The second explanation emphasizes imitation (Guillén, 2003; Haveman, 1993; Lieberman & Asaba, 2006), which involves a market entry signaling the possibility or appropriateness of market entry to other potential entrants, which makes them more likely to enter. This could happen for various reasons (Lieberman & Asaba, 2006). An earlier entrant may be believed to have superior information about opportunities in the market. This assumes that it is difficult and costly for organizations to accurately predict the outcomes of their decisions. Therefore, they tend to rely on easily observable information, such as the decisions of their competitors, to make their prediction. Or such imitative moves may have to do with relative competitive positions: a fear of losing one's competitive advantage, a desire to achieve competitive parity, or perhaps to limit rivalry. Regardless of underlying reasons, the observable result is also that organizations flock together into markets, yet in this perspective an early entry actually directly causes those entries occurring subsequently (Baum, Li, & Usher, 2000; Levitt & March, 1988). For example, US banks may have been hesitant about entering Tokyo in light of the stressed international relations between the US and Japan after the war. The entry by the National City Bank of New York in 1946 served as a powerful signal of appropriateness, encouraging a cluster of new entries.

So both independent and imitative market entry could explain the cluster of entries in Tokyo in the years 1946–1950. This situation of alternative explanations is theoretically problematic because empirical evidence of a correlation between prior entry and new entry cannot facilitate inferences about the value of either theory in explaining the rate of change in the number of foreign organizations in a market (Gimeno et al., 2005; Lieberman & Asaba, 2006).

THEORETICAL FRAMEWORK

We see an overlap between imitation and independent reaction in terms of making information available to a potential entrant. A key difference between independent and imitative market entry is the source of information that an entry decision is based on. We focus on events as the sources of information. For imitation, a prior entry by a competitor is the most relevant event. For independent reaction, events such as demand shocks, regulatory shifts, or perceived market opportunities are relevant events. Does this mean that when either type of event occurs, that the likelihood of subsequent entries is immediately higher than before and that it stays higher forever? Because information is the key driver for both imitation and independent reaction, we base our answer to this question on two principles about the behavior of information over time.

The first principle is that organizations need time to transform information into action. The relevant information must first reach the organization; the organization then seeks additional information to verify it and deepen the context; it uses the information to make a decision; it then must gather the resources needed to execute that decision (Lomi et al., 2005). Organizations transform information into decisions under uncertainty (Levitt & March, 1988). If information availability is low, this uncertainty is higher; the organization is less aware of what goal to set, and is less able to assess, in satisfactory detail, the costs, risks, and challenges associated with achieving that goal. Organizations deciding under higher uncertainty are less likely to make the decision (Guillén, 2002; Martin, Swaminathan, & Mitchell, 1998). Accordingly, as organizations receive more information, they are more likely to attempt a market entry.

The pre-entry ecology literature (Kuilman et al., 2009; Lomi et al., 2005; Sørensen & Sorenson, 2003) views market entry as a transition by organizers from the organizing stage to the operational stage. The organizing stage is a stage in which entrepreneurs gather resources in an

attempt to found an organization. Organizers are transforming information from the environment into the potential organization's structure. If information availability is low, the new organization's structure is more likely to be ill-formed. For instance, the organizer is less aware of the most cutting-edge inventory database systems. An organization with an ill-formed structure is less likely to succeed in transitioning (Kuilman et al., 2009). So according to the pre-entry ecology literature (Kuilman et al., 2009; Lomi et al., 2005; Sørensen & Sorenson, 2003), given a high level of uncertainty, organizers will attempt to enter, but they are less likely to succeed in entering without information. In sum, more entries will be observed if more information is available to potential entrants. In line with this, Ruef (2006) shows that entrepreneurs aiming to start a medical school were most likely to transition to the operational stage two years after the entrepreneur started the organizing stage.

The second principle is that information becomes obsolete over time. Events become less impelling over time as they become a fact of life. In addition, the evaluation of an event changes over time as more and more other changes occur that moderate the value of an event. In line with this, Fier and Woywode (1994) find that the fall of the Berlin Wall and related government programs increased foundings of new businesses in east Germany, but two years later the founding rate returned to normal even though the government programs were still in full operation and the Berlin Wall was still fallen. Thus, information about these events becomes obsolete. More obsolete information should lead to a smaller decrease in the uncertainty of decision-making (Levitt & March, 1988), causing a smaller increase in attempted entries. In addition, more obsolete information should lead to the potential organization's structure being less aligned with the environment (Kuilman & Li, 2006), making transitions to the operational stage less likely to be completed. Thus, the effect of information on market entry slowly becomes negligible as it

becomes obsolete. This obsolescence increases with time, therefore the effect of events on market entry decreases over time. For example, a prior entry reveals some private information of the entrant. Because this private information must have been positive, new entries are likely to follow. However, as more changes occur over time, new potential entrants becomes less certain about whether the basis of the private information has also changed. Therefore, over time, potential new entrants' decisions are influenced less by the prior entry.

These two principles combine to define the speed of an organization's reaction to an event. We need to specify that transformation time and obsolescence do not cancel out each other's effect. We assume that some time after the event, information is greatly more transformed than soon after the event, and that it is only slightly more obsolete. Long after the event, information is only slightly more transformed than some time after the event, but the information is far more obsolete. Thus, we should observe an increase in the likelihood of entry as we move from soon after the event to 'some time' after, as the positive transformation effect dominates the obsolescence effect. We should observe a decrease in the likelihood of entry as we move from 'some time' after the event to 'long after', as the negative obsolescence effect dominates the transformation effect.

Next to this firm-level argument, there is also a parallel population-level argument. We assume a unimodal distribution of reaction speed among organizations in a population, with the mode at the average level (like a bell-curve). This means that the population consists of more organizations with an average reaction speed than organizations with faster-than-average or slower-than-average reaction speed. This assumption is grounded in existing empirical work on start-ups of U.S. hospitals (where the mode is 2 years) (Ruef, 2006: 44), and foreign banks organizing in Shanghai (Kuilman & Li, 2006). This distribution implies that soon after, and long after, an event, the likelihood of entry is low, because there are not many organizations that react

quickly or very slowly. The likelihood of entry is high some intermediate time after an event because most organizations would react at intermediate speeds.

HYPOTHESES

How can these arguments be used to predict how the likelihood of an entry varies with the time lapsed since the immediately prior entry? In ecological studies of organizations (Carroll & Hannan, 2000), time since the most recent prior entry is called the inter-arrival time. Answering this question would entail mapping the two types of events that organizations can respond to onto this inter-arrival period. These two types of events are: 1) market events such as emergence of market opportunities, profit announcements, regulatory shifts that lead to correlated but independent entries, and 2) an entry itself that may influence potential subsequent entrants (via imitation).

These two types of events typically have a particular order in time. The entry to be imitated (i.e. the second type of event) is caused by an earlier event (i.e. the first type of event). The first type of event is most likely to cause the second type of event at that point in time where its information is most completely transformed but not yet obsolete. The second event similarly causes imitative entries at a particular point in time. Thus, events and entries are anchored in time to the prior entry (i.e. the event of the second type). Panel (b) of Figure 1 illustrates a situation where a market opportunity occurs followed by independent market entries, followed by imitative market entries.

In the case of imitative market entry, a pattern of entry likelihood is conditioned on a prior entry. So, at the beginning of the inter-arrival time, the likelihood of entry is low as information availability about that prior entry is still low. But as information become more completely transformed, the entry likelihood increases. When the inter-arrival time becomes longer, and

information starts to become obsolete, the entry likelihood will decline again. In other words, the inter-arrival time initially has a positive effect on the likelihood of entry, but as the inter-arrival time takes on high values, further increases in inter-arrival time have a negative effect.

Hypothesis 1a (Imitative market entry): There is an inverted-u shaped relationship between the inter-arrival time and the likelihood of entry.

In the case of independent market entry however, how the entry likelihood changes over time is quite different. The market opportunity that led to the prior entry in the first place should occur before the moment of the prior entry. Like any entry, the prior entry is most likely to occur at the peak of the previously described increase-decrease pattern. The new entry is a reaction to the same event, and therefore the likelihood of the new entry should follow the same increase-decrease pattern. Therefore, if we average across all prior entries, the likelihood of a new entry should be highest at the same moment that the likelihood of a prior entry is highest. So, at the beginning of the inter-arrival time (i.e. when the prior entry occurs), the likelihood of observing another entry is at its highest point but then starts to decrease, and this decrease continues as the inter-arrival time becomes longer and information about the market opportunity starts to become obsolete.

Hypothesis 1b (Independent market entry): There is a negative relationship between the inter-arrival time and the likelihood of entry.

To stress the distinctness between the predictions that each mechanism makes, consider the low plausibility of the assumptions that would be necessary for each mechanism to predict the other mechanism's pattern. Imitation could predict a negative pattern of the entry likelihood over

the inter-arrival time only if we make the implausible assumption that firms can receive information about a competitor's entry into a foreign market, make a decision based on that information, and start up a subsidiary in a foreign market, all within one month. Independent entry could predict an inverted-u pattern only if, after the likelihood of reaction to an event becomes high enough to result in a prior entry, it then instantly drops to a low level, and then gradually increases over time, to finally decrease over time.

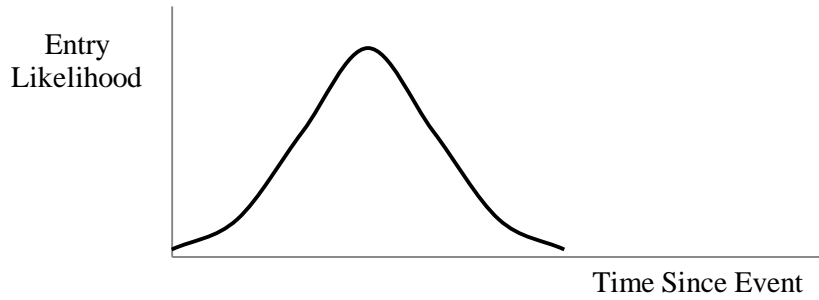
We do not see H1a and H1b as necessarily competing or contradicting hypotheses. It is possible that both mechanisms are at play simultaneously but each dominates a particular range of the inter-arrival time. If this is the case we could observe an initial decrease as predicted by H1b followed by the increase and subsequent decrease as predicted by H1a. Panels (c) and (d) of Figure 1 depict the separate effects of imitation (H1a) and independent market entry (H1b), respectively. Panel (e) of Figure 1 depicts the entry likelihood after the prior entry when the effects are combined. The initial decreasing likelihood of entry as time passes results from the decreasing importance of the underlying market opportunity. Thereafter, the entry likelihood should increase over time because imitation becomes more important. Eventually the likelihood decreases because the likelihood of imitation decreases in the second part of the imitation pattern.

Hypothesis 1c (Independent and imitative market entry): The likelihood of entry will first decline, then increase, and then decline with the duration of the inter-arrival time.

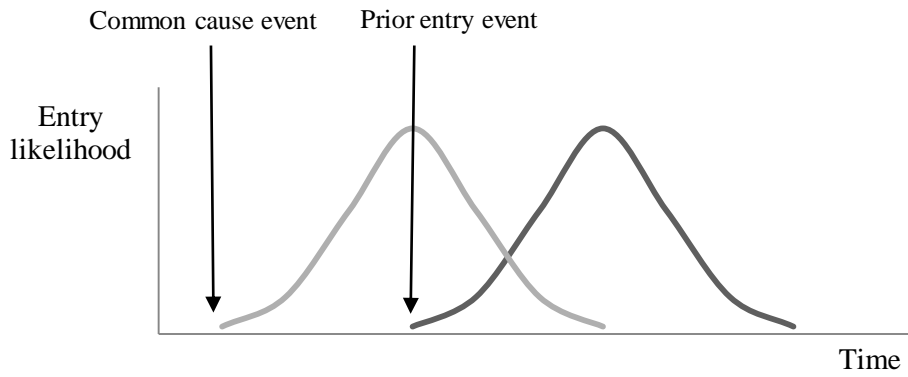
FIGURE 1

Steps Toward Entry Likelihood over the Inter-arrival Time

a) Likelihood of entry after an event



b) Likelihood of entry around the prior entry



c) Likelihood of entry over the time since prior entry if only imitation is true

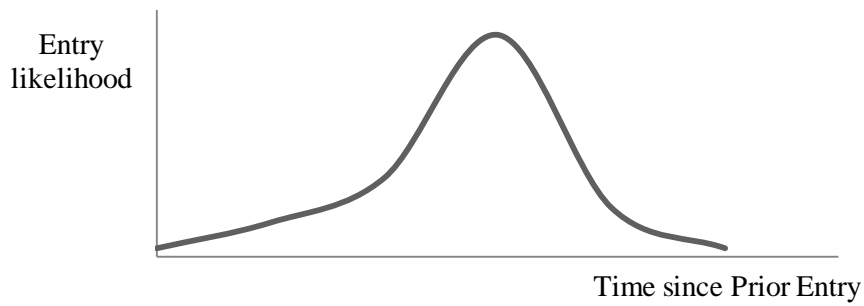
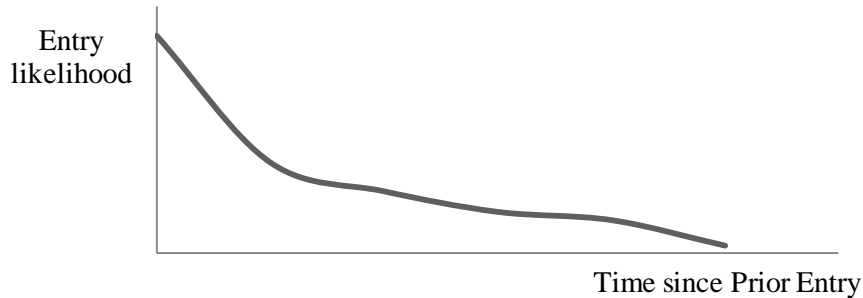
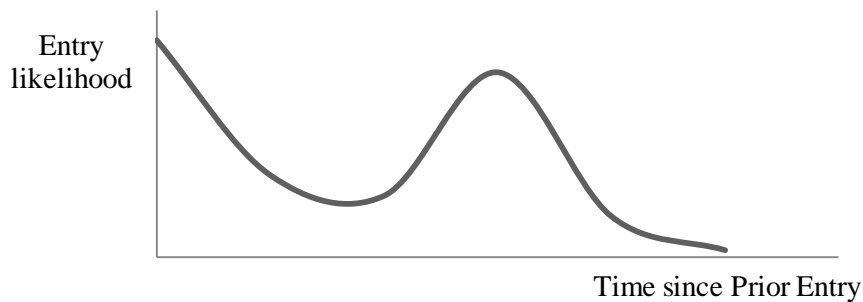


FIGURE 1 (continued)

d) Likelihood of entry over the time since prior entry if only common cause is true



e) Likelihood of entry over the time since prior entry if both are true



Consider how implausible the assumptions must be for each mechanism to cause this pattern by itself. Independent reactions could cause this pattern if there are two groups of organizations; one group with fast reaction speed, and one with slow reaction speed. In addition, there must be fewer organizations with average reaction speed than with fast or slow reaction speed. This is implausible in light of previous empirical work (Kuilman & Li, 2006; Ruef, 2006). For imitation to cause this pattern, there must again be at least some implausibly fast imitators.

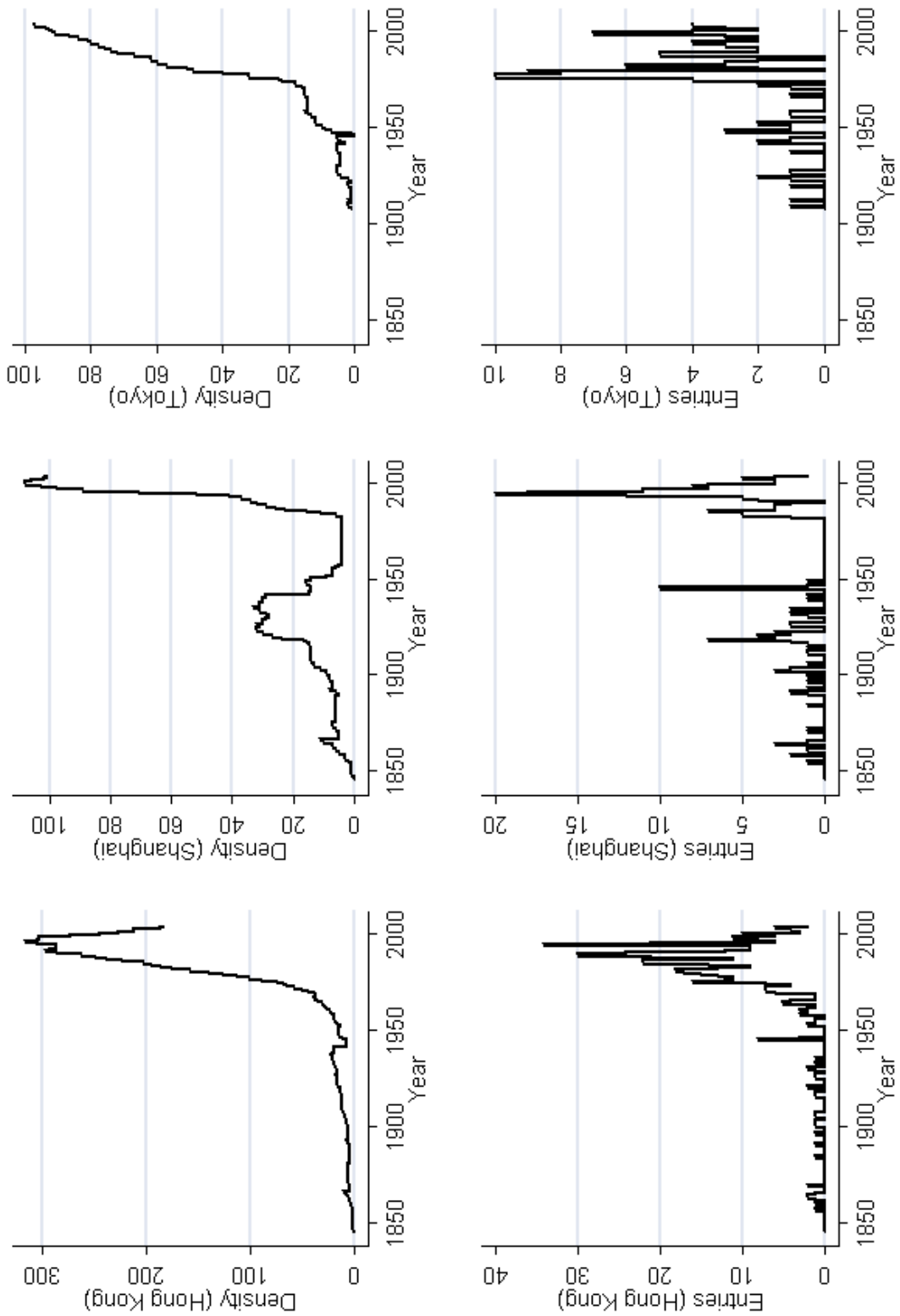
EMPIRICAL SETTING

We test our hypotheses using data on the entries of foreign banks into the banking markets in Shanghai (from 1847 to 2002), Hong Kong (1845–2002) and Tokyo (1907–2002). Figure 2 shows the historical trajectories of the three populations of foreign banks, describing their total numbers (densities) and annual entries.

Hong Kong was among the earliest cities in East–Asia to experience the founding of foreign banks. The International Banking Corporation set up an office there in 1845, following the Opium War which had ended in 1842. A gradual increase in the number of foreign banks followed, initially banks from British territories, but later joined by French, German, American, Dutch and other banks. Foreign banks also entered other East-Asian cities in that era and indeed, as Figure 3 shows, in the 1920s and the early 1930s the number of foreign banks in Shanghai surpassed the number in Hong Kong. Shanghai in that period became a major center of international trade and finance and played a pivotal role in East Asia. The period was characterized by rapid economic development and became known as Shanghai’s ‘golden age’ (Ji, 2003). Hong Kong was at that time ‘essentially a smaller version of Shanghai’ (Jones, 1992: 407).

In Japan, foreign banks moved into Yokohama first, starting in 1863, but the built-up of a significant population of foreign banks in that port city suffered a major blow with the Great Kanto Earthquake of 1923. Afterward, leading foreign banks concluded it would take time to rebuild their operations in Yokohama and they shifted their business to Tokyo instead.

FIGURE 2
Foreign Bank Densities and Entries in Hong Kong, Shanghai and Tokyo, 1845–2002



Political and economic turmoil affected all three populations of foreign banks in the ensuing period. Shanghai lost its position as East Asia's main financial center in the late 1930s and 1940s for a variety of reasons. A currency crisis in 1935 was among the first indicators of the declining role of foreign banks in Shanghai. The Sino-Japanese War that started in 1937 also troubled the banking business, and with the onset of the Pacific War in late 1941, many foreign banks not only in Shanghai, but also in Hong Kong and Tokyo, ceased operations. (In particular, banks from Allied countries had to withdraw from Japanese-controlled cities during the war.)

After the Second World War, some foreign banks soon returned to Hong Kong, but in the case of Shanghai, China's civil war (1945–1949) together with high inflation limited new business opportunities. In October 1949, China came under Communist Party rule and the new regime was, at best, unfavorable to the presence of the remaining foreign banks in China. Foreign banks came to be seen as agents of western imperialism. By 1956, only four 'quasi-foreign' banks were left in Shanghai—The Hong Kong and Shanghai Bank (today's HSBC); the Bank of East Asia; the Chartered Bank of India, Australia, and China (today's Standard Chartered Bank); and the Overseas Chinese Banking Corporation.

Hong Kong and Tokyo were to some extent beneficiaries of the dismantling of Shanghai, but growth in their foreign banking populations was also facilitated by local economic growth and increased foreign trade. Hong Kong at that time offered several advantages, of which the most important were a relatively stable social and political system, as well as economic freedom. Tokyo became a natural entry point for foreign banks with an interest in the Japanese market. As shown in Figure 3, both cities experienced an accelerating pace of entries.

In the early 1980s China ended its period of economic isolation, and many of foreign banks regained an interest in Shanghai as it became the home of an increasing number of headquarters

for foreign banks' Chinese operations. For instance, Citibank's China headquarters was relocated from Hong Kong to Shanghai in 1993. In 1999, HSBC moved its corporate headquarters for China from Hong Kong to Shanghai. Although in both cases the Hong Kong offices retained their importance as headquarters for the broader Asia-Pacific area, the shifts illustrate a clear reallocation of the attention of foreign banks.

DATA AND METHOD

We conduct three empirical tests to assess the relationship between the time since the prior entry and the likelihood of new entry. Specifically, we investigate the establishment of foreign bank offices in Hong Kong, Shanghai and Tokyo over the entire population history. Data on the entire history is necessary to accurately estimate population-level processes such as legitimation and competition (Carroll & Hannan, 2000), that also influence the duration of the inter-arrival time. Our focus on cities instead of countries increases the likelihood of a link between two consecutive entries. For example, in a city-level analysis, two subsequent entries are more likely to be related than two subsequent entries into different parts of a country.

The primary source of data describing the Tokyo population was the extensive work by Kazuo Tatewaki (2002), extended with additional data from the Bank of Japan. The extraordinary combination of comprehensive and detailed data over an extended time period makes the Japanese banking industry an excellent context for the purposes of our study.

Data on the history of foreign banks in Shanghai were drawn from a variety of publications which differed in their organizational and temporal coverage. Some covered the pre-1949 period and others the post-1978 period, but there is no consistent source that covers all banks in both periods. For the pre-1949 period, two sources were particularly helpful: Tamagna (1942) and

Hong, Wang and Li (2003). The precision of the dates reported in these publications, and the limited extent to which additional banks were found in alternative sources such as Ji (2003) and in various local archives in Shanghai, provided considerable confidence that all the relevant data have been included. For the post-1978 period, The Bankers' Almanac was used to compile a first master list of all foreign banks. Although The Bankers' Almanac is a very comprehensive source, timing entries based on the first listing in The Bankers' Almanac at times proved to yield inaccurate dates. A bank's first listing in The Bankers' Almanac was sometimes delayed by one or two years. For this reason, the full master list was checked against articles in Lexis-Nexis and individual banks' annual reports.

For the Hong Kong data, various local archives were consulted, including those of the Hong Kong Companies Registry and the Hong Kong Monetary Authority (HKMA). The most valuable source proved to be the HKMA's annual reports (since 1993), the reports of the Office of the Commissioner of Banking (1987–1992), and listings such as those published in Hong Kong Banking (1985) and the Far Eastern Economic Review (starting in 1960). Jones (1965a, 1965b) has provided detailed information on the early history of foreign banks in Hong Kong.

Each of the cities was treated as being at risk of receiving a foreign entry each month, after having experienced its very first foreign bank entry. The observed market entries were from then on treated as an inter-arrival process (Carroll & Hannan, 2000: 104–107), meaning that the clock is reset with each new entry. The timing analyzed was thus the interval between subsequent entries (i.e. the inter-arrival time).

The aggregate of all inter-arrival times constitutes the distribution of the likelihood of entry over the inter-arrival time. Our hypotheses are about the shape of this distribution. Event-history

analysis (Tuma & Hannan, 1984) was applied to estimate the likelihood of entry in a given period from all the inter-arrival times. This was formally defined as

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(\text{entry}(t + \Delta t) | \text{no entry}(t))}{\Delta t},$$

which reads as the likelihood that an entry occurred in the time period from t to $t + \Delta t$, provided that no entry occurred at or prior to time t . The distribution of the likelihood of entry over the inter-arrival time is the combination of this likelihood over all periods.

Applying event history methods (Tuma & Hannan, 1984) to the analysis of entries or foundings is a practice that dates back to the earliest empirical studies in the field of organizational ecology (Hannan & Freeman, 1987; Carroll & Hannan, 1989b; Wholey, Christianson, & Sanchez, 1993). For instance, Hannan and Freeman's (1987) study of foundings of labor unions in the U.S. between 1836 and 1985 measured the durations between consecutive founding events and modeled the likelihood that such events occurred as an inter-arrival process using Cox's (1975) proportional hazard model. A similar approach was taken in Carroll and Hannan's (1989) study of newspaper foundings in the U.S., Argentina, and Ireland, and in a study of health maintenance organizations by Wholey, Christiansen and Sanchez (1993). Those early studies adopted Cox models because they did not think a particular form of duration dependence was likely for the processes they studied (e.g. Hannan & Freeman, 1987). Han (1998), in contrast, in his study of domestic banks in Japan, looked more closely at how the rate of founding of new banks varied over the time since a prior founding using piece-wise exponential models. He found that the entry likelihood decreased monotonically over time, and attributed this to imitation and to the emergence of market opportunities, both being strongest shortly after a prior founding. Han (1998) did not, however, clearly differentiate between the two mechanisms. The main advantage of using event

history methods over the now more common aggregate count models (such as Poisson and negative binomial models), is that it allows for a closer examination of the time dependence of entry. We applied piecewise-constant exponential hazard specifications in modeling the entry rate. In a piecewise specification, the likelihood of entry is allowed to vary between time periods, but is constant within each period. The piecewise-constant exponential models had the following general form:

$$\lambda(t)=\exp(\alpha_p + \beta'xt) \quad p=1 \dots P,$$

where α is a constant which was allowed to take different values in different periods p , and $\beta'xt$ is a row vector of coefficients (β) and independent variables (x). The periods were defined such that observations were approximately equally distributed across the periods. The models were estimated using the `stpiece` function of the STATA statistical software package (Sørensen, 1999). The model estimates a baseline hazard for each period. Our hypotheses are about how that hazard varies between subsequent periods.

The advantage of this over parametric specifications is that it does not impose any functional form on the relationship between the inter-arrival duration and the hazard of entry. Instead, it allows the researcher to observe how the baseline hazard rate and the effects of any of its time-varying covariates vary over time. A standard exponential hazard model, by contrast, reports only a constant, baseline hazard rate. Averaging the hazard rate over the entire time window hides useful information by assuming that the intercept (i.e. baseline) is the same over the entire at-risk spell. Non-parametric approaches such as Cox models can also capture complex time-dependence of the baseline, but such models do not estimate a parameter to show us this time-dependence. In contrast, our piecewise model shows considerable variation in the estimated intercept entry hazard across time periods, and thereby improves the explanatory power

significantly (compared to an exponential model) in all three populations (model 1 in Tables 2, 3, and 4, respectively).

To hold market attractiveness constant in the empirical analysis, we included the number of existing banks and the country's GDP divided by the number of existing banks, as control variables. Increases in those variables should capture legitimation, improved infrastructure, and market size per firm, all of which would have made the markets more attractive. But as density dependence theory (Hannan & Carroll, 1992) indicates, when the number of organizations increases beyond some threshold, competition should become more severe, which increases resource scarcity. This makes the market less attractive. The number of organizations was therefore squared to represent the negative effect of competition, and the first order term to capture the positive effect of legitimacy (Hannan & Carroll, 1992). Two historical period dummies were also included representing the Second World War and the period of unfavorable government in China. Table 1 summarizes all variables and shows their correlations.

In our Hong Kong data, the month of entry is unknown for most entries. In this case, we randomly assigned (using a uniform distribution) entrants to a month within their year of entry (on which we have data for all entries).

RESULTS

Model 2 in Tables 2, 3, and 4, shows the results for a model with controls for market attractiveness. In all three cities, the initial decreasing pattern is less steep than in the models without control variables. This makes sense if the control variables capture variation in market opportunities instead of the time periods' baseline.

TABLE 1
Descriptive Statistics and Correlations

a) Tokyo^a

Variable	mean	s.d.	1	2	3	4	5
1. GDP/Density	37488	20895					
2. Density	30.93	34.02	-.41				
3. Density ²	2113	3093	-.34	.98			
4. War	0.06	0.23	.17	-.19	-.17		
5. Period 1949–1980	0.33	0.47	-.24	-.15	-.29	-.17	
6. PEIT ^d	24.02	34.89	.42	-.44	-.38	.31	-.20

b) Shanghai^b

Variable	mean	s.d.	1	2	3	4	5
1. GDP/Density	71337	75889					
2. Density	18.26	23.63	-.34				
3. Density ²	891.96	2506.78	-.17	.94			
4. War	0.05	0.22	-.18	.08	-.02		
5. Period 1949–1980	0.31	0.46	.48	-.36	-.23	-.15	
6. PEIT ^d	43.02	74.53	.49	-.26	-.17	-.01	-.00

c) Hong Kong^c

Variable	mean	s.d.	1	4	5	6	7
1. GDP/Density	205.72	158.98					
4. Density	83.38	127	.51				
5. Density ²	23071.4	47271.9	.40	.98			
6. War	0.04	0.19	.05	-.10	-.10		
7. Period 1949–1980	0.21	0.41	.64	-.11	-.20	-.10	
8. PEIT ^d	30.30	65.46	-.26	-.26	-.21	.03	-.14

^a $n=1221$

^b $n=2502$

^c $n=2170$

^d PEIT is an abbreviation for “Prior Entrant’s Inter-arrival Time”

TABLE 2
Results of Piecewise-constant Hazard Models of Market Entry into Tokyo^{a,b}

Variable	Model 1		Model 2		Model 3	
Time period 1 (0,1]	-1.39	* (0.15)	-3.93	* (0.59)	-3.33	* (0.63)
Time period 2 (1,2]	-1.76	* (0.22)	-4.25	* (0.60)	-3.34	* (0.71)
Time period 3 (2,3]	-1.46	* (0.20)	-3.90	* (0.58)	-3.13	* (0.65)
Time period 4 (3,5]	-2.20	* (0.27)	-4.54	* (0.59)	-3.72	* (0.70)
Time period 5 (5,8]	-1.77	* (0.21)	-3.99	* (0.54)	-3.38	* (0.63)
Time period 6 (8,14]	-3.26	* (0.41)	-5.36	* (0.64)	-4.59	* (0.72)
Time period 7 (14,26]	-2.71	* (0.30)	-4.63	* (0.56)	-3.84	* (0.64)
Time period 8 (26,42]	-3.52	* (0.46)	-5.13	* (0.67)	-5.50	* (0.91)
Time period 9 (42,∞)	-3.53	* (0.34)	-5.23	* (0.62)	-5.08	* (0.56)
GDP/Density			0.00	* (0.00)	0.00	* (0.00)
Density			0.07	* (0.02)	0.05	* (0.02)
Density ²			-0.00	* (0.00)	-0.00	* (0.00)
War			0.08	(0.51)	-0.44	(0.78)
Period 1949–1980			0.23	(0.27)	0.15	(0.27)
PEIT×Tp1 ^c					-0.00	(0.01)
PEIT×Tp2					-0.07	(0.06)
PEIT×Tp3					-0.03	(0.03)
PEIT×Tp4					-0.04	(0.03)
PEIT×Tp5					-0.01	(0.01)
PEIT×Tp6					-0.03	(0.02)
PEIT×Tp7					-0.02	(0.01)
PEIT×Tp8					0.01	(0.01)
PEIT×Tp9					0.00	(0.01)
Log pseudo-likelihood	-206.95		-187.38		-182.94	
Degrees of freedom	9		14		23	

^a Robust standard errors are in parentheses. Tests for coefficient significance are two-tailed and are tested on complete population data.

^b $n=131$ market entries

^c PEIT×Tp1 is an abbreviation for “Prior Entrant’s Inter-arrival Time × Time period 1”

* $p < .01$

TABLE 3
Results of Piecewise-constant Hazard Models of Market Entry into Shanghai^{a,b}

Variable	Model 1		Model 2		Model 3	
Time period 1 (0,1]	-0.96	* (0.09)	-3.07	* (0.28)	-2.57	* (0.31)
Time period 2 (1,2]	-1.40	* (0.16)	-3.26	* (0.28)	-2.80	* (0.36)
Time period 3 (2,3]	-2.12	* (0.28)	-3.76	* (0.36)	-3.55	* (0.37)
Time period 4 (3,5]	-1.76	* (0.17)	-3.35	* (0.26)	-2.42	* (0.33)
Time period 5 (5,7]	-2.26	* (0.28)	-3.68	* (0.32)	-2.73	* (0.41)
Time period 6 (7,16]	-3.13	* (0.27)	-4.27	* (0.29)	-3.75	* (0.39)
Time period 7 (16,28]	-3.54	* (0.32)	-4.58	* (0.35)	-4.35	* (0.36)
Time period 8 (28,40]	-2.89	* (0.27)	-3.78	* (0.29)	-3.02	* (0.29)
Time period 9 (40,∞)	-4.61	* (0.44)	-4.70	* (0.40)	-4.50	* (0.46)
GDP/Density			0.00	* (0.00)	0.00	* (0.00)
Density			0.06	* (0.01)	0.05	* (0.01)
Density ²			-0.00	* (0.00)	-0.00	* (0.00)
War			-1.14	* (0.32)	-0.70	(0.39)
Period 1949–1980			-3.15	* (1.20)	-3.23	* (1.18)
PEIT×Tp1 ^c					-0.03	(0.02)
PEIT×Tp2					-0.02	(0.03)
PEIT×Tp3					0.00	(0.00)
PEIT×Tp4					-0.10	* (0.03)
PEIT×Tp5					-0.07	(0.05)
PEIT×Tp6					-0.01	(0.01)
PEIT×Tp7					-0.00	(0.00)
PEIT×Tp8					-0.02	(0.01)
PEIT×Tp9					-0.00	* (0.00)
Log pseudo-likelihood	-322.01		-276.14		-260.97	
Degrees of freedom	9		14		23	

^a Robust standard errors are in parentheses. Tests for coefficient significance are two-tailed and are tested on complete population data.

^b $n=198$ market entries

^c PEIT×Tp1 is an abbreviation for “Prior Entrant’s Inter-arrival Time × Time period 1”

* $p < .01$

TABLE 4
Results of Piecewise-constant Hazard Models of Market Entry into Hong Kong^{a,b}

Variable	Model 1		Model 2		Model 3	
Time period 1 (0,1]	-0.52	* (0.04)	-2.76	* (0.21)	-1.94	* (0.36)
Time period 2 (1,2]	-0.84	* (0.08)	-2.88	* (0.22)	-2.12	* (0.31)
Time period 3 (2,3]	-1.14	* (0.14)	-2.87	* (0.23)	-2.34	* (0.31)
Time period 4 (3,4]	-1.28	* (0.18)	-2.77	* (0.25)	-2.17	* (0.31)
Time period 5 (4,5]	-2.10	* (0.35)	-3.31	* (0.38)	-2.30	* (0.55)
Time period 6 (5,7]	-1.96	* (0.26)	-2.99	* (0.29)	-2.65	* (0.33)
Time period 7 (7,13]	-2.48	* (0.25)	-3.01	* (0.26)	-3.05	* (0.29)
Time period 8 (13,46]	-3.43	* (0.28)	-3.81	* (0.29)	-3.57	* (0.30)
Time period 9 (46,∞)	-4.16	* (0.33)	-4.29	* (0.34)	-4.21	* (0.50)
GDP/Density			0.00	(0.00)	0.00	(0.00)
Density			0.02	* (0.00)	0.01	* (0.00)
Density ²			-0.00	* (0.00)	-0.00	* (0.00)
War			-0.77	(0.54)	-0.68	(0.50)
Period 1949–1980			0.69	* (0.11)	0.51	* (0.11)
PEIT×Tp1 ^c					-0.19	(0.15)
PEIT×Tp2					-0.11	(0.06)
PEIT×Tp3					-0.02	(0.04)
PEIT×Tp4					-0.04	(0.03)
PEIT×Tp5					-0.12	(0.08)
PEIT×Tp6					-0.00	(0.01)
PEIT×Tp7					0.00	(0.00)
PEIT×Tp8					-0.01	(0.01)
PEIT×Tp9					-0.00	(0.00)
Log pseudo-likelihood			-600.19		-586.06	
Degrees of freedom	9		14		23	

^a Robust standard errors are in parentheses. Tests for coefficient significance are two-tailed and are tested on complete population data.

^b $n=500$ market entries

^c PEIT×Tp1 is an abbreviation for “Prior Entrant’s Inter-arrival Time × Time period 1”

* $p < .01$

Next, consider the situation where entrant A imitates entrant B, but a third entry (C) occurs in between A and B’s entries. In our empirical setup, the confounding entry (C) is now defined as the prior entry for A. As a result, A’s inter-arrival time is shorter than the time between A’s entry

and the entry that it actually imitates (i.e. the entry of B). Our analysis then falsely concludes that A is not an imitator. If such situations are common, the influence of common causes will be overestimated and that of imitation will be underestimated. To assess the severity of this issue, the time in between B and C was added as a time-varying covariate indicating the likelihood that the focal entry is a reaction to an entry which occurred before its prior entry. This covariate, called PEIT, was interacted with the time periods. The more closely C followed B, the greater the likelihood that A was a reaction to B rather than C. In other words, the smaller the PEIT, the greater the chance of misestimation. Thus, for example, a negative coefficient on the interaction of a time period with PEIT means that as PEIT increases, the entry hazard decreases. This means that as the chance of misestimation decreases, the entry hazard decreases. This means that the time period coefficient in Model 2 overestimated the entry hazard.

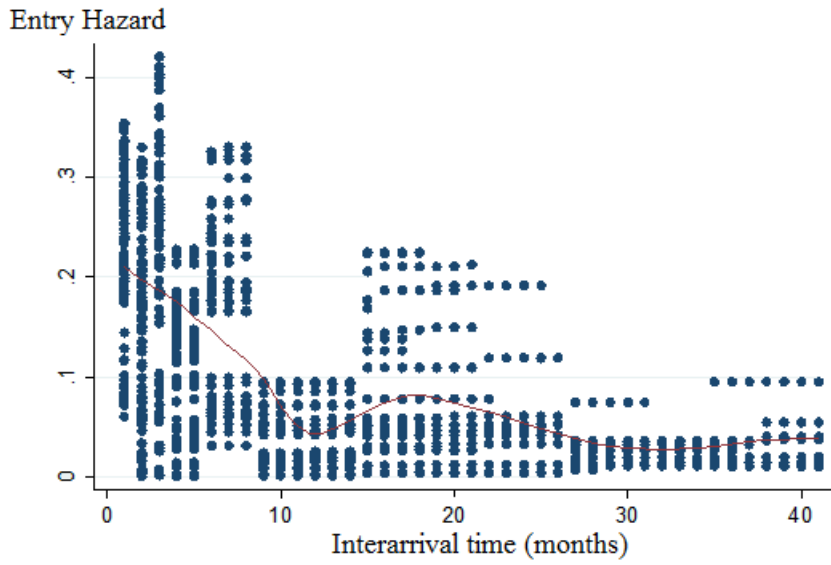
Model 3 shows the results with PEIT included. In all three cities, the interactions with time periods 1 up to 5, sometimes give non-negligible and negative coefficients. Statistical significance is less important because we have complete population data. Instead, a coefficient of -0.08 can already cross the entire range of variation in entry hazard spanned by the time period intercepts, within the domain of the values of the interaction term (PEIT ranges from 1 to about 60 for over 90% of observations (across the three cities, in the subsample of inter-arrival times shorter than eight months)). In other words, a coefficient of -0.08 means that a change in PEIT from 1 to 60 can change a period's estimated baseline hazard from being the highest hazard of all period to being the lowest. Therefore, such coefficient estimates are non-negligible. The negative interaction term coefficients thus indicate that common causes were somewhat overestimated in model 2. That is, the common cause time period intercepts in model 2 would have been lower if confounding entries had been excluded from the data. The interactions with time periods 6 up to 9 are not large

enough to cause the hazard pattern over adjacent periods to change over the range of PEIT. This indicates that confounding entries lead to negligible misestimation on the time periods in the imitation window. It makes sense for misestimation to be lower in those later periods because the periods are longer. So, if A was really a reaction to B rather than C, and B and C entered close together, then A would still count toward the entry hazard in the same period whether we start the time at entry B or entry C. In contrast, if A counts toward the entry hazard in a short period, then changing the starting time from C to B can more easily shift A toward a different period.

The graphical representations of these patterns are shown in Figure 3. To ease the interpretation, we used kernel-weighted polynomial smoothing ('lpoly' command in STATA). For Tokyo and Shanghai, the graphs correspond to the predicted pattern in Figure 1e. That is, the graph initially decreases, then increases and finally decreases. We predicted this pattern if both independent market entry as well as imitation were operating. The increase occurs later and is less steep in Shanghai. A possible reason for this is that, in Shanghai, entries started earlier in history when reaction speeds may have been slower, and later (i.e. between 1949 and 1980) when reactions were potentially faster as communication technology advanced, they were suppressed by the commercially unfriendly regulations (see Figure 2). Indeed, in Shanghai, the imitation spike was most pronounced before the second world war. In contrast, in Tokyo, the imitation spike was most pronounced after it. In Hong Kong, in Figure 3 panel (c), the entry hazard shows a monotonic decline. One plausible explanation for Hong Kong's pattern differing from that of Tokyo and Shanghai is that in Hong Kong we randomly (uniform distribution) assigned many observations to months within a year. This disperses away the natural clustering we could have observed. However, in this way the results for Hong Kong are informative as a baseline to compare how Tokyo and Shanghai deviate from a random distribution.

FIGURE 3
Entry Hazard over the Inter-arrival Time

a) Tokyo



b) Shanghai

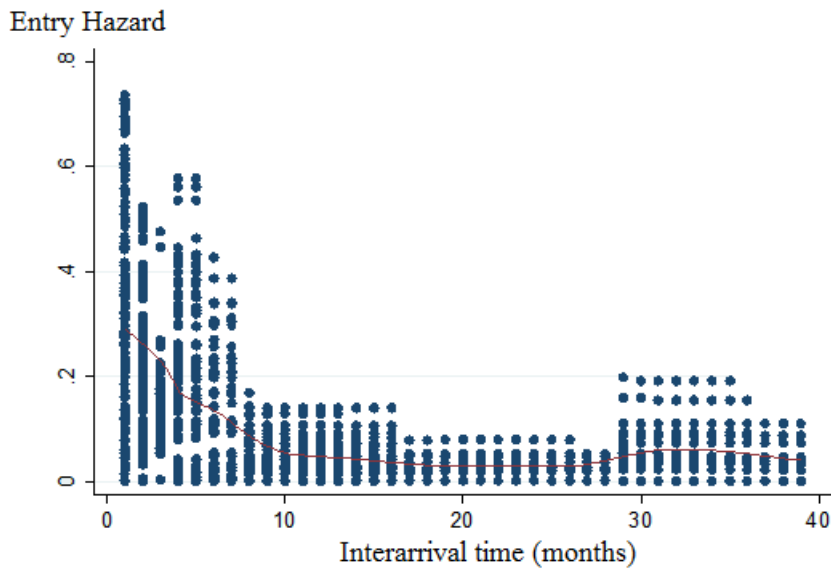
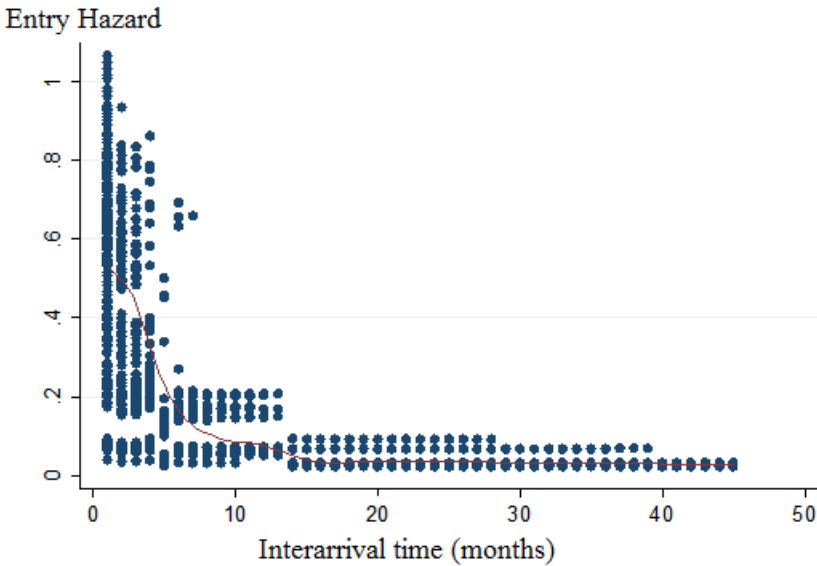


FIGURE 3 (continued)

c) Hong Kong



Because we have data on the full population, each spike is statistically significant. The get a sense of the economic significance we turn to the coefficient estimates. In Tokyo, the spike in the graph occurs in time period 7. The coefficient indicates that the hazard is about .75 higher than in time period 6. Thus, the spike shifts the hazard by 10-13 times density's main effect. For Shanghai, the coefficient of time period 8 is 0.8 higher than time period 7 in model 2 and 1.3 higher in model 3. Thus, the spike shifts the hazard as strongly as an increase in density by 13-26 banks. In addition, this interpretation shows that the graphs understate the similarity in the strength of Tokyo and Shanghai's spikes.

DISCUSSION

To address the thorny issue of causal inference in theorizing about observed market entries leading to conflation of two important types of entry in the extant literature (independent market entry and imitative market entry), we developed a theory to explain the relative timing of

consecutive market entries. The contingency of time turned out to be a useful tool for isolating different types of organizational responses. Specifically, our theory suggests that responses related to imitative market entry occur at a different point in time (relative to a prior entry) than responses related to independent market entry. We argued that information does not transfer instantaneously and completely between places, people and time, and that information must be transferred in the first place (Aharoni, 1966). Information takes time to be transformed from an event into a market entry. We have tested our theoretical ideas in three settings: foreign bank entries in Shanghai, Tokyo, and Hong Kong.

The theoretical – and by extension empirical – angle we developed allows us to uniquely identify each cause. Putting both types of market entry in perspective helps to nuance researchers' interpretation of the correlation between prior market entry and the likelihood of new market entry. Imitation has unique explanatory power, and the attention it receives in the literature (Guillén, 2003; Haveman, 1993; Lieberman & Asaba, 2006) is well-deserved. But, in theory building, clustering of market entries should not be completely attributed to the causal influence of past entries. Instead, sometimes two consecutive entries are independently driven by a common cause: a response to the same market opportunity.

Our results show that the entry likelihood spikes up during the imitation period, but the entry hazard does not become as high as soon after an entry. This suggests that common causes are more important than imitation in our setting. However, our approach was very conservative toward assigning an entry to the imitation period. For example, if organization B imitates organization A, and later organization C enters which also imitated organization A, then organization C is assigned as sharing a common cause with its prior entrant (i.e. organization B).

Thus, that we nevertheless find a spike in entry hazard during the imitation period, is strong evidence for a causal effect of one market entry on another.

Another aim of this paper was to contribute to our understanding of the temporal boundaries of inter-organizational imitation. In our data, a spike in the entry likelihood could be observed 20 or 30 months after a prior entry, and we attribute this to imitation. Thus, the temporal window for inter-organizational imitation is both delayed (because organizations need time to react) and fleeting (because information loses relevance over time). These findings also speak to recent calls for studies on the condition under which one type of cause of clustering of entries in time is more likely than another type of cause (Gimeno et al., 2005; Lieberman & Asaba, 2006). We argue that market opportunities are responsible for clustering if clustering is tightly packed in time, whereas imitation is responsible if clustering is more spread out, but not too much.

In terms of methodological implications, efforts to model inter-organizational imitation using panel data should therefore be careful to use the correct lag to avoid capturing independent entries instead. We recommend using multiple lagged variables when possible to simultaneously capture the different effects of independent and imitative market entries. With yearly data, a simple one-year lag for the effect of past entries on the likelihood of a new one occurring most likely captures independent market entries. Independent market entries may dominate (i.e. the decrease we associate with independent market entries is stronger than the increase we associate with imitation) in the first 12 months after a given entry, so a one-year lag would pick up most correlation among them. Naturally, these optimal specifications of lags vary across industries. Lomi, Larsen and Freeman (2005) compared the results of previous studies and found that lags may be especially long in our empirical context; the banking industry.

In addition, industries might vary in their number of subpopulations. The banking industry is relatively homogeneous (especially within the subset of international banks that we study). However, for example, the brewing industry is split into mass breweries and microbreweries. As a result, that population may not display the unimodal distribution of reaction speed that we assume. Another boundary condition to that assumption is that the range of possible entry modes is limited. The distribution of reaction speed in an industry may be split into a bimodal distribution of fast entry modes and slow entry modes. Another boundary condition is that there are no sudden shifts in entry barriers over time. To the extent that entry barriers influence reaction speed, the distribution of reaction speed may be split into a bimodal distribution of pre- and post-shift in entry barriers. However, if such shifts happen gradually over time, the distribution becomes a unimodal umbrella over the collection of unimodal distributions at each point in time. The war and the period of communism in Shanghai do reduce entries, but do not seem to reduce the likelihood of fast entries more than slow ones.

Another set of boundary conditions relates to the interdependence between the cities. For example, we observe banks that move their office from one city to the other. Such moves may involve different reaction speeds than entries by banks without a presence in a nearby city, creating separate modes for those who enter and those who move. We could avoid that issue by conceptualizing the three cities as constituting a single population. However, treating them as three separate populations is more in line with theory and previous literature. We do include a dummy variable for period of communism in Shanghai in all three cities' model, to account for the possibility that banks use Tokyo and Hong Kong as substitutes during that time (the dummy does have a negative effect in Shanghai). In additional analyses, we also included the density of banks in Tokyo as a covariate in the Shanghai model and vice-versa. The effects are small and

asymmetric; Tokyo's density correlates mostly positively with entries into Shanghai and its inclusion in the model weakens the effect of Shanghai's own density, while Shanghai's density has a weak inverted-u shaped relation with entry into Tokyo and its inclusion in the model strengthens the effect of Tokyo's own density.

We are only able to distinguish between causal and non-causal aspects of the correlation between prior entry and entry likelihood. However, the causal vs. non-causal aspects may correspond to certain mechanisms driving the correlation. If so, our results are evidence for at least one of the causal as well as one of the non-causal mechanisms. For example, the vicarious learning mechanism corresponds to the causal aspect. Therefore, the imitation spike is evidence in favor of vicarious learning that is not spurious due to a common cause. The externalities mechanism is more difficult to link to one aspect because it can be interpreted in both a causal way as well as a non-causal way. One example of an externality is if a bank enters Tokyo and launches a nationwide campaign to hire talented people. In response, a number potential bank employees move into Tokyo to start traineeships. As a result, recruiting talented people has become less costly for a new entrant into Tokyo because many potential employees are nearby. On the one hand, the prior entry causes the externality which causes the likelihood of new entry, so the effect of externalities should be part of the imitation spike. On the other hand, the causal effect does not start at the moment of the prior entry; it starts at the moment that the externality is realized (e.g. the talented people have moved into the city or ended their traineeship). Because externalities involve the reaction speed of talented people in addition to the reaction speed of the focal entrant, its effect is not tightly linked to a specific period in the inter-arrival time. Another interpretation of externalities is in terms of a broader infrastructure, which relates not so much to the most recently entered bank but rather to

the total number of existing banks. That interpretation suggests that the effect of externalities is a common cause that is captured by our density control variable.

The causal and non-causal aspects do not directly correspond to a particular decision-making process. However, the theory we developed may suggest insights into whether or not information about prior entrants is transformed into action faster than information about market opportunities. And would information about market opportunities become obsolete earlier? Such investigation would concretize how the decision-making processes differ between imitations and common cause entries. Such findings could then help understand if one process (e.g. independent assessment of the viability of a strategy) is associated with better performance or survival than the other (e.g. follow the leader)? Recognizing the ability of time as contingency to disentangle these processes in large datasets allows future research to test this intuition.

REFERENCES

- Aharoni, Y. 1966. *The foreign investment decision process*. Boston: Harvard University Press.
- Baron, R. A. 2006. Opportunity Recognition as Pattern Recognition: How Entrepreneurs “Connect the Dots” to Identify New Business Opportunities. *The Academy of Management Perspectives*, 20(1): 104–119.
- Baum, J. A. C., Li, S. X., & Usher, J. M. 2000. Making the Next Move: How Experiential and Vicarious Learning Shape the Locations of Chains’ Acquisitions. *Administrative Science Quarterly*, 45(4): 766–801.
- Belderbos, R., Olffen, W. V., & Zou, J. 2011. Generic and specific social learning mechanisms in foreign entry location choice. *Strategic Management Journal*, 32(12): 1309–1330.
- Carroll, G. R., & Hannan, M. T. 1989. Density Dependence in the Evolution of Populations of Newspaper Organizations. *American Sociological Review*, 54(4): 524–541.
- Carroll, G. R., & Hannan, M. T. 2000. *The Demography of Corporations and Industries*. Princeton, NJ: Princeton University Press.

- Cox, D. R. 1975. Partial likelihood. *Biometrika*, 62(2): 269–276.
- Delacroix, J., & Carroll, G. R. 1983. Organizational Foundings: An Ecological Study of the Newspaper Industries of Argentina and Ireland. *Administrative Science Quarterly*, 28(2): 274–291.
- DiMaggio, P. J., & Powell, W. W. 1983. The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, 48(2): 147–160.
- Fier, A., & Woywode, M. 1994. Enterprise foundings in the Eastgerman Transformationprocess. *ZEW Wirtschaftsanalysen*, 3: 235–260.
- Gimeno, J., Hoskisson, R. E., Beal, B. D., & Wan, W. P. 2005. Explaining the Clustering of International Expansion Moves: A Critical Test in the U.S. Telecommunications Industry. *The Academy of Management Journal*, 48(2): 297–319.
- Guillén, M. F. 2002. Structural Inertia, Imitation, and Foreign Expansion: South Korean Firms and Business Groups in China, 1987-95. *The Academy of Management Journal*, 45(3): 509–525.
- Guillén, M. F. 2003. Experience, Imitation, and the Sequence of Foreign Entry: Wholly Owned and Joint-Venture Manufacturing by South Korean Firms and Business Groups in China, 1987-1995. *Journal of International Business Studies*, 34(2): 185–198.
- Han, J. 1998. *The evolution of Japanese banking industry: An ecological analysis, 1873-1945*. dissertation, Stanford University.
- Hannan, M. T., & Carroll, G. R. 1992. *Dynamics of organizational populations: Density, legitimation, and competition*. Oxford University Press, USA.
- Hannan, M. T., & Freeman, J. 1987. The Ecology of Organizational Founding: American Labor Unions, 1836-1985. *American Journal of Sociology*, 92(4): 910–943.
- Hannan, M. T., Pólos, L., & Carroll, G. R. 2007. *Logics of organization theory: audiences, codes, and ecologies*. Princeton University Press.
- Haunschild, P. R. 1993. Interorganizational Imitation: The Impact of Interlocks on Corporate Acquisition Activity. *Administrative Science Quarterly*, 38(4): 564–592.
- Haunschild, P. R., & Miner, A. S. 1997. Modes of Interorganizational Imitation: The Effects of Outcome Salience and Uncertainty. *Administrative Science Quarterly*, 42(3): 472–500.

- Haveman, H. A. 1993. Follow the Leader: Mimetic Isomorphism and Entry Into New Markets. *Administrative Science Quarterly*, 38(4): 593–627.
- Henisz, W. J., & Delios, A. 2001. Uncertainty, Imitation, and Plant Location: Japanese Multinational Corporations, 1990-1996. *Administrative Science Quarterly*, 46(3): 443–475.
- Hong, X., Wang, X., & Li, A. 2003. *The History of Shanghai Finance*. Shanghai: Shanghai Academy of Social Sciences Press.
- Ji, Z. 2003. *A History of Modern Shanghai Banking: The Rise and Decline of China's Finance Capitalism*. M.E. Sharpe.
- Jones, G. 1992. International Financial Centres in Asia, the Middle East and Australia: A Historical Perspective. *Finance and Financiers in European History 1880-1960*: 405–428. Cambridge: Cambridge University Press.
- Jones, P. H. M. 1965a. Asian banks abroad. *Far Eastern Economic Review*, 165–168.
- Jones, P. H. M. 1965b. Foreign banks in Asia. *Far Eastern Economic Review*, 171–173.
- Kuilman, J., & Li, J. 2006. The Organizers' Ecology: An Empirical Study of Foreign Banks in Shanghai. *Organization Science*, 17(3): 385–401.
- Kuilman, J., Vermeulen, I., & Li, J. 2009. The Consequents of Organizer Ecologies: A Logical Formalization. *The Academy of Management Review ARCHIVE*, 34(2): 253–272.
- Levitt, B., & March, J. G. 1988. Organizational Learning. *Annual Review of Sociology*, 14: 319–340.
- Lieberman, M. B., & Asaba, S. 2006. Why Do Firms Imitate Each Other? *The Academy of Management Review*, 31(2): 366–385.
- Lomi, A., Larsen, E. R., & Freeman, J. 2005. Things Change: Dynamic Resource Constraints and System-Dependent Selection in the Evolution of Organizational Populations. *Management Science*, 51(6): 882–903.
- Lu, J. W. 2002. Intra- and Inter-Organizational Imitative Behavior: Institutional Influences on Japanese Firms' Entry Mode Choice. *Journal of International Business Studies*, 33(1): 19–37.

- Manski, C. F. 1993. Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 60(3): 531–542.
- Martin, X., Swaminathan, A., & Mitchell, W. 1998. Organizational Evolution in the Interorganizational Environment: Incentives and Constraints on International Expansion Strategy. *Administrative Science Quarterly*, 43(3): 566–601.
- Mitchell, T. R., & James, L. R. 2001. Building Better Theory: Time and the Specification of When Things Happen. *The Academy of Management Review*, 26(4): 530–547.
- Ruef, M. 2006. Boom and Bust: The Effect of Entrepreneurial Inertia on Organizational Populations. *Advances in Strategic Management*, 23: 29–72.
- Sørensen, J. B., & Sorenson, O. 2003. From Conception to Birth: Opportunity Perception and Resource Mobilization in Entrepreneurship. *Advances in Strategic Management*, 20: 89–117.
- Tamagna, F. 1942. *Banking and Finance in China*. New York: Institute of Pacific Relations.
- Tatewaki, K. 2002. *One hundred year history of foreign banks in Japan, 1900-2000*. Tōkyō: Nihon Keizai Hyōronsha,.
- Terlaak, A., & Gong, Y. 2008. Vicarious Learning And Inferential Accuracy in Adoption Processes. *Academy of Management Review*, 33(4): 846–868.
- Tuma, N. B., & Hannan, M. T. 1984. *Social Dynamics: Models and Methods*. Orlando: Academic Press.
- Vermeulen, F. 2010, July 9. *Imitation*. <http://dspace.library.uu.nl/handle/1874/45127>.
- Wholey, D. R., Christianson, J. B., & Sanchez, S. M. 1993. The Effect of Physician and Corporate Interests on the Formation of Health Maintenance Organizations. *American Journal of Sociology*, 99(1): 164–200.

Chapter 3

Is it a Bird? Is it a Robin? Imitation when Potential Categories

Vary in Level of Detail

ABSTRACT

For each phenomenon there exists a taxonomy of categories, ranging from broad generic categories to specific detailed categories. Given this large number of options, when an organization interprets the actions of other organizations, which category is actually used? We draw on psychology research about the basic level of categorization to suggest that categorizers tend to use categories from the level of detail where categories have the greatest informativeness and distinctiveness. Categories are more informative the more attributes of an object that can be inferred from knowing the object belongs to that category. Categories are more distinctive the fewer attributes of an object need to be observed in order to determine that the object belongs to that category. In data on market entry by all foreign firms in France from 1999 to 2011, the effect of measurable antecedents of these two factors support this theory in the context of organizations: 1) greater specific-level variety of a category causes a tendency toward using its more detailed subcategories instead, and 2) greater typicality of a subcategory causes a tendency toward using its broader counterpart instead.

INTRODUCTION

Consider organization A, a German firm investing in France. Specifically, it chooses to enter industry X in region Y using a greenfield entry mode. This action is a cue to A's competitor, organization B, which observes this action and may want to imitate it. How does organization B interpret the cue, and precisely what imitative action follows? Does organization B enter France? Guillén (2002) models it that way. Does it enter industry X? Studies such as Hannan et al. (1995) have modeled it that way. Or does B enter industry X in France (studies such as Haveman (1993) model it this way); greenfield into France (as Guillén (2003) models it); greenfield into industry X in France (as Li, Yang, and Yue (2007) model it); greenfield anywhere (studies such as Fligstein (1985) model it that way); or greenfield into industry X in region Y (studies such as Greve (2000) model it that way)? And what if organization B is not from Germany and/or not a competitor? In general, if a competitor enters through greenfield, what determines whether a focal organization interprets this as a cue to enter a foreign market, or to enter through greenfield specifically? There is to date no answer for this question; there is not even a framework for beginning to think about an answer. We aim to develop such a framework.

Foreign market entry is used as a context to make our discussion more concrete. In this context, we consider categories at two different levels of detail: broad categories that exist at the generic level, and detailed categories at the specific level. A detailed category is a subcategory of a broad category. A broad interpretation of foreign market entry is that an organization enters a country-industry. A detailed interpretation adds at least one more dimension: country-industry-entry mode in the hypothetical example. Entry mode means: an organization entering a foreign market by wholly, or partially, acquiring an existing foreign organization, or by building a new

branch in the foreign market from scratch with the cooperation of a foreign organization (e.g. joint venture), or completely on its own (i.e. greenfield).

Previous research shows that social categories influence organizations' actions. For example, whether an organization is a member of one category as opposed to spanning multiple categories has an effect on that organization's external evaluations (Negro, Hannan, & Rao, 2010; Pontikes, 2012; Zuckerman, 1999). We argue that a category's place in a taxonomy and the taxonomy's structure moderate this effect. The categories an organization spans may be related or unrelated at a broader level, which influences evaluations (Wry, Lounsbury, & Jennings, 2013; Wry & Lounsbury, 2013). More fundamentally, whether an organization spans or does not span multiple categories in the first place depends on the level of detail at which the categories are defined. For example, an organization which spans the categories 'Spanish cuisine' and 'Italian cuisine' would not span categories if the categories were defined at the broader level of 'restaurants' and 'supermarkets'. If a researcher defines categories at a level of detail that audiences do not use for evaluations, effects of category spanning are likely to be underestimated. Understanding under what conditions evaluators use which level of detail helps avoid such underestimation. The framework we present helps identify such conditions.

Investigating levels of detail is also important for explaining the diversity of organizations and organizational actions in a population (DiMaggio & Powell, 1983; Greve, 2005; Hannan, Pólos, & Carroll, 2007). This literature argues that social categories drive imitation by penalizing organizations that do not imitate those actions that the social category defines as important (Kovács & Hannan, 2010; Pontikes, 2012; Zuckerman, 1999). Previous research has focused on the degree of imitation as a key determinant of diversity (Carroll & Hannan, 2000), but the level of detail at which imitation occurs also drives diversity. Consider two industries where imitation is equal in

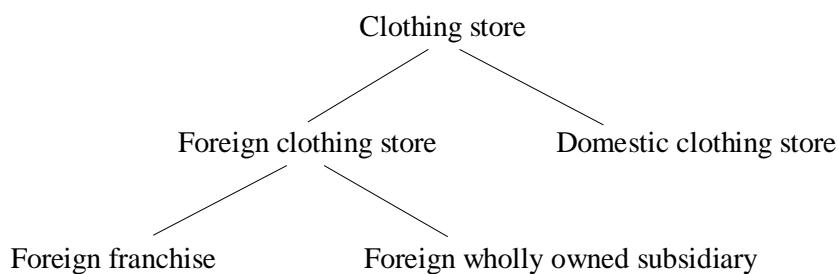
degree but different in level of detail. In one industry, the imitation is broad, such that all organizations have a presence in foreign markets, but the entry modes used can vary widely. In another industry, the imitation is detailed, such that presence in foreign markets varies widely, but those that do go abroad all use joint ventures to do so. This example shows how knowledge of the degree of imitation is a poor predictor of diversity without an understanding of the level of detail. We aim to contribute such an understanding, and therefore this literature's predictions become less poor because of our contribution.

Finally, without taking into account a category's place in the taxonomy, evidence of imitation may be spurious. That is, what appears to be detailed imitation may in fact be a manifestation of broad imitation or vice-versa. In our example context that means that if the number of past and present occurrences of an entry mode are correlated, then the number of past and present occurrences of foreign entries are also likely to be correlated. Thus, a correlation could be taken as evidence of imitation where in fact it does not exist (Hox, 2010). Instead it exists on a different level. One example of where this ambiguity leads to difficulties is Guillén's (2002) conclusion that firms in his data imitate each other's foreign entry, though in another paper Guillén argues that those same firms imitate each other's entry mode choices (Guillén, 2003). The theory we contribute allows researchers to identify conditions under which the former interpretation is more appropriate than the latter. Therefore, our contribution helps avoid spurious inferences.

Our contribution is to describe a framework that maps different interpretations onto a dimension which might be called 'level of detail in a taxonomy', and to formulate a theory that explains why an organization tends to interpret cues at a particular level of detail in a taxonomy. 'Taxonomy' here refers to a hierarchical tree structure of categories and subcategories (see figure 1). Taxonomies vary across industries in their structure. Industries vary in terms of the level of

detail of categories that are used most prominently. We explain how variation in these two aspects are correlated with each other. We think this contribution is useful for research that uses categories (Kennedy & Fiss, 2013; Vergne & Wry, 2014). Specifically, studies of interorganizational imitation show that categories influence who and what is imitated (Haveman, 1993), and that there exists, at least in some industries, a taxonomy of categories that organizations use to define with whom they compete (Porac, Thomas, Wilson, Paton, & Kanfer, 1995). Our contribution is to theorize which level of detail in such a taxonomy is likely to be used; or in Vergne and Wry's (2014) terminology: at which level in a 'classification hierarchy' do the categories have greatest 'saliency'.

FIGURE 1
Example of a taxonomy of categories



The structure of the paper is as follows. First, we explain in the context of interorganizational imitation, what it means for some categories in a taxonomy to be used. Then, we identify the two main drivers of the use of a category (relative to categories from other levels of detail in the same taxonomy). Third, we explain two dimensions in which taxonomy structure can vary, and formulate hypotheses about the correlation between variation in these dimensions and variation in the level of detail of the category that is used most. Then, we measure whether industries vary in the level of detail at which imitation tends to occur. We test hypotheses to predict this variation on data on all foreign firms in France from 1999-2011.

THE ROLE OF CATEGORIES IN IMITATION

We aim to explain category saliency. A category has saliency to the extent that it is used by organizations in a three-stage process that results in imitation. First, scanning their environment for information (e.g. find out if a competitor engaged in a foreign entry). Second, the interpretation of the information. In this stage, the action is labelled (i.e. categorized) in order to understand it and communicate about it (Dutton & Jackson, 1987; Negro, Koçak, & Hsu, 2010; Thomas, Clark, & Gioia, 1993). For example, organizations can categorize a competitor's action as 'a foreign entry' or, more specifically, as 'a foreign acquisition'. The level of detail is determined at this point because the category that is used has a particular level of detail. There were a number of categories potentially able to describe the competitor's action, but all except the resulting category are not used. Third, acting upon the interpretation. Dutton and Jackson (1987) further argue that if an organization categorizes the information as, for example, a 'threat' as opposed to an 'opportunity', that affects the organization's future decisions regarding that issue. Similarly, the category that is used in the interpretation should influence whether an organization imitates its competitor by means of a foreign entry or a foreign acquisition. If the specific level of detail is not used to categorize the competitor's action, it is implied that organizations do not decide their entry mode based on imitation, but perhaps based on which mode is more profitable or best fits their strategy.

Not only the competitor's action can be categorized but also the competitor itself. Organizations can categorize a competitor as 'a member of the same industry' or, more specifically, as 'a member of the same industry and from the same (or a different) home country'. The level of detail an organization uses should influence which of its competitors' action it imitates. Categorizing competitors influences the scanning process underlying imitation, while

categorization of actions influences the interpretation process underlying imitation. We do not differentiate between the stages because we expect them to operate jointly.

In sum, imitative behavior is an observable consequence of category saliency. If detailed categories are salient, the imitation should be detailed; if broad categories are salient, the imitation should be broad. Broad imitation is observed from a correlation between the number of existing foreign firms in an industry and the number of new foreign entries into that industry in a year (Guillén, 2002). Detailed imitation is observed from a correlation between the number of existing foreign firms that entered via greenfield and the number of new greenfields into that industry in a year (Guillén, 2003).

To illustrate, consider a Spanish shoe retail entrepreneur that hears of a competitor engaging in a joint venture in France. The shoe retailer now becomes aware that a joint venture in France is an option that he can consider (Aharoni, 1966). In doing so he is using the category ‘joint venture in France’ to interpret the competitor’s action. The increased awareness leads to an increased likelihood of engaging in that action himself. Alternatively, he could have become aware that ‘an entry into France’ is an option that he can consider. Or, he could have become aware that ‘a joint venture in France’ is an option for shoe retailers like his competitor, but not for himself. This might happen if the competitor is a German shoe retailer. In other words, he could have used the category ‘foreign entry’ or the category ‘joint venture by German shoe retailers in France’ to interpret the competitor’s action. Which interpretation of the competitor’s action is salient to the shoe retailer influences which awareness increases, which influences the likelihood of our shoe retailer engaging in a joint venture.

Next to awareness, news can also affect the shoe retailer’s motivation to act. For example, he hears that the competitor’s joint venture has been profitable. He now considers that joint

ventures into France are likely to be a profitable option. Options that are perceived as more likely to be profitable are more likely to be taken. However, he could have instead come to consider that foreign entry into France is a profitable option. Or he could have come to consider that joint ventures in France are profitable for other shoe retailers like his competitor, but not for himself. In this way, the salience to the shoe retailer of the category 'joint venture in France' relative to the category 'foreign entry into France' can cause joint ventures in France to be perceived as more profitable, while the perception of the profitability of foreign entry into France is not changed by the news.

The greater the number of joint ventures, the more likely that a shoe retailer hears of a competitor engaged in one, or hears of a competitor's joint venture being profitable. The news is categorized and then leads to increases in awareness and motivation associated with its category. Over all shoe retailers, different categorizations consequently result in different correlations between the number of existing joint ventures and new joint ventures. For example, increased awareness and motivation about joint ventures from news of a joint venture should lead to a stronger correlation between existing joint ventures and new joint ventures, compared to the situation where joint ventures increase awareness and motivation about foreign entry. Because in the latter case, shoe retailers may enter via another mode instead.

It is helpful to describe how categories operate at the population level because we use population-level measures. At the population level (e.g. an industry in a country) audiences use categories to constrain the actions of the population's members (e.g. organizations in a particular industry-country) (Hannan et al., 2007). For example, Li and colleagues (2007) reason that 'joint venture' is an identity code that sets expectations for, and circumscribe, organizations. Similarly, Yiu & Makino (2002) have argued that 'joint venture' as a category supports the use of joint

ventures themselves. This is because without an available category, it is difficult for audience member to formulate an interpretation of an organization's action, so that the action fails to gain attention (Zuckerman, 1999). In contrast, with that category, actions can be accurately judged by audiences, such that actions that fit nicely within the category of 'joint ventures' are more appreciated. That appreciation is needed for the density of joint ventures (i.e. the number of joint ventures operating in an industry) to influence the likelihood of new joint ventures.

A category is made available by the population-level collective of audience members. The process by which the population makes categories available to its members is called 'consensus'. An individual may have a category to describe an action, but without consensus, the category cannot lead to a level of appreciation necessary for the action to occur. Thus, only categories that are made available at the population-level can be used. Different populations make available different categories (Hannan et al., 2007; Navis & Glynn, 2010). For example, for some time, the category 'horseless carriages' was the only available category to describe products later categorized as 'automobiles'. Importantly, because of this categorization, horse-carriages and automobiles were seen as competitors (Dobrev, Ozdemir, & Teo, 2006: 581). We add the nuance that each category captures only one aspect of an action. If the 'joint venture' category is unavailable, the joint-venture aspect of a competitor's action fails to gain attention and appreciation. Therefore, the density of joint ventures would be less influential for new joint ventures. However, the density of foreign entries could cause more new foreign entries of which some are joint ventures. Similarly, for some time, the automobile aspect failed to gain attention, but automobiles could be produced as a part of the broader category 'carriages'.

To summarize, categories influence the way organizations react to the actions of their competitors. Different potential categories have different influences. Therefore, which of these

categories is used influences, and can be observed from, organizations' reactions to actions by competitors. Specifically, we focus on organizations' imitative reaction to the number of existing foreign entries (which we refer to as the broad category's density) compared to their reaction to the number of existing foreign joint ventures (i.e. the detailed category's density). Thus, a category is salient to the extent that its density causes imitative reactions.

INFORMATIVENESS AND DISTINCTIVENESS

What causes a category to be salient? We apply theory from research on individuals' tendency towards a certain level of detail in categorizing objects (Murphy, 2002; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), to organizations categorizing competitors' actions. Such application expands on Porac and Thomas' (1995; 1990, 1994) application of that research to the organizational context. The handful of organizational studies that cite this literature only tangentially build on its ideas (Kovács & Hannan, 2010; Thomas et al., 1993). And more recently Wry and coauthors have built on a different stream within the same psychology literature (e.g. Wry et al., 2013).

The theory posits that individuals tend to categorize objects at a level of detail that provides most informativeness without sacrificing too much distinctiveness. Individuals are not conscious of the working of the tendency, and the tendency itself is fully developed after the age of 18 to 24 months (Mervis & Rosch, 1981: 94). We distinguish three steps in the categorization process: 1) observing the attributes of an object; 2) categorizing the object by cross-checking its observed attributes against the attribute lists of potential categories; and 3) predicting unobserved attributes of the object based on the attribute list of the category in which it has been categorized. A category's distinctiveness is the degree to which its attribute list overlaps with other categories'

attribute lists. High distinctiveness makes it easily used in step 2. A category is highly informative if it allows many accurate predictions in step 3.

More detailed categories are more informative because knowing more detail allows more to accurate prediction. A broad category is less informative because there is more variation among the attributes of its members, making the accuracy of prediction in step 3 low. For example, knowing that an object belongs to the broad category ‘mammal’ does not allow many accurate predictions about its unobserved attributes because mammals range from whales to squirrels to humans. On the other hand, it is easier to categorize an object broadly due to broad categories’ high distinctiveness. That is, one broad category’s attribute list is very distinct from those of other broad categories (i.e. there is low overlap). This makes it relatively easy (i.e. cognitively cheap) to decide to which broad category an object belongs. In contrast, many of a detailed category’s attributes are also found in other (related) detailed categories’ attribute lists, so assigning a detailed category is difficult (i.e. cognitively expensive). For example, we observe the following attributes of an object: it is animate, can fly, has feathers, and has a beak. This allows the object to be categorized in the broad category of ‘bird’. No other broad category’s attribute list contains all those attributes. But there are many detailed categories that include the observed attributes on their attribute list. As a result, many more cross-checks would be required to decide whether or not the object belongs to the detailed category ‘robin’. In sum, broad categories have the advantage of distinctiveness, but the disadvantage of being less informative. Detailed categories are more informative but less distinctive (and therefore, cognitively expensive). From very broad categories to a little less broad, distinctiveness decreases only slightly while informativeness increases greatly. From somewhat specific categories to very specific, distinctiveness decreases greatly while informativeness increases only slightly. As a result, there is a sweet spot at which changing

the level of detail results in either losing informativeness without gaining much distinctiveness or losing distinctiveness without gaining much informativeness. People tend to use (categories at) the level of detail that is at this sweet spot in the trade-off between informativeness and distinctiveness (Murphy, 2002; Rosch et al., 1976).

It is necessary to also explain this link in terms of a population-level effect. In fact, Rosch et al.'s (1976) theory builds upon more macro-level research on differences between languages. By language we mean a collection of expressions that is made available to the speakers of that language. Each population has its own language. Research on language finds that expressions with less informativeness or less distinctiveness are made available later in the development of a language (Berlin, Breedlove, & Raven, 1973; Rosch et al., 1976). For example, the expression 'hammer' existed before the expression 'claw hammer'. 'claw hammer' could not have been a salient category as long as this word did not exist. Therefore, language constrained the categorization of claw hammers to the broad 'hammer' category for some period. Thus, language has the same tendency as individuals towards categories at the sweet spot in the trade-off between informativeness and distinctiveness.

Interestingly, expressions that balance informativeness and distinctiveness tend to be 'single word' expressions, while more detailed expressions tend to consist of two words where one of the words tends to be the broad category name (e.g. 'hammer' and 'claw hammer') (Berlin et al., 1973: 217). Expressions that are broader than the sweet spot also tend to be single words, but they also differ linguistically: very broad expressions are more likely to be mass nouns. For example, furniture is a broad expression and is a mass noun (i.e. one must say 'a piece of furniture' instead of 'a furniture') while none of the detailed types of furniture are designated by mass nouns (Markman, 1985). These examples indicate uniquely linguistic markers of how expressions with

less informativeness or less distinctiveness seem to be secondary to expressions at the sweet spot of the trade-off between them. This suggests that, next to the cognitive tendency discussed earlier, there is a population-level mechanism that makes available a category depending on its informativeness and distinctiveness.

HYPOTHESES

The previous section described how experimental research shows that informativeness and distinctiveness drive categorization by individuals in a laboratory setting, but do these drivers also apply in the context of organizations' categorizations in the process of imitation? To find out, we derive implications from Rosch et al.'s (1976) theory about interorganizational imitation. If data are in line with these implications, and there are no plausible alternative explanations, we conclude that the theory helps explain the actions of organizations.

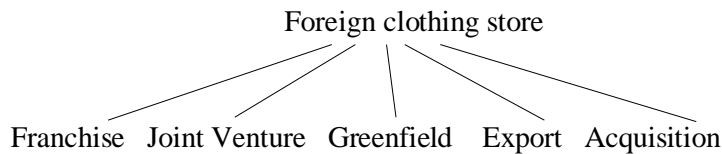
In this section we propose concrete observable antecedents of informativeness and distinctiveness. Because our goal is to demonstrate the usefulness of Rosch et al.'s (1976) theory we chose two core variables in this line of research as antecedents: 1) The number of different detailed categories within a broad category, and 2) how typical of its broad category a detailed category is. Our hypotheses are that variation in those antecedents moderates the effect of density on new entries.

The first antecedent relates to variation in informativeness. Specifically, variation in the number of its subcategories should affect the informativeness of a broad category. Broad categories are less informative than detailed categories because of the large differences between objects within the same broad category. But rather than as a matter of fact, the extent of differences between objects within a broad category can be viewed as a variable. That is, there may be some broad categories with more differences between member objects than others, and those broad

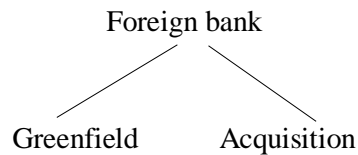
categories should be more disadvantaged in terms of informativeness. We refer to the extent of differences between objects in a category as the ‘specific-level-variety’. We focus on those differences that are relevant to the detailed categories at our chosen specific level of detail. In our example context, high specific-level-variety means that many different entry modes are used in an industry (see figure 2).

FIGURE 2
Visualization of specific-level variety

a) High specific-level variety



b) Low specific-level variety



In industries with low specific-level-variety, all organizations are viewed as similar, so broad and detailed categories allow equally accurate predictions. The lack of improvement in prediction accuracy from using detailed categories means it is not worth the additional cognitive cost. But in industries where a large number of entry modes are being used, organizations are more different. Therefore, broad categories are greatly disadvantaged in terms of informativeness. This implies that with high specific-level variety, detailed categories allow much more accurate predictions than broad categories. Therefore, the benefits of detailed categories are now more likely (compared to if specific-level variety were low) to offset the additional cognitive cost of detailed categories. As a result, detailed categories are more likely to be used. In other words, as

specific-level variety increases, the sweet spot in the trade-off between informativeness and distinctiveness shifts away from ‘foreign entry’ and toward ‘joint ventures.

Returning to the Spanish shoe retailer, he will tend to be interested in the specific entry mode used by his competitor if shoe retailers in general use a wide variety of entry modes. In contrast, if every shoe retailer who invests in France uses the joint venture mode, the Spaniard is less likely to attend to that specific information. The shoe retailer is thus influenced to categorize the competitor’s action differently depending on the specific-level variety. As a result, the effect of detailed category density on entries (i.e. the saliency of the detailed category) should depend on specific-level variety. To tease out this effect from ways in which specific-level variety may impact both the detailed category and the broad category in the same industry, we formulate the hypothesis relative to broad category saliency.

At the population-level, detailed categories are more likely to be made available as specific-level variety increases. For example, Tzeltal Mayans encounter, in their daily life, more specific-level variety in terms of plants than suburban Californians do. As a result, the Tzeltal Mayan’s language contains (i.e. makes available) more expressions for different types of plant than the language of suburban Californians (Dougherty, 1978). Analogously, we expect that populations that encounter more variety in the foreign entries, make available more detailed categories of foreign entries (e.g. joint venture). If detailed categories are not available, detailed density’s influence on new entries is weakened. To be more precise, greater specific-level variety increases the likelihood of detailed categories are made available relative to the likelihood that broad categories are made available. Therefore, greater specific-level variety increases detailed density’s effect relative to broad density’s effect.

H1: The specific-level-variety within a broad category is positively related to the saliency of the detailed categories relative to the broad category.

Hypothesis 1 takes the broad category's disadvantage (i.e. low informativeness) and frames it as a variable by saying that some broad categories are more disadvantaged than others. The next hypothesis takes the detailed category's disadvantage (i.e. distinctiveness) and frames it as a variable. That is, some detailed categories are more disadvantaged than others.

Research in psychology (Mervis & Rosch, 1981; Murphy, 2002; Rosch et al., 1976) as well as in organization ecology (Hannan et al., 2007; Kuilman & Li, 2009) confirms that category members often differ in their degree of membership. For example, robins are more typical members of the category 'bird' than penguins. The detailed category that is the most typical of a broad category is the one that has the most attributes in common with other detailed categories in that broad category. Atypical detailed categories, such as penguin, have fewer attributes in common with other subcategories of birds (Murphy & Brownell, 1985; Smith & Medin, 1981).

Detailed categories are disadvantaged in terms of cognitive economy. Their lack of distinctiveness means that many attributes need to be cross-checked in order to categorize an object on the detailed level. The idea of 'typicality' nuances this argument: especially typical detailed categories have many attributes in common with other detailed categories, and therefore typical detailed categories especially lack distinctiveness. Distinctiveness makes categorization cognitively cheap, so typical detailed categories are especially disadvantaged in terms of cognitive economy. Categorizers tend to avoid this disadvantage by adjusting the level of detail they use. Therefore, members of typical detailed categories tend to be categorized on a broader level of detail instead. In contrast, atypical detailed categories are less disadvantaged and members of

atypical detailed categories are more likely to be categorized on a specific level of detail. In other words, as typicality increases, the sweet spot in the trade-off between informativeness and distinctiveness shifts toward ‘foreign entry’ and away from ‘joint venture’.

An example of this effect at the macro-level is the subcategories of massage businesses. One of the subcategories that we test for is an organization’s home country. One can imagine industries where different home countries correlate with different collections of practices and therefore each country has an industry subcategory. For example, Thai massage is a distinct subcategory in the massage industry, whereby the home country is a modifying term; it modifies the broad category expression ‘massage’ to form a subcategory. The effect of typicality is clear when people refer to ‘the normal kind of’ massage, or ‘classic’ massage. Only a few languages make available an expression for this ‘normal kind of’ massage subcategory: ‘Swedish massage.’ Swedish massage is the most typical type of massage, and the effect of typicality prevents many languages from making available the special subcategory expression. Because of its typicality, this type of massage has lost its association with a particular home country. In general, people are often left to use words like ‘classic’ or ‘normal’ to distinguish the most typical subcategory from others because an actual expression for that subcategory is not available, But mostly, such words are omitted such that the expression for the most typical subcategory is the same as the broad category.

Thus, because typical subcategories are less likely to be available such that people defer to the broad category, and because subcategories help subcategory density influence new entries, the effect of the density of a typical subcategory on new entries is weaker. Instead, those entries are more likely to be a side-effect of the broad category density causing generic foreign entries of which some happen to be typical.

At the organization-level, we expect that the interpretation of typical foreign market entries is less likely to involve the detailed category and people are more likely to use the broad category instead. In other words, entrepreneurs from an atypical home country imitate only their fellow countrymen's foreign entries, while entrepreneurs from typical home countries are more likely to imitate entries from any country.

H2: The typicality of a detailed category is negatively related to the saliency of the detailed category relative to the broad category.

DATA AND METHOD

We analyze the entries by foreign firms into all industries in France from 1999 to 2011. The data were collected by the Institute national de la statistique et des études économiques (INSEE), the French national statistics agency, in their annual survey of firms in France. This survey aims to capture the complete population of foreign-owned firms in France. They were supplemented by data from Bureau van Dijk. We accessed the raw firm-level data through the CASD program of INSEE.

In the data, each firm has a unique identification number. We treat these firms as separate entities though some may have been owned by a common parent company. An entry is recognized when a firm with a new identification number is included in the survey. The number of such entries that occur in a subcategory in a year is the dependent variable.

The first independent variable is the number of firms in a broad category (i.e. an industry, using the most specific classification of industries) in a year. We refer to this as the broad category's density. The second independent variable is the number of firms in a subcategory in a year (detailed density). Models such as ours, with density as an independent variable to predict

market entry, are a proven method to study the effect of categories (Carroll & Hannan, 2000; Kuilman & Li, 2009; Li et al., 2007). It is an especially useful model because the variables have similar meaning across industries. In contrast, methods with other dependent variables are more context specific. For example, analyst attention (Zuckerman, 1999), or rankings in a magazine, or getting venture capitalist backing (Pontikes, 2012), are just not meaningful for the paper manufacturing industry, for instance.

The advantage of focusing entirely on foreign entries is that the broad density is less correlated with the degree of competition because it does not include the domestic competitors. Detailed density should also have a low correlation with competition because it excludes even more competitors. In contrast, if we had chosen to include domestic entries, the broad density would strongly pick up a competitive effect while the detailed density would not. Such asymmetry could have created the different reactions to the moderators we hypothesize, but for reasons beyond our theory.

To form the subcategories, we divided each industry in three ways: by legal form, by firm size (number of employees in France of the firm's parent company), and by home country. Each legal form group is an official 4-digit 'catégorie juridique', which are groups of legal forms that corporations can have in France (e.g. SARL, which is equivalent to 'Ltd' in English-speaking countries). The cutoff point for the firm size groups are 1, 5, 20, 200, 500, 1500, 5000, and 10000 employees. For each industry, we only created a subcategory if at least one firm existed in that subcategory in that year. We aim to have subcategories which are plausible as potentially relevant categories for organizations when interpreting actions. The legal form division is plausible because of its link to organizational form. Organizational forms such as joint venture as opposed to greenfield have been conceptualized as being a social category (Yiu & Makino, 2002). The size

and home country groupings are plausible because prior literature has already shown that organizational imitation is affected by variation in these dimension (e.g. Haveman, 1993; Li et al., 2007; Porac et al., 1995). An additional advantage of these divisions is that each subcategory is unambiguous and mutually exclusive, so we avoid the complications of having firms that are a member of multiple subcategories.

It is plausible that there are differences between industries in the extent to which words are made available to describe these subcategories. For example, in the alcoholic-beverage industry the country of origin is always present on wine bottles, and some styles of drink are named after regions (e.g. champagne, American pale ale), and organizations are divided and named based on their size (e.g. microbreweries vs. mass breweries). In contrast, in the vegetable industry, vegetables are rarely named after a region nor are the producing organizations labelled according to their size. Our hypotheses mean that we expect those differences between the alcoholic-beverage and vegetables industries to be related to differences in the specific-level variety and typicality of their categories.

We excluded industries where only one subcategory was populated because such observations would be perfectly collinear between broad category and subcategory and therefore would not add much information. Furthermore, in 2007 the industry classification system changed. The pre-change and post-change industries are treated as separate unique industries. As a result, firms change industry. However, because the firm identification numbers already existed in the data, industry switches are not counted as entries. Ultimately, in the legal form data there are 1379 industries and 7530 subcategories, with an average of 22.4 observations per industry and 4.1 observations per subcategory. The corresponding numbers for firm size groupings are 1409 industries and 8571 subcategories, with averages of 25.1 and 4.1 observations. For the home

country data the corresponding numbers are 1405, 13,490, 38, and 4. The 4.1, 4.1 and 4 observations per subcategory represent the average number of years that a subcategory was represented in the data.

Our hypotheses are about two moderators of the effect of both densities on entries. First, the measure for specific-level variety is the number of different subcategories within each industry populated by at least one firm during at least one year. For example, if in industry A, organizations use five different legal forms, specific-level variety is high. Second, the measure for typicality is the percentage of industry density accounted for by each subcategory.¹ For example, if in industry A, one of the five legal forms is used by 90% of the organizations while the other four legal forms share the remaining 10%, the one legal form is considered more typical than the other four. To reduce overlap with the density measure, we use, for all years, the typicality-score of the first year that an industry-subcategory combination is observed. The theoretical justification for this is that a subcategory's typicality should be somewhat sticky relative to the fluctuating number of

¹ This measure requires that the same legal form or home country or firm size has different typicality in different industries. Technically, this means that the attribute overlap of one, say, size with other sizes depends on the industry. We can make that plausible by introducing one assumption: that firms of a particular size differ systematically in terms of a number of other attributes from firms of other sizes within the same industry. Although size itself is not different between industries, the attributes associated with it are. As a result, the typicality of a size group in terms of attribute overlap can differ between industries. For example, consider a large restaurant (i.e. with a high number of employees) compared to a small one. Attributes associated with the large restaurant could be 'longer opening hours', 'part of a restaurant chain', or 'very formal food preparation procedures'. Attributes associated with the small restaurant could be 'family owned', 'little seating space', 'mostly makes deliveries'. In contrast, in the banking industry, large and small banks differ in different attributes than those. This example shows how we can plausibly think about size to associate with many other attributes that are specific to the industry. In line with this, Porac et al. (1995, p.222) find that knitwear managers define categories by those dimensions (e.g. size) that allow inferences about other attributes (e.g. opening hours). One last premise that this measurement needs is that typical size-groups' other attributes are more common in other size-groups than atypical size-groups' other attributes. Continuing the restaurant example, this means that if small restaurants are typical and large restaurants are atypical, then medium size restaurants should be more likely to have the attributes 'family owned' and 'does deliveries' than the atypical attributes 'part of a restaurant chain' and 'extensive opening hours'.

organizations within that subcategory (Porac et al., 1995: 206), especially given that we observe a subcategory for only four years on average.

We add the variable ‘taxonomy’ which counts for each industry (at the 4-digit level) the number of other 4-digit industries within the same 3-digit class. With this variable we take into account the place of our observations in the broader taxonomy. By interacting this variable with the density variables we control for the possibility that taxonomy correlates with specific-level variety or typicality and creates spurious focal interaction coefficients. For the same reason, we include an interaction term between the broad and detailed density. The models also include squared terms of all main variables to prevent any nonlinear effect to be absorbed by the interaction terms. In addition, year-fixed effects are included to improve the accuracy of the density estimates. Finally, we exclude 1999 from the regression models because all firms would have been counted as new entries.

Because of our interest in comparing the effects of the same variable (density) at different levels of analysis (industry vs. subcategory), random effects at the industry and subcategory level are crucial. We also include an industry-level random effect on the coefficient of subcategory density and on the coefficient of the interaction between subcategory density and industry density. Such random effects are basically additional error terms. The addition of those random effects makes the model a multi-level model (Hox, 2010).

First, the industry-level random error allows us to estimate the broad density and the detailed density in the same model even though the latter is a subset of the former. This is because the industry-level random error absorbs all the unexplained industry-level variation in entries. Because the detailed density correlates with industry-level factors, that variation would otherwise be included in (i.e. would bias) the coefficient estimate of detailed density. For example, if students

in school A have on average higher test scores than school B, and school A has more girls than school B. A simple model would be unable to determine whether being a girl causes higher scores, or being at school A cause higher scores. A multi-level model can make such determinations by separating the variation in test scores into between-school and between-students variation.

Second, random effects on coefficients allow, for instance, the coefficient estimate of detailed density to vary randomly across industries (normally distributed). Without that random error, if that variation correlates with the moderator variable, that variation would be included in (i.e. bias) the coefficient estimate of the interaction term (Hox, 2010).

RESULTS

Table 1 shows the correlations among the variables and summary statistics. The broad and detailed density are highly correlated. This means that the random errors we add are needed. Typicality's correlation with the density measures is substantial as expected, but not extreme. This means that making typicality non-time varying was necessary and effective in separating these related variables.

Table 2 shows how our focal variables interact with the effect of density on the likelihood of new entries. Specifically, new entries into a 4-digit industry using a particular legal form (column 1), or within a particular firm size range (column 2), or from a particular home country (column 3).

TABLE 1
Descriptive Statistics and Correlations

a) Legal Form

Variable	mean	s.d.	min	max	1	2	3	4	5
1. Entries	1.21	6.13	0	250					
2. Industry Density	76.82	212.97	2	2557	.35				
3. Detailed Density	11.26	39.06	2	1407	.73	.44			
4. Specific-Level Variety	4.65	3.04	1	18	.21	.50	.23		
5. Typicality	.43	1.45	.0012	1	-.01	-.07	.07	-.12	
6. Taxonomy	6.60	3.92	1	17	-.04	-.05	-.04	-.05	-.01

b) Firm Size

Variable	mean	s.d.	min	max	1	2	3	4	5
1. Entries	1.21	4.43	0	162					
2. Industry Density	63.19	163.42	2	2557	.65				
3. Detailed Density	7.04	19.41	0	477	.70	.78			
4. Specific-Level Variety	5.18	2.70	1	15	.21	.35	.29		
5. Typicality	.20	.18	.0034	1	-.01	-.16	.03	-.33	
6. Taxonomy	6.41	4.05	1	17	-.06	-.05	-.03	.01	-.00

c) Home Country

Variable	mean	s.d.	min	max	1	2	3	4	5
1. Entries	.81	3.55	0	157					
2. Industry Density	105.87	271.02	2	2557	.29				
3. Detailed Density	4.58	14.72	0	590	.67	.36			
4. Specific-Level Variety	9.17	7.51	1	53	.21	.54	.32		
5. Typicality	.14	.15	.0004	1	.03	-.22	.07	-.34	
6. Taxonomy	5.45	4.00	1	17	-.05	-.07	-.03	.04	-.01

TABLE 2
Results of Multi-Level Ordinary Least Squares Analysis of Entries into a Subcategory^{a,b}

Variables	Legal Form		Firm Size		Home Country	
Industry Density	-.0025	(.0038)	.0063	** (.0020)	.0005	* (.0003)
Industry Density ²	.0000	** (.0000)	.0000	(.0000)	-.0000	(.0000)
Detailed Density	.0216	(.0272)	.0018	(.0342)	.1051	* (.0548)
Detailed Density ²	.0001	* (.0000)	-.0001	** (.0001)	-.0001	(.0001)
Ind. Density*Det. Density	-.0005	** (.0001)	-.0013	** (.0002)	-.0017	** (.0002)
Specific-Level Variety	-.1256	(.0755)	-.1224	** (.0318)	-.0180	** (.0075)
Specific-Level Variety ²	.0211	** (.0085)	.0138	** (.0043)	.0008	** (.0002)
Typicality	.0252	(.0653)	-.6377	(.5376)	-.1655	(.5919)
Typicality ²	-.0043	** (.0016)	2.13	** (.944)	2.26	** (1.06)
Ind. Density*S.-l. Variety	-.0004	** (.0002)	-.0006	** (.0003)	-.0000	** (.0000)
Det. Density*S.-l. Variety	.0109	** (.0039)	.0280	** (.0049)	.0110	** (.0046)
Ind. Density*Typicality	.0197	** (.0070)	.0050	(.0077)	-.0068	(.0237)
Det. Density*Typicality	-.0029	(.0037)	-.0566	(.0489)	-.2624	** (.0577)
Taxonomy	.0101	(.0156)	.0066	(.0080)	.0031	(.0045)
Industry Density*Tax.	.0001	(.0004)	-.0000	(.0002)	-.0001	* (.0000)
Detailed Density*Tax.	-.0047	** (.0013)	-.0062	** (.0015)	-.0059	** (.0016)
Year fixed effects	Incl.		Incl.		Incl.	

^a Robust standard errors are in parentheses. Significance denoted as if this were not complete population data.

^b With additional random errors at the industry and detailed category levels, and an industry-level random error on the coefficients of detailed density and the interaction between the broad and detailed density.

* $p < .05$

** $p < .01$

Hypothesis 1 was that the specific-level variety of a broad category would increase the saliency of the detailed category relative to the broad category. Specific-level variety has a more positive interaction effect with the detailed density than with the broad density in all three datasets. Thus, as specific-level variety increases, the detailed category becomes a more important cause of imitation relative to the broad density. Therefore, hypothesis 1 is supported. (Because we have complete population data, all differences are significant with $p\text{-value}=0$). Hypothesis 2 was that

the typicality of a detailed category would increase the saliency of the broad category relative to the detailed category. Typicality has a less negative interaction with broad density than with detailed density, so hypothesis 2 is also supported.

TABLE 3
Effect size of interactions in terms of number of entries

a) One standard-deviation increase in Specific-level Variety

Effect of one standard deviation increase in	Broad Density	Detailed Density
1. Legal Form	-0.2	+1.3
2. Firm Size	-0.3	+1.5
3. Home Country	-0.1	+1.2

b) One standard-deviation increase in Typicality

Effect of one standard deviation increase in	Broad Density	Detailed Density
1. Legal Form	+6.1	-0.2
2. Firm Size	+0.1	-0.2
3. Home Country	-0.3	-0.6

The coefficients are quite small and p-values are not informative, so we should calculate effect sizes to get a better sense. Table 3 shows how the effect on entries of a one-standard-deviation increase in a density changes if a moderator increases by one standard deviation. In the legal form data, if the specific-level variety increases by one standard deviation, a one-standard-deviation increase in industry density predicts approximately 0.2 fewer entries, while a one-standard-deviation-increase in detailed density predicts 1.3 more entries. This was calculated by multiplying the coefficient on the respective interaction term (from table 2) with the standard deviation of specific-level variety and with the standard deviation of the respective density (from Table 1). Similarly, in the firm size (home country) data, industry density leads to 0.3 (0.1) fewer entries, and detailed density leads to 1.5 (1.2) more entries, if specific-level variety is one standard deviation greater. The standard deviation of the dependent variable, entries, is approximately 6, 4,

and 3, for the legal form, home country, and size data, respectively. Thus, the effect sizes of the interaction term vary from 25% to 40% of the standard deviation of the dependent variable.

For a one-standard-deviation-increase in typicality, these numbers are: a 1SD-increase in industry density predicts 6.1, 0.1, and -0.3 more entries, while a 1SD-increase in detailed density predicts 0.2, 0.2, and 0.6 fewer entries, in the legal form, firm size, and home country data, respectively. As predicted, typicality favors the broad density in each case. In the home country data, typicality favors industry density by a seemingly small 0.3 entries, but this is still 10% of the standard deviation of the dependent variable. (Again, these differences have $p\text{-value}=0$ because of complete population data.) In the legal form data, the effect size of the interaction term is approximately 100% of the standard deviation of the dependent variable.

One way to interpret these numbers is that the theory (i.e. the moderating variables it suggests) helps us predict anywhere from 0.3 up to 6 entries that we would have otherwise not predicted from the same information about density. Even with only two variables, that is already a substantial fraction of yearly entries (1.2 entries per year on average; 99% of the observations show less than 20 entries per year in a subcategory).

Robustness checks

The distribution of the residuals suggests that a negative binomial model is the best model specification for the data. Unfortunately, random effects in a negative binomial model turned out to be computationally intractable. We present the multi-level linear model's results as the main results. To assess the effect of not using a negative binomial model, we compared how estimates from our variables in our data differed between OLS and Poisson models (of which negative binomial models are a subset). A multi-level Poisson model, not presented here, on a reduced sample with only an additional industry-level random error and additional industry-level random

error on the coefficient of detailed density gives coefficients that are similar in their relative strength to the coefficients reported for the full model. One caveat is that the marginal effects associated with these coefficients vary non-linearly over the ranges of all the other variables. A multi-level OLS model with the same reduced specifications estimates substantively similar coefficients.

We also log-transformed the dependent variable to make the distribution of the residuals more suitable to OLS. These results support the hypotheses in all six cases. However, the distribution of the residuals does not resemble standard normal more than before the transformation. We also ran models that excluded subcategories with more than 20 entries. This prevented many large residuals, forcing the distribution to resemble standard normal. Results from these models also completely supported the hypotheses.

Our approach relies on the idea that the correlation between density and entries captures the extent of imitation. What are the concerns about the validity of this premise, and what happens if it is false? The main concern is the possibility of an alternative explanation of the correlation in terms of a common cause. For example, low profitability of an action means a low likelihood of that action in the past, causing a low density, as well as in the future, causing few new entries. Conversely, high profitability means a high likelihood both in the past and future. As a result, density and entries could be correlated without imitation. Fortunately, our hypotheses are not about explaining any such correlation, but about moderators of the correlation. However, such alternative forces could influence our results. If the correlation between past and new entries is partly driven by a common cause such as profitability, that would make the estimates of the moderator effects more noisy, but still unbiased. Therefore, this alternative explanation is not enough to threaten the validity of our results. However, the alternative explanation may be extended to become

threatening to the moderator effects. For instance, if a particular legal form is more profitable than other legal forms, its share of the total industry density would also be high. That share was our measure of typicality. Therefore, it is plausible that profitability causes the dependent variable 'new entries', the independent variable 'density', and the moderator 'typicality'. In this explanation, typicality does not change the effect of the independent variable on the dependent variable, so there is no moderation. However, statistically, moderators are interaction terms. Interaction terms have a positive coefficient when the simultaneous movement of the interacted variables (e.g. density and typicality) correlates with the dependent variable over and above the effects of each individual interacted variable. We can imagine a scenario in which the alternative explanation causes a significant interaction effect in this way. Namely, in cases where density and typicality are high simultaneously, profitability is probably also high because it is causally linked to both. In addition, this high profitability causes new entries to be high. Conversely, in cases where density is high while typicality is low (or vice versa), profitability must be having a weak influence; otherwise typicality would have been high. Accordingly, profitability should also have a weak influence on new entries. Therefore, the dependent variable correlates with profitability especially in cases where the independent variable and the moderator move simultaneously. As a result, entries will be especially high if the interacted variables are high simultaneously, and entries will be especially low if the interacted variables are low simultaneously. Thus, the interaction term can be significant without imitation and without our theory. Fortunately, our hypotheses are not about whether moderators are significant, but about whether a moderator affects detailed density more than it affects broad density. The extended alternative explanation involves the profitability of a specific legal form, which should influence the density of that specific legal form. Therefore, the alternative explanation predicts that typicality interacts more strongly with detailed density

than with broad density. However, our theory predicts the opposite. So, this potential alternative explanation is not a threat after all. Analogous potential alternative explanations based on firm size or home country instead of legal form are similarly possible but ultimately opposite.

DISCUSSION

We found and explain differences between industries in terms of the level of detail of salient categories. Our explanation has implications for research on interorganizational imitation. Without our theory, the prediction that firm B will imitate firm A's greenfield entry, still allows for a broad range of behaviors. Firm B's behavior could range from very broad imitation (launching a joint venture in another country and another industry) to very detailed (renting office space one floor above firm A's offices). This range of behaviors differs only a little from the range which could be predicted if imitation were not considered at all. Our theory facilitates predictions about the level of detail in imitation. This allows researchers to further delimit the range of their prediction of imitators' behavior.

By extension, the results also have implications for research on the diversity of populations of organizations, because diversity is as a result of imitative behavior by many organizations (DiMaggio & Powell, 1983; Greve, 2005; Terlaak & Gong, 2008). A number of studies have been interested in estimating how the strength of imitation differs between levels of detail (Cattani, Pennings, & Wezel, 2003; Chan, Makino, & Isobe, 2006; Durand & Vergne, 2014; Garcia-Pont & Nohria, 2002; Hannan et al., 1995; Henisz & Delios, 2001; Li et al., 2007; Porac et al., 1995; Swaminathan & Wiedenmayer, 1991). By introducing variables that predict those differences, we developed such interest from an empirical exercise into a theory. For example, Bigelow, Carroll, Seidel, and Tsai (1997) estimated the relative effects of the national and regional densities of car manufacturers in the U.S. on new findings. To their surprise they found that the regional density

is not a significant predictor in the Midwest region (1997: 393–395). The framework we developed helps explain that surprising finding: the Midwest region is the most typical region for car manufacturing in the U.S., and typicality leads to broader categorization, and hence a less influential detailed density.

Considering only one level of detail at a time may lead to results that are spurious due to an adjacent level of detail's effect (Guillén, 2002, 2003). In addition, density on the generic level may have a substitutive relationship with density on the specific level (Rossman, 2014). We dealt with those effects by including additional random effects and an interaction between the generic and specific density. The results are sensitive to removing either, suggesting that those factors are significant. In addition, the theory allows researchers to form expectations about the extent of these effects. For instance, in case of high specific-level variety, detailed categories are preferred, so broad density is less likely to substitute for detailed density. For highly typical detailed categories, broad categories are preferred relative to detailed categories, so the substitutive relationship may be stronger.

The concepts of specific-level variety and typicality are similar to contrast and typicality (as defined in Hannan et al. (2007)), and are different in the same way. Contrast is based on the similarity of all existing members to the hypothetical perfect member (i.e. an object whose attributes match completely with the attributes of the category) (Hannan et al., 2007). Specific-level variety is based on the similarity of all existing members to each other (in line with Rosch et al.'s (1976) framing). Thus, we used specific-level variety instead of contrast because it was more in line with the theory we draw on. However, contrast should have an effect similar to that of specific-level variety. This is not the case for typicality. Hannan and colleagues (2007) define typicality in terms of how many attributes overlap with a hypothetical perfect category member.

Rosch and colleagues (1976) define typicality in terms of how many attributes overlap with all other existing members. In an additional analysis we can see how different definitions of typicality imply different predictions. We divided the broad category into geographical units using France's 101 départements rather than, for instance, legal form. We did not find support for the typicality hypothesis in that analysis (though the results for specific-level variety did support the hypothesis). For example, the département that includes Bordeaux is considered highly typical in the wine industry. Our hypothesis suggested that this high typicality induces broad categorization. Concretely, foreign wineries entering Bordeaux should relate to the density of wineries over the whole of France rather than the density of wineries in Bordeaux's département (compared to an atypical département which should be relatively more based on their local density). We find the opposite: the number of entries of wineries in Bordeaux is especially strongly correlated with density in Bordeaux specifically rather than in France as a whole. Over all industries, typicality works opposite of prediction for detailed categories based on the départements. Another example from our data is the density of banks in Paris being especially strongly correlated with bank entries into Paris specifically. It makes intuitive sense that organizations attend to the region specifically. Indeed, the Bordeaux wine region and Parisian banking are 'typical' especially because they are not like other regions; they are uniquely well-suited for their industry. This uniqueness means geographic units are not typical in the strict theoretical sense of 'greatest overlap in attributes with other geographic units'. These cases are typical according to Hannan et al.'s definition (2007), but not according to Rosch et al.'s definition (1976). Thus, the assumptions discussed in footnote 1 seem to be violated in this case. In this way, confounding Hannan et al.'s (2007) and Rosch et al.'s (1976) definitions may cause the results to appear counterintuitive, but that does not imply that the assumption is violated for all studies using geographic divisions. For instance, Bigelow and

colleagues (1997) found regional differences in the relative effects of national and regional densities in the U.S. automobile industry that are consistent with our typicality hypothesis.

Future work in related literatures might usefully apply the framework we used. For example, research on how density (the number of firms in a population) can be weighted by firm size to create a measure of population mass (the sum of all firms' size in a population), which then drives legitimation and competition (Baum & Mezias, 1992; Cattani et al., 2003). Our theory predicts that weighted densities are useful when categorization is detailed and weights are not useful when categorization is broad. One especially burgeoning type of weight is the degree of category membership. Organizations with a higher degree of membership in a category contribute more to the legitimization of that category than organizations with low degree of membership (Kuilman & Li, 2009). We suggest an additional explanation for why that should be. Organizations with low membership may be less typical (this is true by definition in Hannan et al.'s (2007) framework, but may also be true by correlation in Rosch et al.'s (1976) framework). If so, such organizations are categorized with more detail, and contribute primarily to their subpopulation's legitimacy instead of the whole population's legitimacy.

Our perspective may also add insight into the weaker effect of lenient categories (Pontikes, 2012). Lenient categories are those that overlap with many other categories, which means they are highly typical by Rosch et. al's (1976) definition. We find that highly typical subcategories tend to be overlooked in favor of the associated broad category. Therefore, organizations or their actions in a lenient subcategory may be evaluated with regard to their fit in the broader category instead of within the lenient subcategory. That would make an organization's fit with a lenient subcategory less influential.

Imitation is only one behavior through which researchers can observe the effects of social categories. Our theory should be applicable if categories are observed in a different way. One such way is through their effect on the likelihood of being evaluated (Zuckerman, 1999). For example, we could replace ‘density (high vs. low)’ with ‘position in category space (category spanning vs. unambiguous position)’ and replace ‘new entries’ with ‘discount by the audience’. Kovacs and Hannan (2010) hypothesize that the effect of category spanning on discount is stronger if category contrast is greater. The results of our study suggest an additional possible mechanism for that relationship. Specific-level variety is low when contrast is high (though the reverse is sometimes not true). We found that if specific-level variety is low, the broad category is salient instead of the detailed categories. Spanning categories that are salient should lead to higher discount than if those same categories are not salient. For example, take-out retailers may be seen as spanning the categories ‘supermarket’ and ‘restaurant’. However, if audiences find those categories too broad, the salient category may be ‘take-out retailer’. In that case, take-out retailers do not span relevant or salient categories, so no discount is observed.

Our findings also relate to the partial adoption of innovations and adoption with customization or adaptation (Ansari, Fiss, & Zajac, 2010; Westphal, Gulati, & Shortell, 1997). We would frame this phenomenon as imitating an innovation in a broad or a detailed way. Broad imitation allows more room to choose the details. In contrast, if the detailed category is salient, organizations have little room to customize an innovation without falling outside of the category. This framing suggests that adoption customization is more likely if specific-level variety is low (i.e. there are few different subtypes of the innovation), and if the innovation is typical of its group of innovations.

A final future application of the framework we present is in finding antecedents of informativeness or distinctiveness that cause differences between audience segments. One antecedent that the psychology literature has identified is expertise. Expertise of categorizers has been found to increase the distinctiveness they perceive categories to possess. Consequently, expertise causes a stronger preference for detailed categories (Tanaka & Taylor, 1991). This could help explain the divergent findings about audience segments of very active evaluators (Kovács & Hannan, 2010), or the findings about changes in evaluations over time as audiences become more familiar with a new domain (e.g. Rao, Monin, & Durand, 2005), or finding that the relevant peer pressure shifts from home country competitors to business group peers as a firm gains international experience (Guillén, 2002).

Apart from demonstrating hypothesized correlations, we also found a pattern in the effect of the 'taxonomy' variable (the number of other broad categories sharing a higher level group). The detailed density has a smaller effect on entries if its broad category is part of a larger higher level group. In a sense, this is the broader variant of specific-level variety - generic-level variety. We can explain why generic-level variety would strengthen the broad category relative to the higher level group, though why this also makes broad categories favored relative to detailed categories is difficult to argue. One possibility is that taxonomy correlates with specific-level variety because if there are many broad categories, there are fewer differences remaining for the detailed categories to order. However, our correlation table shows that such a correlation does not exist, and because specific-level variety is also in the model, we have inadvertently controlled for this explanation.

There may be firms or managers for which our arguments do not hold. The influence of categories, and imitation in general, on market entry decisions may vary at the firm-level (with,

for example, firm size), or among managers (based on, for example, manager narcissism). Our results show this influence at the average of such variation.

Empirical studies on levels of detail in imitation must beware of how imitation pressures may be strengthened or weakened differently on different levels. For example, the pressure to imitate a foreign entry is counterbalanced by a differentiation pressure, because imitating foreign market entry increases competition in that market. But this counterbalancing does not apply to imitating specific modes of entry (keeping constant the total number of foreign entries) (cf. Li et al., 2007; Lieberman & Asaba, 2006). This asymmetry causes imitation at the generic level to be counterbalanced more than imitation at the specific level. This gives the impression that imitation shifts to a more specific level of detail while really it is the degree of imitation which changes asymmetrically. Our strategy to mitigate this issue is to focus on only foreign entry. The number of domestic firms should be the dominant driver of competition, and because that number is not included in the count of foreign firms, we create noise between the number of foreign firms and competition such that they are less systematically correlated.

We aim to apply Rosch et al.'s (1976) theory to the context of organizations. Our results support this application for the aspect of interorganizational imitation, but that should not be interpreted as implying that the theory applies to all aspects of organizations. In general, the theory relies on laboratory experiments. The context of organizations may be too different to generalize their theory to. One such difference may be the importance of rhetoric. Rhetoric invites conscious, deliberate, instrumental categorization (choosing the categorization that is most likely to convince the audience). So it might encourage categorization tending toward the generic level of detail to prevent detailed scrutiny (Porac, Wade, & Pollock, 1999). Next to this, if informativeness and cognitive cost have great (monetary) value, or if the time period over which the categorization

takes place is long, then categorization may be calculative and instrumental instead of a non-conscious tendency. The main goal of empirical study of this topic should be to find out to what extent such factors are relevant concerns. For example, future research could study the effect of time-pressure on audience evaluations of category spanning at different levels of detail. Barring alternative explanations, we could not have found support for our hypotheses if those factors were empirically influential concerns. Thus, we believe our work helps move toward this main goal.

Conclusion

We have developed a theory which helps to understand at which level of detail in a taxonomy categories have the greatest saliency. Our results show that the effect of a social category depends on its position in the taxonomy and the taxonomy's structure (i.e. specific-level-variety and typicality). We aimed to answer the question: should researchers using social categories in the organizational context complement their theory with insights from research in psychology on levels of detail (as summarized by Murphy (2002))? Our results should move the consensus toward 'yes', because (part of) this theory formed the basis for accurate predictions of (hitherto unexplained) variation in organizational actions.

REFERENCES

- Aharoni, Y. 1966. *The foreign investment decision process*. Boston: Harvard University Press.
- Ansari, S. M., Fiss, P. C., & Zajac, E. J. 2010. Made to Fit: How Practices Vary As They Diffuse. *Academy of Management Review*, 35(1): 67–92.
- Baum, J. A. C., & Mezias, S. J. 1992. Localized Competition and Organizational Failure in the Manhattan Hotel Industry, 1898-1990. *Administrative Science Quarterly*, 37(4): 580–604.
- Berlin, B., Breedlove, D. E., & Raven, P. H. 1973. General Principles of Classification and Nomenclature in Folk Biology. *American Anthropologist*, 75(1): 214–242.

- Bigelow, L. S., Carroll, G. R., Seidel, M.-D. L., & Tsai, L. 1997. Legitimation, Geographical Scale, and Organizational Density: Regional Patterns of Foundings of American Automobile Producers, 1885–1981. *Social Science Research*, 26(4): 377–398.
- Carroll, G. R., & Hannan, M. T. 2000. *The Demography of Corporations and Industries*. Princeton, NJ: Princeton University Press.
- Cattani, G., Pennings, J. M., & Wezel, F. C. 2003. Spatial and Temporal Heterogeneity in Founding Patterns. *Organization Science*, 14(6): 670–685.
- Chan, C. M., Makino, S., & Isobe, T. 2006. Interdependent behavior in foreign direct investment: the multi-level effects of prior entry and prior exit on foreign market entry. *Journal of International Business Studies*, 37(5): 642–665.
- DiMaggio, P. J., & Powell, W. W. 1983. The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, 48(2): 147–160.
- Dobrev, S. D., Ozdemir, S. Z., & Teo, A. C. 2006. The Ecological Interdependence of Emergent and Established Organizational Populations: Legitimacy Transfer, Violation by Comparison, and Unstable Identities. *Organization Science*, 17(5): 577–597.
- Dougherty, J. W. D. 1978. Saliency and relativity in classification. *American Ethnologist*, 5(1): 66–80.
- Durand, R., & Vergne, J.-P. 2014. Asset Divestment As A Response To Media Attacks In Stigmatized Industries. *Strategic Management Journal*, n/a–n/a.
- Dutton, J. E., & Jackson, S. E. 1987. Categorizing Strategic Issues: Links to Organizational Action. *The Academy of Management Review*, 12(1): 76–90.
- Fligstein, N. 1985. The Spread of the Multidivisional Form Among Large Firms, 1919-1979. *American Sociological Review*, 50(3): 377–391.
- Garcia-Pont, C., & Nohria, N. 2002. Local versus global mimetism: the dynamics of alliance formation in the automobile industry. *Strategic Management Journal*, 23(4): 307–321.
- Greve, H. R. 2000. Market Niche Entry Decisions: Competition, Learning, and Strategy in Tokyo Banking, 1894-1936. *The Academy of Management Journal*, 43(5): 816–836.
- Greve, H. R. 2005. Interorganizational Learning and Heterogeneous Social Structure. *Organization Studies*, 26(7): 1025–1047.

- Guillén, M. F. 2002. Structural Inertia, Imitation, and Foreign Expansion: South Korean Firms and Business Groups in China, 1987-95. *The Academy of Management Journal*, 45(3): 509–525.
- Guillén, M. F. 2003. Experience, Imitation, and the Sequence of Foreign Entry: Wholly Owned and Joint-Venture Manufacturing by South Korean Firms and Business Groups in China, 1987-1995. *Journal of International Business Studies*, 34(2): 185–198.
- Hannan, M. T., Carroll, G. R., Dundon, E. A., & Torres, J. C. 1995. Organizational Evolution in a Multinational Context: Entries of Automobile Manufacturers in Belgium, Britain, France, Germany, and Italy. *American Sociological Review*, 60(4): 509–528.
- Hannan, M. T., Pólos, L., & Carroll, G. R. 2007. *Logics of organization theory: audiences, codes, and ecologies*. Princeton University Press.
- Haveman, H. A. 1993. Follow the Leader: Mimetic Isomorphism and Entry Into New Markets. *Administrative Science Quarterly*, 38(4): 593–627.
- Henisz, W. J., & Delios, A. 2001. Uncertainty, Imitation, and Plant Location: Japanese Multinational Corporations, 1990-1996. *Administrative Science Quarterly*, 46(3): 443–475.
- Hox, J. J. 2010. *Multilevel Analysis: Techniques and Applications*. Routledge.
- Kennedy, M. T., & Fiss, P. C. 2013. An Ontological Turn in Categories Research: From Standards of Legitimacy to Evidence of Actuality. *Journal of Management Studies*, 50(6): 1138–1154.
- Kovács, B., & Hannan, M. T. 2010. The consequences of category spanning depend on contrast. *Categories in Markets: Origins and Evolution*, vol. 31: 175–201. Emerald Group Publishing Limited.
- Kuilman, J. G., & Li, J. 2009. Grades of Membership and Legitimacy Spillovers: Foreign Banks in Shanghai, 1847–1935. *Academy of Management Journal*, 52(2): 229–245.
- Li, J., Yang, J., & Yue, D. 2007. Identity, Community, and Audience: How Wholly Owned Foreign Subsidiaries Gain Legitimacy in China. *The Academy of Management Journal*, 50(1): 175–190.
- Lieberman, M. B., & Asaba, S. 2006. Why Do Firms Imitate Each Other? *The Academy of Management Review*, 31(2): 366–385.

- Markman, E. M. 1985. Why superordinate category terms can be mass nouns. *Cognition*, 19(1): 31–53.
- Mervis, C. B., & Rosch, E. 1981. Categorization of Natural Objects. *Annual Review of Psychology*, 32(1): 89–115.
- Murphy, G. L. 2002. *The big book of concepts*. Cambridge, MA, US: MIT Press.
- Murphy, G. L., & Brownell, H. H. 1985. Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(1): 70–84.
- Navis, C., & Glynn, M. A. 2010. How New Market Categories Emerge: Temporal Dynamics of Legitimacy, Identity, and Entrepreneurship in Satellite Radio, 1990–2005. *Administrative Science Quarterly*, 55(3): 439–471.
- Negro, G., Hannan, M. T., & Rao, H. 2010. Categorical contrast and audience appeal: niche width and critical success in winemaking. *Industrial and Corporate Change*, 19(5): 1397–1425.
- Negro, G., Koçak, Ö., & Hsu, G. 2010. Research on categories in the sociology of organizations. *Categories in Markets: Origins and Evolution*, vol. 31: 3–35. Emerald Group Publishing Limited.
- Pontikes, E. G. 2012. Two Sides of the Same Coin How Ambiguous Classification Affects Multiple Audiences' Evaluations. *Administrative Science Quarterly*, 57(1): 81–118.
- Porac, J. F., & Thomas, H. 1990. Taxonomic Mental Models in Competitor Definition. *Academy of Management Review*, 15(2): 224–240.
- Porac, J. F., & Thomas, H. 1994. Cognitive categorization and subjective rivalry among retailers in a small city. *Journal of Applied Psychology*, 79(1): 54–66.
- Porac, J. F., Thomas, H., Wilson, F., Paton, D., & Kanfer, A. 1995. Rivalry and the Industry Model of Scottish Knitwear Producers. *Administrative Science Quarterly*, 40(2): 203.
- Porac, J. F., Wade, J. B., & Pollock, T. G. 1999. Industry Categories and the Politics of the Comparable Firm in CEO Compensation. *Administrative Science Quarterly*, 44(1): 112–144.
- Rao, H., Monin, P., & Durand, R. 2005. Border Crossing: Bricolage and the Erosion of Categorical Boundaries in French Gastronomy. *American Sociological Review*, 70(6): 968–991.

- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. 1976. Basic objects in natural categories. *Cognitive Psychology*, 8(3): 382–439.
- Rossmann, G. 2014. The Diffusion of the Legitimate and the Diffusion of Legitimacy. *Sociological Science*, 1: 49–69.
- Smith, E., & Medin, D. 1981. *Categories and Concepts*. Harvard University Press.
- Swaminathan, A., & Wiedenmayer, G. 1991. Does the pattern of density dependence in organizational mortality rates vary across levels of analysis? evidence from the German brewing industry. *Social Science Research*, 20(1): 45–73.
- Tanaka, J. W., & Taylor, M. 1991. Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, 23(3): 457–482.
- Terlaak, A., & Gong, Y. 2008. Vicarious Learning And Inferential Accuracy in Adoption Processes. *Academy of Management Review*, 33(4): 846–868.
- Thomas, J. B., Clark, S. M., & Gioia, D. A. 1993. Strategic Sensemaking and Organizational Performance: Linkages among Scanning, Interpretation, Action, and Outcomes. *The Academy of Management Journal*, 36(2): 239–270.
- Vergne, J.-P., & Wry, T. 2014. Categorizing Categorization Research: Review, Integration, and Future Directions. *Journal of Management Studies*, 51(1): 56–94.
- Westphal, J. D., Gulati, R., & Shortell, S. M. 1997. Customization or Conformity? An Institutional and Network Perspective on the Content and Consequences of TQM Adoption. *Administrative Science Quarterly*, 42(2): 366.
- Wry, T., & Lounsbury, M. 2013. Contextualizing the categorical imperative: Category linkages, technology focus, and resource acquisition in nanotechnology entrepreneurship. *Journal of Business Venturing*, 28(1): 117–133.
- Wry, T., Lounsbury, M., & Jennings, P. D. 2013. Hybrid Vigor: Securing Venture Capital by Spanning Categories in Nanotechnology. *Academy of Management Journal*, amj.2011.0588.
- Yiu, D., & Makino, S. 2002. The Choice between Joint Venture and Wholly Owned Subsidiary: An Institutional Perspective. *Organization Science*, 13(6): 667–683.
- Zuckerman, E. W. 1999. The categorical imperative: Securities analysts and the illegitimacy discount. *American Journal of Sociology*, 104(5): 1398–1438.

Chapter 4

The Heterogeneous Diffusion of Money

ABSTRACT

Diffusion of new ideas has attracted the attention of management scholars. Our key insight is that money also diffuses, but the diffusion process is qualitatively different. This allows us to understand the boundary conditions of known diffusion processes, as well as understand wealth differences as a function of known antecedents of diffusion. The current explanation of wealth differences is based on a metaphor of distribution. We argue that diffusion is a richer metaphor than distribution in terms of accurate predictions it suggests about wealth differences. We test these ideas on data of money flows between soccer clubs transferring players.

INTRODUCTION

‘Trickle down’ is a well-known but ill-understood metaphor for one process underlying wealth differences. We aim to see if we can better understand this process by conceptualizing it as a diffusion process. In other words, we draw an analogy between how innovations diffuse over a group of organizations and how money trickles down a group of organizations. Then, we see if the diffusion conceptualization suggests new variables that influence the trickle-down process. Those variables should be new relative to the literature on the velocity of money, which is the most related literature in terms of the phenomenon. That literature focuses on macro-level variables, thereby overlooking characteristics of organizations. In contrast, the literature that we draw from has a strong connection to research on organizations, so we can propose a number of new variables.

In terms of theory, our application is also interesting to diffusion scholars because it highlights boundary conditions to existing work as well as shows how diffusion ideas work beyond these boundaries. ‘Existing work’ here refers to research on the diffusion of information, knowledge, innovation, and management practices between organizations (Abrahamson & Fairchild, 1999; Baum, Li, & Usher, 2000; Burns & Wholey, 1993; Fligstein, 1985; Greve, 2005; Haunschild & Miner, 1997; Haunschild, 1993; Kennedy & Fiss, 2009; Rao, Davis, & Ward, 2000; Sorenson & Stuart, 2001; Strang & Soule, 1998).

We make this contribution by theorizing and testing how known correlates of diffusion (Strang & Tuma, 1993) operate in the context of money. If those correlates are useful in explaining the data, then our study has improved the understanding of wealth differences and the boundary conditions of diffusion models. Data on soccer clubs’ transfer spending allow us to track not only each instance of a transfer of money, but also the subsequent transfer behavior of the receiver of that money.

First we discuss in more detail the differences and similarities between the diffusion of money and various other materials (e.g. innovations) studied by diffusion scholars. These differences are the boundary conditions that are crossed by moving to the context of money. Second, we go over main correlates of diffusion and explain one by one how their influence is affected by these boundary conditions. This serves as hypothesis development. Third, we discuss the data in more detail. Fourth, we discuss the results of the hypothesis testing. Finally, we discuss the implications for diffusion research and our understanding of wealth differences.

APPLYING THE DIFFUSION MODEL

In this section, we explain how insights from research on diffusion can be used to understand the process underlying wealth differences, and also how diffusion of money differs from diffusion of other materials.

Conceptualizing the phenomenon in diffusion terms

The phenomenon we focus on is coins and bills ‘changing hands’, and it is digital money being subtracted and added on different bank accounts. We may conceptualize this phenomenon using the metaphor of diffusion. It may help understand the usefulness of the diffusion conceptualization to consider why an alternative conceptualization is less useful. Consider the alternative metaphor called ‘distribution’, meaning that a central sender allocates all wealth to all others (i.e. receivers) based on some allocation rule. For example, the centralized state gives welfare to all citizens below some poverty line. Importantly, the state is in contact with all potential claimants of welfare. A receiver’s fit with allocation rules helps predict its wealth only if the centralized sender is in contact with all actors. In contrast, diffusion is a more accurate metaphor if senders are not in contact with all other actors in the whole network. Instead, each actor independently allocates their spending to their direct contacts (i.e. transaction partners), and those

contacts then independently further allocate the wealth they have received. Thus, the diffusion metaphor goes beyond the receivers' characteristics in predicting their wealth. It highlights how the degree of contact with a sender moderates the effect of the receiver's characteristics. In addition, the receiver's wealth is further moderated by characteristics of the sender(s) with whom they are in contact.

Diffusion models are often used to study the diffusion of an idea or practice before all actors in a population have become aware of it (termed saturation). Under this interpretation, the diffusion of money would mean that we track when an actor gains access to money for the first time in their life, which is only plausible when studying newly introduced currencies. This is not what we study. When diffusion materials can be lost (e.g. people forget the idea, or are cured of an infection), diffusion models can also be used to study the diffusion of ubiquitous materials (e.g. influenza (Viboud et al., 2006)). An actor loses the money he or she spends, so money qualifies. Such a use of diffusion logic investigates the same mechanisms as currency adoption but with a different empirical approach.

The empirical setting is about soccer players transferring to new clubs for amounts so inordinate that they could never reflect some true economic value of the player. This disconnect with value makes it less plausible to argue that it is the players who diffuse and that the money flows are a side-effect. Such a disconnect was also found in the market for Champagne grapes (Ody-Brasier & Vermeulen, 2014). Soccer clubs also get money from outside of the focal system (i.e. player transfers) from ticket sales, investors, sponsorship and so on. They also release money to outside the world of soccer transfers via, for example, stadium mortgage payments, and wage payments. But the diffusion of money throughout the transfer system is our focus.

To appreciate the usefulness of the diffusion model conceptualization, consider again the contrasting conceptualization of ‘distribution by allocation rule’. Distribution would mean that money is assigned (e.g. by FIFA) to clubs based on the quality of players leaving the club like welfare benefits are assigned to citizens based on pre-specified characteristics. Clearly, the distribution model does not match with what actually happens, and is therefore not appropriate for soccer transfers. More generally, the distribution conceptualization is the default model to think about the division of wealth (because it is mostly macro-economists that think about the division of wealth), but the diffusion model may be more informative and more accurate in many settings. We aim to show that soccer transfers is one such setting.

We can imagine the diffusion process starting from the infusion of wealth into the transfer system from the outside. For example, a wealthy investor donates money to Chelsea F.C. for transfers. Chelsea can now demand more players than otherwise. The demand curves of clubs selling players to Chelsea shifts up. Such a shift means that those sellers can now ask a higher price for the same player. After a transfer and Chelsea, having spent the donation, is back to being as wealthy as before. The selling club received a higher price such that they can now themselves demand more players than they otherwise could have. The wealth, or in other words, the ability to demand, has diffused from Chelsea to the selling club. The interesting part of the process is to predict where that wealth will go next. Which club will experience an increase in wealth next? Perhaps, the initially selling club is interested in a player from Chelsea, causing Chelsea’s demand curve to shift up, and after that new transfer, Chelsea is back in possession of the extra wealth. Or perhaps the selling club uses half of their new wealth to bid extra for two new players. The sellers of both those players then experience an increase in their demand curve, and ultimately wealth, equal to half the original donation. Or perhaps the selling club uses its new wealth for mortgage

payments and the new wealth leaves the transfer system. We can use characteristics of sellers (i.e. susceptibility) and buyers (i.e. infectiousness) and their relationship (i.e. social proximity) as a basis for predictions about these questions.

Crossing boundary conditions

When applying the diffusion model to the money context, we risk applying it outside its scope. Specifically, some of the premises involved in the knowledge-context-manifestation of the diffusion model may not hold. We focus on three collections of such premises (i.e. boundary conditions) that are crossed when moving to the money context: 1) behavior by adopters, 2) mode of diffusion, and 3) zero-sum transfers. The diffusion model itself is the same at its core, but the money-context-manifestation of the model differs in these ways from the knowledge-context-manifestation.

First, different materials' diffusion requires different behavior by adopters. There are a wide range of possible materials that diffuse. For instance, innovation, diseases (Viboud et al., 2006), baby names (Berger & Le Mens, 2009), or money. Greve (2005) provides a framework for research on the diffusion of information, innovation, knowledge, and experience. All of these diffusion materials often require actors to actively decide to adopt actions that build on the information, knowledge, or experience. If actors decide to not apply the information in their actions, the information may not diffuse further, because other actors can mostly observe only actions. For example, if an organization knows about a diffusing management practice but does not implement it, its competitors are less able to infer knowledge about the practice from observing the organization's behavior. In contrast, another material like a viral infection does not require decisions or actions by senders or receivers for the infection to diffuse further. Instead, it requires the behavior of physical contact to diffuse. The correlates of diffusion identified by Greve often

operate on decisions and actions. For example, low organizational performance enhances diffusion because it fosters a willingness to act on the diffusing information (Greve, 2005: 1028; Levitt & March, 1988). So, if the behavior involved in diffusion differ between materials, then the effect of correlates of diffusion should also differ.

Analogous to the way diffusion of knowledge requires actors to adopt innovative practices, the diffusion of money requires actors to spend; to demand products or services. However, spending involves different considerations by senders and receivers than adopting innovations. For example, adopting an innovation involves more uncertainty, risk, and effort. As a result, the correlates of diffusion of knowledge that operate via the behavior of organizations may or may not work as correlates of diffusion of money. For example, performance may affect the extent to which diffusing information is used, but its effect on the extent to which money diffuses has not been studied. In sum, the difference in behavior is one of the boundary conditions that separate the knowledge-context-manifestation of the diffusion model from a money-context-manifestation of a diffusion model.

A second boundary condition is the mode of diffusion. Commonly discussed modes of diffusion in the context of innovation are: 1) observing how frequently an innovation is used (by high status firms), 2) observing the outcomes of other organizations' adoption (Haunschild & Miner, 1997), 3) information spillovers, and 4) knowledge sharing. These modes of diffusion are called transfer modes because all of them involve the material transferring from one actor to another. Next to transfer, organizations can also come into contact with new knowledge through search (Sorenson, Rivkin, & Fleming, 2006). Search means organizations independently discover the knowledge from its original source. Search falls within the propensity part of heterogeneous diffusion models, while transfer is represented by the susceptibility, social proximity, and

infectiousness parts (Strang & Tuma, 1993). The key insight of Sorenson and colleagues' (2006) is that the search mode of diffusion substitutes for transfer more often if knowledge is less complex. Likewise, within the transfer modes, observation and spillover are likely substitutes for knowledge sharing if knowledge is less complex (Hansen, 1999). Thus, characteristics of the diffusion material determine the diffusion mode (search, observation, spillover, sharing, or something else entirely). More importantly, the effect of correlates of diffusion depend on the mode of diffusion. So, if materials differ in their mode of diffusion, then the effect of correlates are also likely to differ.

Applying these insights, money's characteristics suggests 'transaction' as the most likely diffusion mode. Other transfer modes do not have a valid counterpart in the money context. For example, in contrast to observing knowledge being used, observing someone else's wealth cannot increase your own wealth. Search does have a valid counterpart in the money context: the gains from creating real economic growth are 'discoveries' of new money by the entrepreneurs to whom the profits accrue. In our empirical setting, we constrain the system to include only money from player transfers. Therefore, income from sources other than transfers are also treated as forms of search.

The first and second boundary condition interact. One characteristic of transactions as a diffusion mode is that they require agreement from both the buyer and the seller. In contrast, in the knowledge context, organizations may imitate competitors' innovation in spite of secrecy. This characteristic of transactions implies that the preferences of competitors can affect an organization's susceptibility and its likelihood to adopt. Thus, the increased importance of the competitors' preferences is another difference between the innovation and money contexts.

The third boundary condition is the zero-sum nature of money transfers. It means that I do not lose the knowledge I share with others, but I do lose the money I transact to others. One implication of this is that the amount of money diffusing does not decrease; it does not become diluted. In the knowledge context, dilution refers to the idea that a receiver of knowledge can only diffuse the fraction of the sender's message that they understood. As knowledge passes to more and more actors, each of which imperfectly understands their own sender, the original knowledge quickly becomes diluted. Simultaneously, there is an opposing effect whereby receivers of knowledge may add their own insights to the diffusing knowledge. As a result, more knowledge may diffuse with each successive transfer. In the context of viral infections we know that a full-blown disease has to begin from scratch with a new incubation period every time the virus transfers to another host. In contrast, money stays the same. A more concrete consequence is that while the value of knowledge (or at least, organizations' estimation of that value) varies with the number of organizations that have previously adopted it (Rao, Greve, & Davis, 2001), the value of money does not. (Because each new adoption means an equal loss of money, adoption does not cause inflation). This allows researchers to avoid noise and bias from external processes like legitimation or competition that are alternative explanations for the adoption of innovations in a population (Strang & Soule, 1998). This characteristic of money also interacts with the first boundary condition because the prospect of losing money changes spenders' considerations.

Thus, buyers lose the money they diffuse in transactions. However, on the other side of this coin, buyers get non-monetary value in return. Consequently, the diffusion of money is also simultaneously the diffusion of the product or service that is sold. This characteristic of transactions influences which control variables are required. For example, we perform a robustness check from the perspective of players being sold (in addition to the perspective of money being

transferred). An alternative way to exclude the possibility that products or services are the driving force, we could study money flows in gift giving networks, in which people give money and get nothing in return. However, such a setting has the drawback that money can be substituted by other kinds of gifts or assistance in kind or personal loans. If independent variables of interest correlate with the likelihood of substitution, then their coefficients confound effects on diffusion and effects on substitution. Moreover, a study of money as capability to demand products and services has more policy implications than a study of money as gift.

HYPOTHESES

In this section we aim to show that heterogeneous diffusion is a useful model to explain spending and differences in wealth. The model is useful if it suggests new hypotheses about spending that are supported by the data (Lave & March, 1975). We also aim to show that correlates of diffusion in the context of innovation work differently in the context of money, and specifically that this difference is in line with the boundary conditions that are crossed.

Greve summarizes the heterogeneous diffusion model as follows: “The rate of learning from an origin organization to a destination organization is the product of the infectiousness of the origin organization, the social proximity of the origin and destination organizations, and the susceptibility of the destination organization” (2005: 1027). In the knowledge context, below aspiration performance is used as a covariate to proxy for susceptibility, and similarity is used as a covariate to proxy for social proximity. Accordingly, our application of the diffusion model to the money context involves looking at the effect of below aspiration performance (H2 and H3) and of similarity (H4).

The most basic diffusion insight is that receiving information about an innovation increases the likelihood of adopting it, and thereby diffusing it further. Because information and knowledge

flows are difficult to observe, researchers assume that recent adoption should lead to flows by means of one of the diffusion modes. An organization receiving a flow becomes more likely to adopt. Therefore a high number of recent adoptions close to an organization makes it more likely to adopt (e.g. Haunschild, 1993).

In the money context, this basic insight becomes: receiving money increases the likelihood of spending money. The spending behavior by another organization leads to flows by means of a transaction and the selling organization then becomes more likely to adopt that behavior (i.e. spend money itself). The soccer setting has the convenient property of allowing easy direct measurement of both the flow (i.e. a transfer deal) and the adoption (i.e. using the proceeds to buy new players). We have discussed at length how diffusion is different in the money context, but this baseline hypothesis is a key instance of how core ideas about diffusion are applicable.

H1: The amount of transfer money a club receives in a season is positively associated with that club's transfer spending in that season.

But when is this somewhat obvious statement not true? We use diffusion ideas to think about conditions under which this hypothesis is more or less valid. The heterogeneous diffusion model suggests that H1 may be false for actors with low susceptibility. Susceptibility has two aspects: 1) likelihood of receiving flows, 2) likelihood of adopting behavior once a flow is received. First, we focus on variables that affect an organization's likelihood to adopt the spending behavior given that a flow of money has diffused to it. In other words, which factors influence an organization to spend rather than save money? Such factors are moderators of the effect predicted in H1.

One such factor is below aspiration performance (Audia & Greve, 2006). Below aspiration performance is argued to induce risk-taking. The argument's underlying logic is that barriers to change (e.g. risk aversion) are reduced when performance is below aspiration level, so organizations are more willing to adopt an innovative (i.e. risky) practice (Levitt & March, 1988). In the knowledge context, this argument has been applied to diffusion. An organization performing below aspiration is more likely to use knowledge that diffuses to it (e.g. by adopting the innovation based on that knowledge) (Kraatz, 1998).

In line with the previous discussion of the first boundary condition, we do not expect this argument to generalize to the context of money. The considerations underlying the adoption of money (i.e. spending) differ from those underlying the adoption of innovations. The most relevant difference here is that money can be spent both carefully (buying, for example, a number of cheaper players with solid records), as well as adventurously (e.g. buying one expensive talent). Therefore, changes in risk aversion should not have an effect on the amount of money spent.

Below aspiration performance is also theorized to cause stress, anxiety, and dissension, especially if the low performance threatens the survival of the organization (Audia & Greve, 2006). Such stress disrupts organizational decision making. Such disruption reduces the ability to act on incoming diffusion materials (e.g. to spend incoming money) (Staw, Sandelands, & Dutton, 1981). For example, if managers are fired in response to below aspiration performance, decisions that used to involve those managers may be delayed until their replacement is fully operational. Below aspiration performance increases susceptibility in the innovation context only because the reduced risk aversion dominates the association between low performance and high stress, anxiety, and dissension (Audia & Greve, 2006). If risk aversion is excluded in the money context, we expect the negative effect of stress to dominate.

H2: For clubs with below aspiration performance, the amount of transfer money a club receives in a season is less positively associated with that club's transfer spending in that season.

We operationalize below aspiration performance as demotion (i.e. relegation to a lower league). Another diffusion explanation may underlie an effect for demotion. This is the social proximity explanation, which considers how diffusion tends toward some pathways (i.e. dyadic ties) over others. Demoted clubs may find themselves forced into the lower league's transfer network as well. This may cause a kind of liability of newness (Stinchcombe, 1965) for demoted clubs in their new environment. Liability of newness involves being unconnected, or having low social proximity, to other actors in the network. As a result, a demoted club may not be able to find buyers for its players, or clubs that sell players, as easily as in its old network. This could break the correlation between the money a club receives and its expenditure.

We can test for the strength of the social proximity explanation by assuming that promoted clubs face a similar liability of newness when they move to the higher league's transfer network. Promoted clubs have surely performed at or above aspiration, so if the effect of H1 is also weaker for promoted clubs, the social proximity explanation is more plausible than the aspiration level explanation.

H3: For recently promoted clubs, the amount of transfer money a club receives in a season is less positively associated with that club's transfer spending in that season.

The previous hypotheses discussed how organizations react to incoming flows. The next step is to predict where in a network flows occur. Because we can observe the flows separately

from the adoption, we can directly test predictions that affect the flow. This allows us to get better traction on diffusion hypotheses than prior research in two ways. First, some characteristic may make an actor more likely to receive a flow but also less likely to act on it. For example, young organizations tend to be unconnected to established knowledge networks, so they are less likely to receive diffusing knowledge (Stinchcombe, 1965). At the same time, young organizations have great flexibility to act on new knowledge (Amburgey, Kelly, & Barnett, 1993; Greve, 2005). Researchers that infer flows from adoption patterns would be unable to detect the cancelled-out effect of an organization's age. Second, heterogeneous diffusion research identifies social proximity as a source of opportunities for flows to occur and hence diffusion of a practice (Greve, 2005). However, social proximity also non-causally correlates with the likelihood of adopting a practice because two similar actors in a shared environment are more likely to independently adopt similar practices than two dissimilar actors in different environments. This presents a troubling alternative explanation for diffusion studies that rely on a correlation between one actor's adoption and the likelihood of new adoption by another. We avoid this by directly observing a flow instead of past adoption.

There are two main arguments for why social proximity between two organizations increases the likelihood of flows of knowledge about innovations between them (Greve, 2005). The first is based on the premise that the knowledge of a similar organizations is more relevant than that of dissimilar ones (Haveman, 1993). Another version of this argument is based on the premise that organizations are better able to understand and therefore apply knowledge received from similar others (Lane & Lubatkin, 1998). For example, clothing manufacturers ignore information or knowledge they could receive about innovations in biotechnology, and they would in any case lack the ability to apply it by adopting the innovation. In line with our discussion of

the boundary conditions regarding behavior and regarding dilution, neither of these premises hold in the context of money. The value of a particular dollar bill is the same for everyone. The second argument is based on the premise that socially proximate organizations are more likely to communicate (Haunschild, 1993; Kraatz, 1998). In line with our discussion of the interaction between the first and second boundary conditions, this premise becomes more important in the context of money, because the transaction mode of diffusion mode requires active participation by both seller and buyer. Communication facilitates active participation toward shared goals such as a transaction. In contrast, consider a case where stealing is the diffusion mode for money (or, in the context of innovation, stealing trade secrets). Communication and social proximity are unlikely to increase the diffusion of money in this case.

In sum, there is at least one argument for the effect of social proximity that is still strong in the context of money.

H4: The greater the social proximity between two clubs, the more likely they are to transact.

DATA AND METHOD

The flows of money between actors are usually hard to trace, but it is uniquely well-documented and publicly available in the domain of soccer transfers. Our source is www.transfermarkt.de. From this website we obtained information on all player purchases made by all clubs that played, in 2013, in the highest league of the top six European countries in terms of club soccer performance (England, Germany, Spain, Italy, France, and the Netherlands). The data are very comprehensive for all clubs back to 2001. Older data are only complete for the most famous clubs. The transfer price of each purchase is a flow of money that diffuses from the buyer

to the seller. The transfer price is the compensation for the initial owner of the player to cancel its contract such that the player can sign a new contract.

The context of money avoids some of the factors that confound diffusion research in the innovation context. These factors mirror the boundary conditions outlined above. One instance is that although knowledge diffuses to an actor, a researcher may be unable to observe it if the likelihood of diffusion occurring correlates with the adoption of behaviors based on non-focal knowledge. For example, consider a researcher collecting data on the effect of board interlocks on the diffusion of a particular management practice between interlocked organizations. Knowledge has multiple diffusion modes that can substitute for each other and thus create noise or bias if diffusion mode choice is correlated with the independent variables. In the example, the effect may be insignificant if the diffusing practice is more likely to diffuse via observation or search rather than the knowledge sharing or spillovers that are associated with interlocks. A related issue arises when a variable appears insignificant because it has a positive effect on one diffusion mode and a negative effect on another and these effects cancel out. In the context of money, substitutes for the transaction mode of diffusion are not prevalent enough to cause such issues.

The soccer transfer setting is especially suitable for studying the diffusion of money if we compare it with consumer purchasing behavior, or investments by firms. Neither of those alternatives have publicly available comprehensive data. Furthermore, the spending behavior of those who receive the money from purchases or investments is even more difficult to track. In contrast, the soccer clubs' transfer spending goes only to other soccer clubs, allowing us track the receivers' spending, as well as creating a somewhat closed network. (Though sometimes clubs spend money outside of the observed set of firms, for example, on transfers from clubs in secondary leagues or from clubs outside Europe.) Secondly, soccer clubs have few other variable

costs that are substitutes for transfer expenditure (Kuper & Szymanski, 2012). This is useful because a researcher normally needs to control for spurious factors that could cause both the independent variable (e.g. below aspiration performance) and the adoption. In the diffusion of an idea there are usually few other uses for the idea except for adopting the innovative behavior the idea suggests. In contrast, in case of diffusion of money, all other investment opportunities are potential other uses (i.e. substitutes). Because substitutes (negatively) correlate with the adoption, all factors that affect both below aspiration performance *and any* substitute are potential causes of spuriousness. This may cancel out or otherwise bias the effect of interest. Soccer clubs' relative lack of substitutes for transfer expenditures thus helps the internal validity of our study. In addition, the extent of alternative inward and outward money streams (e.g. merchandising income to finance transfers or using transfer income to pay the stadium lease) is relatively measurable because this correlates closely with rank in the league tables. Third, because soccer players are unacceptable as collateral, clubs cannot use bank loans to finance transfers. This eliminates the possibility that the search diffusion mode creates noise, although some clubs do manage to finance transfers with debt via complex accounting and ownership practices. Fourth, European soccer does not have a draft or first pick system that exogenously restricts transfers, so our effects are not distorted by regulation. In fact, there is some regulation against buying non-European players, so the regulations increase the closure of the network. Europe-wide regulations promoting the financial stability of clubs (UEFA's so-called financial fair play) came into effect only after our observation window.

Measures

The independent and dependent variable in diffusion studies are usually prior adoption by other organizations and the likelihood of adoption by a focal organization, respectively. Some

studies have a sharper independent variable, such as ‘prior adoption by other organizations of which board members also sit on the focal organization’s board’ (Haunschild, 1993). That allows a researcher to more accurately capture the diffusion mode (e.g. knowledge sharing via board interlocks). Our design allows an ideally sharp independent variable: the yearly amount of money spent by other clubs in the transactions with the focal club, which is equivalent to the money that the focal club receives from other clubs in our data. Our dependent variable is the yearly amount the focal club spent. These club-year-level data are used to test hypotheses 1-3. The expenditure and received variables are in thousands of euros.

H2 introduces below aspiration performance. Commonly, studies calculate below aspiration performance based on deviation from a moving average of that organization’s past performance (Audia & Greve, 2006). Other studies use a unique characteristic of their empirical setting (e.g. Kraatz, 1998). Our setting suggest such a clear, easy to track, but idiosyncratic indicator: demotion (i.e. relegation to a lower league). Although different soccer clubs have different aspiration levels, demotion should qualify as below aspiration performance for any club (Kuper & Szymanski, 2012). One drawback of this operationalization is that some non-demoted clubs may have performed below aspiration. But this introduces noise which only makes our estimates more conservative. To double-check, we also used a deviation from moving average measure, and obtained similar results.

Transfer-level data are used to test H4. The independent variable, social proximity, is operationalized as the difference in rank between the two transacting clubs. For soccer clubs, there is powerful social construction of groups according to rank. The groupings are associated with similarity in the clubs’ social status, comparison and aspiration groups, club size, and player quality (Kuper & Szymanski, 2012). This makes clubs socially proximate to the other clubs in

their group (Greve, 2005). H4 is not a hypothesis about correlations, but about a difference in the mean the social proximity of transfers we observe and the mean social proximity of the transfers we would observe if clubs chose their transaction partners randomly. There are a few considerations that affect the random selection. First, in a league of 20 clubs, the number ten cannot have a rank difference higher than 10, and has a rank difference of one up to nine twice. The number three cannot have a rank difference higher than 17, and has a rank difference of one and two twice. As a result, smaller rank differences are inherently more likely. Because four leagues have 20 clubs, and two have 18, the potential average rank differences are not the same for all of the clubs. In fact, because players are also transferred internationally, simply averaging the rank difference over leagues would not be valid. In light of such considerations we numerically calculated all possible rank differences each club can have with all other clubs in the data, and averaged that over all clubs.

Our main control variable is a club's final rank in the league table in a particular season. Higher ranked clubs should have higher quality players to sell, and demand higher quality players when they buy. In addition, such clubs should have more money available for transfers from non-transfer activities, and need less transfer income to cover non-transfer expenses. We also use club-fixed effects to capture even more of this heterogeneity. Second, although we selected only clubs competing in the highest leagues in 2013, some of those clubs transacted with clubs from lower leagues, and they may themselves have been in a lower league in previous years. The variable '*League*' captures this.

RESULTS

Table 1 shows the summary statistics and correlations.

TABLE 1
Descriptive Statistics and Correlations

a) Transfer-level data

Variable	mean	s.d.	1	2	3	4	5	6	7
1. Price	1334757	3721584							
2. Buyer Rank	8.37	5.43	-.24						
3. Seller Rank	7.74	5.30	-.08	.02					
4. Rank Difference	6.28	4.42	-.12	.36	.13				
5. Year	29.43	13	.11	.01	-.01	-.01			
6. Year ²	1034.9	564.12	.11	.00	.00	-.01	.97		
7. Buyer League	1.17	.40	-.17	.08	.09	.03	.03	.00	
8. Seller League	1.11	.33	-.07	.03	-.04	.00	-.01	-.03	.14

b) Club-year-level data

Variable	Mean	s.d.	1	2	3	4	5	6	7	8	9
1. Expenditure	3923.56	12773.73									
2. Received	2442.85	8265.65	.58								
3. Received _{t-1}	2495.11	8345.82	.43	.38							
4. Rank	8.26	5.31	-.19	-.16	-.15						
5. Rank ²	96.45	101.71	-.16	-.14	-.13	.96					
6. League	1.13	.38	-.15	-.09	-.13	-.07	-.04				
7. Demoted	.03	.17	-.06	.03	-.05	-.11	-.10	.48			
8. Received*Demoted	195.54	2220.11	-.03	.18	.02	-.06	-.05	.24	.51		
9. Promoted	.04	.21	-.01	-.08	-.05	.19	.19	-.05	-.04	-.02	
10. Received*Promoted	56.24	916.25	.03	.06	.02	.03	.03	-.01	-.01	-.06	.29

Table 2 gives the results predicting a club's total expenditure on transfers in a year. The positive coefficient of the variable '*Received*' indicates support for H1. For demoted clubs, the coefficient of the interaction with '*Received*' is negative, indicating support for H2. '*Received*' also has a negative interaction with being promoted, indicating support for H3. This suggests that the social proximity explanation is strong. The negative interaction effect of '*Received*' was stronger for demoted clubs than for promoted clubs (this difference is statistically significant). This

is consistent with the idea that demoted clubs may experience reduced susceptibility on top of their reduced social proximity.

To test H4, we compare the average rank difference in transactions we observe, to the average rank difference of randomly selected pairs of organizations in the data. We calculate that the average rank difference of randomly selected pairs in our data is equal to 6.53. In Table 1, the mean value of the variable rank difference equals 6.28, well below random. (Because we have complete population data, this difference is also statistically significant). This supports H4, which predicted that transactions with high social proximity (i.e. small rank difference) should be more likely. The observed average rank difference is 6.73 for the subsample of loans (which are excluded in all other analyses) and 6.05 in the subsample of transfers. Additional analyses

TABLE 2
Results of Ordinary Least Squares Analysis of Expenditure ^a

Variables			
Received	.45	**	(.07)
Received _{<i>t-1</i>}	.05		(.05)
Rank	-888.18	**	(383.27)
Rank ²	29.40	*	(16.24)
League	-6995.55	**	(1087.39)
Demoted	2827.36	*	(1405.11)
Received*Demoted	-.79	**	(.13)
Promoted	3458.85		(2102.42)
Received*Promoted	-.55	**	(.26)
Club fixed effect	Included		
Season fixed effect	Included		
 R ²	 .65		

^a *n*=1253. Robust standard errors are in parentheses. Tests for coefficient significance are two-tailed and are tested on complete population data.

* *p* < .05

** *p* < .01

further show that clubs tend toward socially proximate transaction partners more strongly when transactions involve more money. This result also supports the premise underlying H3 that low social proximity impedes transactions. The alternative that more socially proximate clubs may be more intense rivals and therefore less willing to transact appears not to be important enough to cancel out the correlation predicted by the diffusion model. However, rivalry may explain why clubs tend to choose transaction partners with greater rank difference if that partner is from their own country.

We also tested the hypothesis using another indicator of social proximity: whether the clubs are from the same country. Clubs transfer within their country as opposed to abroad more frequently than if they were to choose transaction partners at random (77% at home versus an expected 15% at random). Thus, the hypothesis is also supported for this indicator. There is some between-country variation in this percentage, with Spanish clubs transacting with Spanish clubs only 66% of the time, while French clubs buy 86% of their players from other French clubs. Another dimension on which this tendency varies is the club's rank, with more highly ranked clubs being more likely to transact with foreign clubs. That makes sense as top clubs compete for a smaller pool of more talented players and compensate by increasing their geographic scope.

DISCUSSION

We have theorized and tested how diffusion works when we cross boundary conditions related to 1) adopter behavior, 2) diffusion mode, and 3) transfer. Combined, our hypotheses depict how socially proximate organizations are more likely to transact with each other (H4). The receiving organization is now more likely to spend money itself (H1), unless below aspiration performance and liability of newness prevent it from being the sender organization (H2 & H3) in

a new transaction (H3 back to H4). These insights inform careful generalization of diffusion research to and from the context of money.

Our approach allows us to get unique empirical traction on theoretical expectations about diffusion. For instance, we are able to observe flows of money, allowing us to exclude the alternative explanation that similar organizations adopt similar practices (i.e. spending transfer money) because of similar preferences. We find that flows (of transfer money) which cause a practice (such as spending transfer money) flow more frequently between similar organizations. The boundary conditions that separate the context of money from that of innovation insulate our empirical setting from factors that have confounded diffusion research in the innovation setting. But then, how can we generalize our results back beyond these boundary conditions? We found that below aspiration performance does not spur adoption in the context of money. If we used the premises underlying the negative effect of below aspiration performance, except swap out the premises that are context-specific (i.e. swap ‘money can be spent both carefully and in a risky way’, and ‘adoption of an innovation is risky’), then these premises would suggest a positive effect of below aspiration performance. Thus, our result also supports the reasoning for why below aspiration should spur investment in the context of innovation.

Robustness checks

Audia and Greve (2006) have identified another condition under which below aspiration performance does not spur investment: when the organization’s survival is threatened. Demotion (our measure of below aspiration performance) does put pressure on a club’s survival. In additional analyses we attempted to exclude this explanation by comparing the susceptibility of clubs recently demoted with the susceptibility of clubs which stayed in the lower leagues. We find that lower league clubs have susceptibility similar to that of higher league clubs rather than similar to that of

demoted clubs. Thus, the effect of demotion is not caused by the threat to survival posed by being in a lower league. Instead, the characteristics of the diffusion material (specifically, that money can be spend both carefully and adventurously) constitute a new contingency with a similar effect as the contingency found by Audia and Greve (2006).

We also check for robustness with a measure of below aspiration performance which takes the value 1 when a club ranks four or more places below the average of its rank over the previous four years. That measure also negatively moderates the relationship between receiving and spending transfer money. The alternative measure has high overlap with the demotion variable. In addition, its effect becomes very small when we exclude those poor performers that are also demoted. That overlap makes the alternative measure less useful for ruling out alternative explanations, but it does give confidence in the convergent validity of our measures of below aspiration performance.

As noted in the method section, correlations may be spurious if the independent variable correlates with any substitute for transfer expenditure. Demotion decreases visitor and television income, which should decrease the likelihood that transfer income is used for transfer expenditure. Demotion also decreases wage costs. That should tend to free up transfer income for transfer expenditure. Considering promotion as well allowed us to evaluate such substitutes more accurately, because promotion normally has the opposite effect on them. For example, if reduced visitor and television income, rather than stress, which causes the weakened relationship between transfer income and transfer expenditure, we should have found a positive moderating effect of promotion.

These additional analyses provide more confidence that the boundary conditions of the money context explain why the results diverge from those of prior research (with respect to the

‘threat to survival’ explanation, for example). However, we infer support for the risk rigidity mechanism from the difference between the effects of demotion and promotion. The main concern about this inference are between-country differences in the effects of promotion we discovered in additional analyses. The weak negative effect of promotion (as predicted) is the result of aggregating strong (stronger than demotion) negative effects in four countries and slight positive effects in two other countries (Spain and Italy). One of the promoted teams in Italy was Juventus, which was demoted as a penalty for a betting scandal. If we exclude Juventus from the analysis, then the effect for promoted clubs is more negative than for demoted clubs. Because of this sensitivity, we cannot confidently exclude social proximity as an alternative explanation to the below aspiration performance and risk rigidity explanation for H2. The negative effect of demotion is more stable between countries than the effect of promotion (also if we use the alternative measure). Perhaps the social proximity mechanism is highly variable between countries, such that the relative stability of demotion’s effect is evidence for a different (additional) mechanism to underlie its effect, such as risk rigidity.

One argument which would go against our core contribution is the alternative conceptualization of vacancy chains. That conceptualization reinterprets the transfer phenomenon from the player perspective (rather than a money perspective). The argument is that our results are spurious because a player’s departure causes both 1) receiving money, and 2) a vacancy that needs to be filled up by new players. Then, new players coming to fill up the vacancy cause expenditures. As a result, receiving money and expenditures are correlated, but not causally linked as the diffusion model suggests. To assess the significance of this alternative we added ‘the number of players that departed from a club in a year’ as a covariate to the model in Table 2. The coefficients of all the other variables are robust to this addition. The covariate itself correlates significantly

with expenditures (unless ‘the number of arriving players’ is also included in the model). Additional analyses show that number of departures also powerfully predicts the number of arriving players ($\beta=.24$), and the amount of money received. The number of arriving players is significantly associated with expenditures but not with the amount of money received. Thus, the ‘vacancy chain’-model seems useful, because 1) departures do cause receiving money as well as new players arriving, and 2) players arriving affects expenditure. However, the diffusion model is robust, because the diffusion variables are still as powerful after the vacancy-chain variables are added. This makes sense because vacancies may also be filled by means of loans without transfer price. This reduces the correlation between vacancies and transfer expenditure.

Another alternative explanation is that engaging in transfers causes clubs to gain information about how to accurately value players. Accurate valuation reduces uncertainty of transfers. Reduced uncertainty increases the likelihood that the club engages in transfers. Thus, selling a player can lead to buying a player via more accurate valuation instead of money diffusion. This alternative explanation could also explain a negative moderation effect of demotion and promotion. If demoted clubs are more rushed to sell their players, perhaps because players are eager to leave given the reduced salary in lower leagues, then sales could reflect the market’s valuation of the players less accurately. Thus, demoted clubs gain less information from sales, and sales are therefore a less informative basis for new spending. In line with this explanation, though not inconsistent with a diffusion explanation either, we observe that clubs that spent more in the previous year spend more in a focal year. In addition, it predicts the vacancy chain findings: if the number of departing players increases, more transfers have been made, so the reduction in uncertainty should be greater, so more new purchases should be made. In fact, this alternative explanation must work via the number of sales and purchases. If accurate valuation were to have

a direct effect on expenditures, it is unclear why that effect should be positive. In other words, the accurate valuation may in fact be lower than the valuation by clubs without the information. As a result, the same control variables that capture the vacancy chain explanation also capture this alternative explanation. The fact that the hypothesized effects are still observed while these controls are included means that the diffusion of money explanation is robust to this alternative explanation as well.

Implications

Given that both knowledge and money are ingredients for innovation, there are interesting trade-offs if a variable that increases susceptibility to knowledge flows also decreases susceptibility to money flows. For example, it appears that organizations that perform below aspiration would be more susceptible to knowledge flows (Kraatz, 1998) but less susceptible to money flows. One caveat is that this tension only appears because our dependent variable combines both risky as well as careful money flows. In other words, if the dependent variable had been ‘expenditure on risky/innovative investments only’, our results might not have indicated this tension.

We are also able to understand how new wealth spreads in a system. One concept that has helped economists understand wealth in a system is the ‘velocity of money’. The velocity of money is higher (i.e. diffusion is faster) if more transactions are made, keeping constant the money supply (i.e. the number of bills in circulation). Such studies have generally focused only on country-level variables to explain the velocity in a country at a particular time (Friedman, 1959). Applying insights from the heterogeneous diffusion model allows us to identify and test new (organization-level) correlates of velocity. For instance, the velocity may decrease in systems where organizations have low social proximity (as we find in H4).

A diffusion perspective also improves our understanding relative to a sociological perspective of embeddedness. This perspective also suggests effects of social proximity variables (Sorenson & Stuart, 2001; Uzzi, 1999). However, it does not consider characteristics of organizations that influence their reaction to incoming flows of money (as we do in H2 and H3). As such, perhaps the embeddedness perspective is best viewed not as an alternative to a diffusion model but as input for it.

Another interesting finding from our subsequent analyses is that the spending of lower league clubs depends more on last season's transfer income (i.e. '*Received_{t-1}*'), while the spending of first league clubs depends more on the current season's transfer income. One possible explanation for this is that top clubs have better routines (Greve, 2005) and can therefore handle incoming and outgoing diffusion more quickly. Another explanation is that first league clubs have greater trust in their transaction partners, allowing them to spend money (on buying players) they are not sure they are going to receive in the near future (e.g. sales of players are still being negotiated). That would be consistent with the idea that liability of newness explains H2 and H3 if liability of newness involves being unfamiliar with the new environment's routines and/or a lack of trusted connections. This makes for an interesting idea that diffusion research could contribute to the velocity of money research; velocity may be enhanced by developing diffusion routines in organizations. For example, firms with high price-setting capability (Dutta, Zbaracki, & Bergen, 2003) may diffuse more money more quickly (in other words, they have higher infectiousness).

Our results also indicate variables that underlie wealth differences. Prior research in this area tends to explain inequality as a result of income inequality as a result of education and family background (Autor, Katz, & Kearney, 2008; Fernandez, 2001). The heterogeneous diffusion model setup suggests a whole new range of additional variables. For instance, H2 and H3 suggest that

actors with low susceptibility and low social proximity (to sources of money) have less purchasing power, because they seem less able to spend the money they receive. H4 suggests that actors with low social proximity are less likely to have money diffuse to them. Moreover, it follows from H2 and H3 that actors who rely on others with low susceptibility and low social proximity as a source of money, are less likely to receive money. This suggested research direction follows in the spirit of a study by Fernandez-Mateo (2007) where workers with equal skills were found to receive different wages depending on their staffing intermediary's relationship with employers. Thus, actors in brokerage positions (e.g. staffing intermediaries) channel the diffusion of money, affecting the distribution of workers' wealth over and above these workers' own characteristics.

We apply diffusion ideas to the context of money. In return, there are also some possibilities for research focusing the phenomenon of money (i.e. finance and macroeconomics) to inform specifications of the diffusion model. For example, there is some literature in finance that aims to explain why firms save money. Riddick and Whited (2009) find that income uncertainty predicts saving by firms. Thus, income uncertainty could be used in explaining an organization's susceptibility to money flows. Perhaps, lower league clubs could not base their spending on current season's transfer income because that income is too uncertain. Another example is shareholder-value maximization orientation, which influences the likelihood that firms use free cash flow to pay dividends to shareholders instead of using it for unrelated diversification (Kavadis & Castañer, 2014). One application of this idea is that actors with shareholder orientation are less susceptible to adopt spending behavior, because they are less likely to use money received from transfers, preferring instead to extract money from the transfer system and return it to shareholders. Similarly, correlates of the velocity of money found by macroeconomists (Friedman, 1959) could be used in diffusion models.

REFERENCES

- Abrahamson, E., & Fairchild, G. 1999. Management Fashion: Lifecycles, Triggers, and Collective Learning Processes. *Administrative Science Quarterly*, 44(4): 708–740.
- Amburgey, T. L., Kelly, D., & Barnett, W. P. 1993. Resetting the Clock: The Dynamics of Organizational Change and Failure. *Administrative Science Quarterly*, 38(1): 51–73.
- Audia, P. G., & Greve, H. R. 2006. Less Likely to Fail: Low Performance, Firm Size, and Factory Expansion in the Shipbuilding Industry. *Management Science*, 52(1): 83–94.
- Autor, D. H., Katz, L. F., & Kearney, M. S. 2008. Trends in U.S. Wage Inequality: Revising the Revisionists. *Review of Economics and Statistics*, 90(2): 300–323.
- Baum, J. A. C., Li, S. X., & Usher, J. M. 2000. Making the Next Move: How Experiential and Vicarious Learning Shape the Locations of Chains' Acquisitions. *Administrative Science Quarterly*, 45(4): 766–801.
- Berger, J., & Le Mens, G. 2009. How adoption speed affects the abandonment of cultural tastes. *Proceedings of the National Academy of Sciences*, 106(20): 8146–8150.
- Burns, L. R., & Wholey, D. R. 1993. Adoption and Abandonment of Matrix Management Programs: Effects of Organizational Characteristics and Interorganizational Networks. *The Academy of Management Journal*, 36(1): 106–138.
- Dutta, S., Zbaracki, M. J., & Bergen, M. 2003. Pricing process as a capability: a resource-based perspective. *Strategic Management Journal*, 24(7): 615–630.
- Fernandez-Mateo, I. 2007. Who Pays the Price of Brokerage? Transferring Constraint through Price Setting in the Staffing Sector. *American Sociological Review*, 72(2): 291–317.
- Fligstein, N. 1985. The Spread of the Multidivisional Form Among Large Firms, 1919-1979. *American Sociological Review*, 50(3): 377–391.
- Friedman, M. 1959. The Demand for Money: Some Theoretical and Empirical Results. *Journal of Political Economy*, 67(4): 327–351.
- Greve, H. R. 2005. Interorganizational Learning and Heterogeneous Social Structure. *Organization Studies*, 26(7): 1025–1047.
- Hansen, M. T. 1999. The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits. *Administrative Science Quarterly*, 44(1): 82–111.

- Haunschild, P. R. 1993. Interorganizational Imitation: The Impact of Interlocks on Corporate Acquisition Activity. *Administrative Science Quarterly*, 38(4): 564–592.
- Haunschild, P. R., & Miner, A. S. 1997. Modes of Interorganizational Imitation: The Effects of Outcome Salience and Uncertainty. *Administrative Science Quarterly*, 42(3): 472–500.
- Haveman, H. A. 1993. Follow the Leader: Mimetic Isomorphism and Entry Into New Markets. *Administrative Science Quarterly*, 38(4): 593–627.
- Kavadis, N., & Castañer, X. 2014. Ownership Effects on Unrelated Diversification: An Institutions? Perspective. *Finance and Strategy*, vol. 31: 253–288. Emerald Group Publishing Limited.
- Kennedy, M. T., & Fiss, P. C. 2009. Institutionalization, Framing, and Diffusion: The Logic of TQM Adoption and Implementation Decisions among U.S. Hospitals. *Academy of Management Journal*, 52(5): 897–918.
- Kraatz, M. S. 1998. Learning by Association? Interorganizational Networks and Adaptation to Environmental Change. *Academy of Management Journal*, 41(6): 621–643.
- Kuper, S., & Szymanski, S. 2012. *Soccernomics: Why England Loses, Why Spain, Germany, and Brazil Win, and Why the US, Japan, Australia, Turkey-and Even Iraq-Are Destined to Become the Kings of the World's Most Popular Sport* (Second Edition edition). Nation Books.
- Lane, P. J., & Lubatkin, M. 1998. Relative absorptive capacity and interorganizational learning. *Strategic Management Journal*, 19(5): 461–477.
- Lave, C. A., & March, J. G. 1975. *An introduction to models in the social sciences*. New York: Harper & Row.
- Levitt, B., & March, J. G. 1988. Organizational Learning. *Annual Review of Sociology*, 14: 319–340.
- M. Fernandez, R. 2001. Skill-Biased Technological Change and Wage Inequality: Evidence from a Plant Retooling. *American Journal of Sociology*, 107(2): 273–320.
- Ody-Brasier, A., & Vermeulen, F. 2014. The Price You Pay Price-setting as a Response to Norm Violations in the Market for Champagne Grapes. *Administrative Science Quarterly*, 59(1): 109–144.

- Rao, H., Davis, G. F., & Ward, A. 2000. Embeddedness, Social Identity and Mobility: Why Firms Leave the NASDAQ and Join the New York Stock Exchange. *Administrative Science Quarterly*, 45(2): 268–292.
- Rao, H., Greve, H. R., & Davis, G. F. 2001. Fool's Gold: Social Proof in the Initiation and Abandonment of Coverage by Wall Street Analysts. *Administrative Science Quarterly*, 46(3): 502–526.
- Riddick, L. A., & Whited, T. M. 2009. The Corporate Propensity to Save. *The Journal of Finance*, 64(4): 1729–1766.
- Sorenson, O., Rivkin, J. W., & Fleming, L. 2006. Complexity, networks and knowledge flow. *Research Policy*, 35(7): 994–1017.
- Sorenson, O., & Stuart, T. E. 2001. Syndication Networks and the Spatial Distribution of Venture Capital Investments. *American Journal of Sociology*, 106(6): 1546–1588.
- Staw, B. M., Sandelands, L. E., & Dutton, J. E. 1981. Threat Rigidity Effects in Organizational Behavior: A Multilevel Analysis. *Administrative Science Quarterly*, 26(4): 501.
- Stinchcombe, A. L. 1965. Organizations and social structure. *Handbook of organizations*, vol. 44: 142–193.
- Strang, D., & Soule, S. A. 1998. Diffusion in Organizations and Social Movements: From Hybrid Corn to Poison Pills. *Annual Review of Sociology*, 24: 265–290.
- Strang, D., & Tuma, N. B. 1993. Spatial and Temporal Heterogeneity in Diffusion. *American Journal of Sociology*, 99(3): 614–639.
- Viboud, C., Bjørnstad, O. N., Smith, D. L., Simonsen, L., Miller, M. A., et al. 2006. Synchrony, Waves, and Spatial Hierarchies in the Spread of Influenza. *Science*, 312(5772): 447–451.

Closing Chapter

The introduction described an approach to research that tacitly guided my three papers. I hope the approach is interesting enough (and I hope the resulting papers help demonstrate that) to inspire more researchers to think critically about their own approach and perhaps to see what this approach might bring for them. The three papers show three new interpretations of concepts and the new premises they suggest. I hope the interpretations are interesting enough to inspire future research to search for more new premises that the concepts may suggest. The papers also show three phenomena about which the concepts suggested premises that helped derive accurate predictions. I hope the papers convinced readers about the usefulness of these concepts in suggesting accurate hypotheses, so that future research applies those concepts to make more predictions and to come up with plausible explanations for puzzling aspects of those phenomena.

Short summary of the results

The first paper showed that when a foreign market entry occurs, the likelihood of a new foreign market entry by another firm is high. We argue this is because the initial market entry indicates that an event (e.g. regulatory change or a demand shock) occurred that might also trigger another firm. Then, this likelihood decreases over time, until after about two years (in our banking industry data). We argue this is because the information from the event slowly becomes obsolete. Then, after two years, the likelihood gradually increases again. We argue this is because imitators are gradually becoming more likely to succeed in reacting to the initial entry. Then after a few months of gradual increase, the likelihood reverts back to a decreasing pattern. We argue this is because the information that led to imitation is becoming obsolete.

The second paper showed that the effect of the number of foreign firms in an industry on the number of foreign entries into that industry depends on the structure of the taxonomy of social categories in that industry. For example, in some industries there are many subcategories in the taxonomy. We find that the more subcategories in an industry, the stronger the effect of the number of foreign firms in a subcategory on the number of foreign entries (relative to the effect of the number of foreign firms in that industry in general). Another characteristic of taxonomy structure is the degree to which a subcategory is typical of the industry in general. A typical subcategory has many attributes in common with the other subcategories. In contrast, an atypical subcategory is ‘the odd one out’. For example, penguins are atypical birds. We find that the more typical a subcategory is, the weaker the effect of the number of foreign firms in that subcategory on the number of foreign entries (relative to the effect of the number of foreign firm in that industry in general).

The third paper showed how money diffuses between soccer clubs. Soccer clubs spend more money on transfers if they receive more money from transfers, controlling for the number of players departing and unobserved time-invariant club characteristics and soccer performance. How much a club spends of the money it receives, depends on its soccer performance and its social proximity to potential transaction partners. Furthermore, the paper shows that soccer clubs prefer to transact with socially proximate clubs, as measured by difference in rank in the league table, as well as geographic distance. These factors should help evaluate the idea of ‘trickle-down’ economics in arguments about wealth differences. If all economic actors behave like soccer clubs behave in the context of transfers, then the results suggest that the creation of wealth is more likely to spread to actors who are socially proximate to the wealthy creator. In addition, wealth is less likely to spread if actors are isolated from potential transaction partners. These implications

suggest that, if wealthy people are most likely to create new wealth, and wealthy people are more likely to be socially proximate to wealthy people than to poor people, then the new wealth is likely to spread to other wealthy people instead of to poor people, and wealth differences will grow. In addition, if actors that receive money are isolated from spending opportunities, wealth will spread less to anyone. Practically, this means wealth differences can be decreased by initiatives that increase social proximity between wealth creators and poor people and that provide an opportunity for spending by wealth creators.

Insights emergent from the interactions between the papers

The first paper studies three cities within the same continent. Banks tend to have one office that is the continental-level headquarters. The decision to build such continental headquarters may have different underlying motivations than those for other offices and may involve different reaction speeds. Possibly, this sets the stage for a particular kind of interdependence between the cities. For instance, banks may imitate continental headquarter strategies of other banks, but once they have an office in East-Asia, they enter East-Asian countries independently. Or alternatively, once they have one office in East-Asia, they are more informed about competitors' actions in that region, and are therefore more likely to imitate. If the importance of imitation or common cause in bank's entry decisions differs between continental headquarters and other offices, then actions in one city can influence the results in the other city. The results of the first paper are at the average of such interactions. A focal bank's behavior might differ not only between headquarters or normal offices, but also between whether the prior entry was a continental headquarters or not. The theory in the second paper can inform some speculations about those differences. That theory says that actions can be interpreted at the generic or the specific level of detail. Continental-level strategies are generic relative to country-level strategies. The second paper would argue that a bank, once it

has a continentally leading office (e.g. in Shanghai), tends to interpret actions with more detail. Thus, if a competitor sets up a continentally leading office in Tokyo, the focal bank interprets this as a country-level strategy. In contrast, a bank without existing presence in the continent would interpret this as a continental-level strategy. If a bank interprets its competitor's entry into Tokyo as a country-level strategy, then imitations are limited to Japanese cities. In contrast, if it is interpreted as a continental-level strategy, then that bank may imitate by entering Shanghai. In that case, because the competitor's entry is not taken into account in our Shanghai data, we fail to observe a case of imitation. If such situations occur, that does not mean we systematically misclassify the entry, unless the competitor's entry into Tokyo is also non-causally correlated with the timing of entry into Shanghai immediately prior to the imitative entry into Shanghai. A causal correlation would lead to a correct classification in two ways. First, if the competitor enters Tokyo and follows up (within months) by also entering Shanghai, then the entry into Shanghai that imitates the competitor's entry into Tokyo is (accidentally correctly) classified as imitating that competitor's entry into Shanghai. Second, if the competitor's entry into Tokyo causes the prior entry in Shanghai, then the imitative entry into Shanghai is (correctly) classified as sharing a common cause with its prior entrant (unless the speed at which the competitor's entry causes prior entries differs systematically from causing the focal entry).

In the second paper we measure imitation as the correlation between density and entry. In the first paper, we recommend to measure imitation as the correlation between the number of entries two years ago and entry. The best way to resolve this inconsistency is to rephrase the first statement: we explain the correlation between density and entry in terms of imitation. An explanation of the correlation between density and entry in terms of legitimacy is also possible,

but was abbreviated to streamline exposition. An explanation in terms of profitability was explicitly considered but discarded.

The relationship described in the first paper, between the effect of information and time since an event, may be supported by insights from the third paper. The effect of information is the outcome of the diffusion of information. The factors that influence diffusion are described in the third paper. The extent to which organizational actions influence the diffusion of information, combined with the extent to which organizations respond to that information, equals the extent to which imitation occurs. Therefore, the difference between Tokyo and Shanghai in their inter-arrival time pattern may be caused by factors identified in the third paper. For example, theory in the third paper suggests that if two entrants are more socially proximate, diffusion between them is more likely. Social proximity involves frequency of encounters. Therefore, diffusion between them may also be faster. Unfortunately, results from additional analyses on whether banks from the same home country imitated each other more or faster were inconclusive.

The second paper features a result that rules out profitability as the driver of the effect of density on entries. Ruling out common causes such as profitability is a main goal of using time as a moderator in the first paper. Thus, the first and second paper support each other in ruling out an alternative explanation. The fact that the first paper has results that we interpret as inconsistent with a common cause makes it plausible for the second paper to interpret its result as inconsistent with common cause, and the fact that the second paper has results that we interpret as inconsistent with a common cause makes it plausible for the first paper to interpret its results in that way.