AGE-RELATED STRUCTURAL INERTIA: A DISTANCE-BASED APPROACH

ABSTRACT. This paper proposes a distance-based characterization of age-related structural inertia as an increasing constraint on the speed of change as organizations age. Our framework regards organizations as points in multi-dimensional metric spaces of architectures. Organizational change means movement in this space. The speed of change is the ratio of the distance between positions in a space and the time it took for the organization to make the move. We illustrate how our distance-based approach can be used to formulate theories of age-related organizational inertia by using this representation to develop a model for a possible mechanism: age-related cultural resistance to change based on the dynamics of exposure of organizational members to architectural features. Our proposed mechanism is distinct from prevailing explanations and leads to new predictions. We also illustrate the value of our distance-based approach in a reanalysis of Sørensen and Stuart's (2000) study of age variations in firms' patenting behavior. On the basis of patent citations, we construct a space that allows us to characterize the positions of organizations, and the speed at which they change. We find that organizational age has a negative effect on the speed of change.

1. INTRODUCTION

Age-related structural inertia plays a central role in organization theory. It has been invoked to explain a variety of organizational outcomes, such as the positive relation between organizational age and the failure hazard, the observation that older organizations tend to have more trouble adapting to changing environmental conditions, or the fact that the innovations of older organizations are more similar to their previous innovation efforts (for reviews, see Barnett and Carroll (1995); Carroll and Hannan (2000); Sørensen and Stuart (2000); Gilbert (2005)). Although the intuition behind age-related structural inertia—that older organizations have more trouble changing their architectures than younger ones—is fairly straightforward, this concept has so far escaped attempts at defining it precisely and connecting it to variables that can be more easily observed.

In this paper, we address this challenge by conceptualizing age-related structural inertia in terms of speed of change in a space of organizational architectures: we will say that an organization experiences age-related structural inertia if the maximal speed at which it can change its architecture in this space decreases with age. Our definition relies on a construction that sees organizational

architectures as positions in a metric space, a space with a distance measure. Then we illustrate the benefits of this clear definition by developing a formal model that provides a new explanation for age-dependent structural inertia. Most existing explanations that invoke an age-dependent factor such as organizational size, complexity, set of routines to explain why structural inertia increases with age. As such, in these explanations, the effect of time is completely mediated by another causal factor. By contrast, our model gives a more central role to time by focusing on the length of time organizational members have been exposed to features of the organizational architecture. We show that even when there is turnover in organizational membership, the distribution of exposures of organizational members to architectural features increases with time. This, together with the assumption that people are resistant to changes of features to which they have been exposed for a long time, is enough to imply a positive relation between age and structural inertia.

Although our explanation for age-related structural inertia is novel, we see the main contribution of the paper as providing the necessary framework to formalize theories of age-related structural inertia. We hope that other scholars will build on this framework to specify alternative theories. The formal language should help clarify how the theories differ and to specify empirical tests that should parse between various explanations.

The paper is organized as follows. Section 2 develops the distance-based characterization of organizational change and structural inertia. Then, after reviewing the existing explanations for age-related structural inertia in Section 3, we propose a describe our model in Section 4. Section 5 points to the potential difficulties in making inferences about the occurrence of structural inertia on the basis of data about what happens in offer space. It provides a set of sufficient conditions that warrants such inferences. In Section 6, we illustrate our distance-based approach with a reanalysis of Sørensen and Stuart's (2000) data on patenting by dedicated biotech firms. Finally, in Section 7, we discuss the assumptions of our model and its theoretical and empirical implications.

2. A Distance-Based Characterization and a Target Theorem

We first define the architectural space and provide a characterization of structural inertia. Our proposed definition builds on a characterization of organizations as points in a multi-dimensional metric space of architectural features.

What kinds of features characterize an organization's architecture? In general, an organizational architecture is a set of codes specifying the ontology of the organization (what are its units) and

the control relations holding among units, and possibly a lot more. Such codes can be represented in terms of features that take on discrete values, either binary or multi-valued. A particular architectural code is a string of values of the features just as in the familiar representations from information theory and genetics.In formal terms we regard an organizations as a vector of feature values. Suppose that K relevant dimensions characterize an organizational architecture. A description of the architecture of an organization (x) at a time point (t) is given by the indexed set $\mathbf{a}_{x,t} = \langle a_{1,x,t}, \ldots, a_{K,x,t} \rangle$, where $a_{k,x,t}$ denotes the value of the k-th architectural feature. The space of organizational architectures is denoted by $\mathcal{A}$.[1]

To conceptualize organizational change, one has to compare organizational architectures at different points in time. In order to do so, one needs a distance measure. In this paper, we will use the Hamming distance. The Hamming distance between two vectors of length K counts the number of positions at which the values are different. Of course this is a somewhat restrictive definition of the distance between architectural codes. Nonetheless, it allows us to develop the the implications of the core intuitions in a framework with potentially broad application. For any two organizational architectures $\mathbf{a}, \mathbf{a}' \in \mathcal{A}$, we will denote by $d(\mathbf{a}, \mathbf{a}')$ the distance between $\mathbf{a}$ and $\mathbf{a}'$. The space of architectures together with the distance d is denoted by $< \mathcal{A}, d >$ and defines a metric space.

For simplicity, we adopt a discrete conception of time. We then rely on the elementary notion of speed of change in discrete time as the distance traveled in a single period.

Definition 1 (Speed of architectural change). The actual speed of movement in architectural space of a producer x at time t is given by $z_x(t) = d(\mathbf{a}_{x,t}, \mathbf{a}_{x,t+1})$.

Potential speed in architectural space need not be the same as actual speed because an organization might change less than it could for any number of reasons. For example, managers might believe that the organizational environment will remain stable for a while and thus might decide not to reconfigure the organization even if they could. We define the maximal or *potential* speed as follows. First let $\mathbf{A}_{x,t}$ denote the set of architectures that the organization x can reach in one time unit at time t. Then the maximal distance that can be transited in one period is the distance from the origin to the most distant element in $\mathbf{A}_{x,t}$.

Definition 2 (Potential (maximal) speed of architectural change). $\psi_{x,t} = \max_{\mathbf{a}' \in \mathbf{A}_{x,t}} d(\mathbf{a}_{x,t}, \mathbf{a}')$.

---

[1]This representation resembles that used in applications of NK models to organizational issues (Levinthal 1997; Levinthal and Warglien 1999).

With the notion of potential speed in hand, we can formalize our definition of age-related structural inertia and introduce a corresponding predicate to refer to occurrences of organizational inertia:

Definition 3 (Age-related structural inertia)**.** An organization x experiences age-related structural inertia if and only if its *potential* speed of architectural change declines with age.

$$\textsc{si}(x) \leftrightarrow \forall\, t, t'[(t < t') \to \psi_{x,t} > \psi_{x,t'}];$$

This definition is useful for at least two distinct but related purposes. First, it provides some guidance for constructing empirical measures that could be used as evidence for (or against) age-related structural inertia. Second, it provides the 'right hand side' of the main prediction made by a theory of structural inertia. Indeed, such theory would lead to a theorem that specifies the conditions under which structural inertial increases with age:

Target theorem 1. *[Age-related structural inertia] Under some (not-yet-specified) assumptions,* $\mathcal{C}(x)$*, an organization experiences structural inertia:*

$$\mathcal{C}(x) \to \textsc{si}(x).$$

3. Existing explanations

Among those who contributed to the considerations concerning the relationship between organizational age and inertia there appears to be a consensus that inertia grows with age in a *population* of organizations: as an organizational cohort ages, inertia proliferates. The consensus breaks down though when attention shifts to the why question.

Hannan and Freeman (1984) point to environmental selection. Selection favors structurally inert organization. Consequently older cohorts contain more inert organizations than do younger ones. Importantly, this argument works even if individual organizations do not experience change in the strength of structural inertia over their life times. Indeed it does not address aging per se. An alternative selection-based argument builds on notions about size and complexity. New organizations tend to be small and relatively simple. As organizations age both size and complexity increase. Simple and small organizations face weaker inertial pressures than complex and large organizations. If small organizations face higher hazards of mortality, then older cohorts will have higher average levels of structural inertia than younger ones because they are larger and more complex (Hannan and Freeman 1984; Péli et al. 2000).

Schumpeter (1934) and Hannan and Freeman (1984) argued that structural inertia increases as a consequence of external organizational dynamics. As an organization ages, its external audience develops an increasingly taken-for-granted image of the organization and the organization gets increasingly embedded into a web of inter-organizational relationships. Both the taken-for-granted image and this web of ties generates external resistance to organizational change since breaking up ties, deviating from the taken for granted image imposes serious costs that most organizations are reluctant to pay. In other words they increasingly avoid the costly alterations of organizational configurations, they are becoming increasingly inert. Several studies have shown that such mechanism leads organizations to avoid investing in innovation (Gilbert and Newbery 1982; Reinganum 1983; Christensen and Bower 1996).

Others argue that the key mechanisms concern internal organizational dynamics that lead to increasing inertia of individual organizations. Organizations typically rely on procedural logics rather than consequential ones, and well-established procedures create obstacles to organizational change (Cyert and March 1963; Nelson and Winter 1982). The content of organizational memory—procedures, scripts, discourse, routines and shared organizational culture—provides the source of inertia (Gilbert 2005; van der Steen 2009; Naslund and Pemer 2012). The more extensive is the content, the stronger is the inertial pressure. However the stock of procedures does not stay constant over the lifetime of the organization; new organizations lack well-established procedures. Procedures and rules are invented as solutions to problems, challenges the organization has faced and stored in organizational memory (Levitt and March 1988; March et al. 2002). New organizations are blessed with ignorance, not much has been stored in memory. This in turn means that memory creates few obstacles to change. Older organizations have faced more problems and devised and stored more solutions. In that case, memory creates numerous obstacles to change and organizations have become more inert. A common feature of these existing explanations is that time is not a direct cause for structural inertia. Instead, age affects a factor (such as size, complexity, set of routines, embeddedness in web of external relations) which in turns affects structural inertia. Our explanation differs, in that it gives a more central role to time. Using a simple dynamic model, we demonstrate that the length of exposure to organizational features by organizational members increases with organizational age even when there is turnover. This and the individual-level assumption that people are more resistant to change when they have been exposed to a feature for a long time is enough to imply a collective outcome: age-dependent structural inertia.

4. A Possible Mechanism: Age-Related Cultural Resistance to Change

Our explanation for age-related structural inertia starts by noting that people have strong tendencies to prefer the status quo both inside and outside organizations. There are many reasons why this might be the case. We list just a few here. First, people have a natural tendency to like what is familiar to them: this is the mere-exposure effect (Zajonc 1968). Second, people learn from experience, which leads their job-related performance to increase with time. This in turn increases the perceived costs of change. Third, people come to like their environments because not doing so would create an uncomfortable state of cognitive dissonance (Festinger 1957). The motivation to reduce cognitive dissonance together with an ubiquitous self-enhancement motive leads people to see themselves as competent and even to systematically overestimate their competence (Aronson 2007). The operation of dissonance-reduction mechanisms requires that people hold the job already, and therefore this over optimistic bias does not apply to the same extent for future jobs. This implies that people will likely predict their competence and job-satisfaction to be lower after the implementation of organizational change. Fourth, organizational change creates a lot of uncertainty. But people are generally risk-averse and ambiguity averse. This tendency comes at least partly from the fact that people generally find hard it to imagine what does not exist, a tendency epitomized by Kahneman's (2011) formula: WYSIATI ("What You See Is All There Is"). Fifth, even if the change did not induce any uncertainty, it would likely imply that some aspects of the job would improve while others become less attractive. But people are subject to the endowment effect that makes them systematically prefer the status quo, even if an outsider would view the new option as equally attractive as the older one (Samuelson and Zeckhauser 1988; Tversky and Kahneman 1981). Sixth, people might anticipate experiencing regret if they do not oppose the change and their future conditions turn out to be inferior to the status quo.

In summary, there are numerous reasons why individuals would prefer the status quo rather than organizational change. In the rest of the paper, we follow Hannan et al. (2007) and refer to this individual preference for the status quo as *cultural resistance*. Such resistance varies in intensity from mild grumbling to unwillingness to cooperate in implementing change, to active efforts to thwart the change. While some of the above factors do not clearly depend on how long an individual has been exposed to the status quo (factors 4, 5, and 6 noted above), the first three factors predict that individuals will be increasingly resistant to change as they become more accustomed to the prevailing state of affairs. Therefore, we will assume that each individual in the organization

becomes increasingly likely to oppose a change in any organizational feature value the longer she has been exposed to it.

We will show that this individual-level assumption is sufficient to explain that organizations change more and more slowly with time even if the membership turns over. We now turn to the description of our model.

4.1. **Model Structure.** We consider a single organization. In every period, the organization potentially updates some of its architectural features. But members are resistant to change. And resistance to change tends to increase with the length of time with which organizational members have been exposed to features. The reason has to do with exposure, as we discussed above. As an agent gains experience with some architectural feature value, the probability that she takes it for granted increases. If people take a feature value for granted, then they will tend to oppose attempts at changing it. This increased tendency to resist change slows down organizational change. This, in turn, leads to a tendency for longer exposures, which slows down organizational change even more.

We assume, for mathematical simplicity, that the organization has infinitely many members and infinitely many features. This simplification allows us to keep the state description of the organization as simple as possible. With this assumption it is enough to look at the evolution of the distributions of the tenures of members and features. Without this simplification, the state description would have to include the dates at which each member joined the organization and at which each feature was altered. Although such a representation would make our model more realistic, it would lead to almost intractable mathematical complications that would prevent us from developing formal proofs for our results.[2] We now turn to a description of these state variables and of the mechanism that lead to state transitions.

State Variables. Organization x at the beginning of period t is described by the following state variables:

- $\mathbf{a}_{x,t} = \langle a_{1,x,t}, \ldots, a_{K,x,t} \rangle$: the architecture of the organization;
- $W_{x,t}$: the tenure distribution of organizational members;
- $U_{x,t}$: the tenure distribution of architectural features.

---

[2]This simplifying assumption is akin to making a mean-field approximation, used widely in statistical physics, which essentially ignores some of the randomness around the main trajectory of the dynamics of the system.

State Transitions. Two things happen in each period: some new members join and some existing members leave the organization (turnover in membership), and the organization alters some of its architectural features. The dynamics of tenure of memberships is driven by the rate of turnover in membership ($m_{x,t}$ in our notation), defined as the proportion of organizational members that are new at the beginning of period t. Because all members are new in the first period, $m_{x,1} = 1$. The key rate for architectural change is the rate of feature updating ($z_{x,t}$ in our notation), which is defined as the proportion of organization's feature values that take on new values in period t.

We make four assumptions regarding the dynamics of membership and feature values.

Definition 4 (Model assumptions). The following four assumptions, denoted by $\Omega(x)$, describe the dynamics of the organization:

(i.) In each period, the to-be-replaced members are randomly selected, without replacement, until the proportion of the members that have been selected is equal to $m_{x,t}$.

(ii.) $m_{x,t}$, the rate of turnover of organizational members, is (a.) positive, (b.) lower than one, and (c.) remains constant in all periods starting in period 2 ($\forall t \geq 2, m_{x,t} = m$).[3]

(iii.) In each period, the features to be updated are randomly selected, without replacement, until the proportion of the members that have been selected is equal to $z_{x,t}$.

(iv.) $z_{x,t}$, the rate of feature replacement, is (a.) positive, and (b.) lower than one.

In summary, in each period, the focal organization updates some of its features and its membership partially changes. We start by showing that the tenure of members tends to increase over time under the stated model assumptions. Then, we specify how the actual rate of updating of features, $z_{x,t}$, depends on the other elements of the model. With this specification, the increasing tenure distribution in membership triggers a self-reinforcing dynamics where the maximal rate of architectural feature updating (i.e. the potential speed of change) decreases, exposure of members to features increases, and resistance to change increases. These changes cause the maximal rate of architectural updating to decrease even more. This finally leads to our structural inertia theorem.

4.2. **Dynamics of Organizational Membership.** The first lemma that builds to our target theorem states that member tenures tend to increase over time. According to the intuition that undergirds our model, some members, original and newer entrants, get replaced in every period.

---

[3]Our results also hold if this auxiliary assumption is relaxed, provided that turnover is independent from feature updates and non-increasing. We make this stronger assumption for simplicity of exposition.

Because all members are equally likely to leave and be replaced, the tenure distribution progressively shifts to the right.

Lemma 1. *The distribution of the tenures of the members (presumably) increases with the age of the organization.*[4]

$$\mathfrak{P} \, x \, \forall \, t, t' \, [\Omega(x) \wedge (t < t') \rightarrow W_{x,t} \prec W_{x,t'}].$$

4.3. **Modeling Exposure and Individual Resistance to Change.** A member's exposure to any feature value is generally zero at the beginning of a member's involvement in the organization and at all subsequent times at which the members perceives that the feature has some different value. The maximum possible exposure to a particular value of a feature equals the minimum of two quantities: (1) the length of a member's tenure in the organization in the case of continual perception of the fact and (2) the time that has elapsed since the feature has had this value has come to existence.

Definition 5. Let $w_{y,t}$ denote the number of (whole) periods that has elapsed between the time at which member y joined the focal organization and time t, and let $u_{v,t}$ denote the number of (whole) periods that has elapsed between the time at which the value of the feature changed to v and time t. Then, $e_{v,y,t} = g_x(\min(w_{y,t}, u_{v,t}))$, where g_x is a non-negative integer-valued increasing function.

We build on the core intuition that resistance slows change. A potentially valuable way of formalizing this notion employs the language of defaults to express taken for granted assumptions. Following prior work, we use a modal language (Polós et al. 2010). The key modality for the present argument concerns taken-for-grantedness. We first introduce a notation for this general concept. Then we apply it to the specific case of structural configurations. According to our reading of the intuitions behind the prior work, an agent likely treats some "fact" as a default if she repeatedly perceives that the fact remains true over a period. In developing this argument, we use the following notation:[5]

Notation. $\boxed{\text{D}}_y^t \, (a_t = v)$ stands for "The focal agent y takes for granted that the value of feature a equals v at time t."

---

[4]$\prec$ denotes the usual stochastic order.
[5]Polós et al. (2010) provide the possible-world semantics for this default operator.

As we explained above, members are more likely to take an architectural feature as a default when they have had extensive exposure to it. We formalize this intuition in the following postulate:

Postulate 1. *Consider an agent y, an architectural feature* a *and a feature value v. The probability that agent y takes feature value* $\mathbf{a} = v$ *as a default (normally) is increasing in her exposure to the value v*[6]

$$\mathfrak{N} x \exists D_x \left[(D_x \in \mathfrak{D}) \right] \wedge \forall y, a, t \exists v \left[(\Pr\{\llbracket \boxed{\mathrm{D}} \rrbracket_y^t (a_t = v) \rrbracket = 1\} = D_x(e_{v,y,t})) \right].$$

where $\mathfrak{D}$ denotes the set of increasing functions from the non-negative integers to [0,1]; and $\llbracket \boxed{\mathrm{D}} \rrbracket_y^t (a_t = v) \rrbracket$ refers to the standard notation for the truth value of the formula $\boxed{\mathrm{D}} \rrbracket_y^t a_t = v$: it equals one if the formula is true and zero otherwise.

We model resistance to change by members at the level of particular architectural features. Let $\pi_{v,y,t}$ denote the intensity with which the agent y will resist any attempt to change the value of the focal architectural feature at time t.

Postulate 2. *An agent's culturally based resistance to an attempt to change an architectural feature (normally) is uniquely determined by whether the agent takes the current feature value as a default.*

$$\mathfrak{N} x \exists \hat{\nu}_x, \check{\nu}_x \forall y, v, t \left[(\hat{\nu}_x > \check{\nu}_x > 0) \rightarrow \pi_{v,y,t} = \hat{\nu}_x \llbracket \boxed{\mathrm{D}} \rrbracket_y^t (a_t = v) \rrbracket + \check{\nu}_x (1 - \llbracket \boxed{\mathrm{D}} \rrbracket_y^t (a_t = v) \rrbracket) \right].$$

4.4. **Aggregating to the Organizational Level.** We are interested primarily in organizational evolution, rather than in what happens at the level of the organizational member, the architectural feature, or the member-feature pair. Therefore, instead of focusing on specific units, we examine the dynamics of the *distributions* of the tenures of members and of the features. In order to do so, we introduce some additional notations that refer to the state of organization x at time t:

- $E_{x,t}$: the distribution of exposure of a randomly selected pair of a member and a feature value (assuming that features and members are independently picked, and that all the members and all the features are equally likely to be picked);
- $R_{x,t}$: the distribution of the resistance to change that characterizes a randomly selected member-feature pair (assuming that features and members are independently picked, and that all the members and all the features are equally likely to be picked);

---

[6]See Appendix A for details on the "normally" quantifier and other quantifiers used in the statements of postulates, propositions and theorems.

Postulates 1 and 2 jointly imply that longer exposure leads to stronger resistance to change at the level of the organization. We regard this implication as a formalization of Selznick's (1957) claim that even practices chosen on purely technical grounds acquire a moral standing if retained long enough.

Proposition 1 (Selznick's law)**.** *The distribution of resistance to change in an organization's architecture (presumably) increases with the distribution of exposures in that organization.*

$$\mathfrak{P} \, x \, \forall \, t, t' \, [\Omega(x) \wedge (E_{x,t} \prec E_{x,t'}) \to R_{x,t} \prec R_{x,t'}].$$

Proof. The proof of this proposition, and of all other formal results, is in Appendix B.   □

Then we implement the intuition that strong cultural resistance complicates and limits architectural change by invoking the following postulate.

Postulate 3. *The larger is the distribution of cultural resistance within an organization, (normally) the lower is the potential speed of architectural change.*

$$\mathfrak{N} \, x \, \forall \, t, t' \, [(R_{x,t} \prec R_{x,t'}) \to \psi_{x,t} > \psi_{x,t'}].$$

This postulate connects resistance to change prevailing in the organization at the beginning of period t and potential speed of architectural change in that period. Therefore potential speed of change refers to what happens between the beginning of period t and the beginning of period $t+1$.

4.5. **Actual versus Potential Speed of Change.** At a theoretical level, we decided to define structural inertia in terms of *potential* speed instead of *actual* speed, because we believe that the concept of structural inertia should reflect constraints on what an organization can do. Potential speed need not be realized, it is an unobservable construct. This implies, in turn, that whether an organization faces age-related organizational inertia cannot generally be directly observed. What can be observed is whether actual speed goes down with age. But observing that actual speed goes down with age only warrants inferences about structural inertia when there is a clear connection between actual speed and potential speed.

So we need some persistent feature of an organization that makes a systematic link between potential and actual speed. We recognize that some "fast" organizations persistently change as much as they can; other "slow" organizations use up only some of the potential. If organizations

can switch from fast to slow modes, then we cannot make inferences over organization lifetimes. So we restrict our argument to apply to cases in which each organization has a stable propensity to take up the potential. We make this narrowing of the argument by introducing and including in the antecedents of the relevant formulas a predicate, $\text{CP}(x)$, that reads as "the organization x's actual rate of change is a constant (non-zero) proportion of its potential rate."

Definition 6 (Constant Change Proportion)**.**

$$\text{CP}(x) \leftrightarrow \exists\, c_x \left[(c_x > 0) \wedge \forall\, t \left[(z_{x,t} = c_x \psi_{x,t}) \right] \right].$$

4.6. **Dynamics of Architectural Feature Updating.** We now have all the elements in place to describe the dynamics of the tenure distribution of architectural feature:

Lemma 2. *For organizations that have a constant change proportion, the tenure distribution of features (presumably) stochastically increases with organizational age.*

$$\mathfrak{P}\, x\, \forall\, t, t' \left[\Omega(x) \wedge \text{CP}(x) \wedge (t < t') \rightarrow U_{x,t} \prec U_{x,t'} \right].$$

The proof of this result is more complicated than the proof of Lemma 1. The reason is that in the case of organizational membership, we assumed that the turnover rate was constant, which lead to a simple proof. By contrast, the rate of feature updating goes down with time as the distribution of exposure increases over time. The increase in the exposure distribution is partly driven by the increase in the distribution of the tenures of memberships and to a positive feedback loop that makes increases in the tenure distribution of architectural feature lead to further increases of the tenure distribution of features. These interacting stochastic processes unfortunately imply that it is almost impossible to state a simple, intuitive explanation for this result. In this case the formal proof provides the insight.

4.7. **Main Result.** The fact that the two tenure distributions of membership and architectural features are increasing implies that the distribution of exposures also increases. This, in turn, implies that the distribution of resistance to change increases. That is, members become more and more inclined to oppose updates of features. These results are formalized in the following lemma:

Lemma 3. *For organizations that have a constant change proportion,*

[A.] *the distribution of exposures by organizational members to architectural features (presumably) increases with organizational age:* $\mathfrak{P} x \, \forall t, t' \, [\Omega(x) \wedge \text{CP}(x) \wedge (t < t') \rightarrow E_{x,t} \prec E_{x,t'}]$.

[B.] *the distribution of cultural resistance within the focal producer (presumably) increases with organizational age:* $\mathfrak{P} x \, \forall t, t' \, [\Omega(x) \wedge \text{CP}(x) \wedge (t < t') \rightarrow R_{x,t} \prec R_{x,t'}]$.

The foregoing argument yields the central theoretical result:

Theorem 1. *[Structural inertia theorem] Organizations that have a constant change proportion are subject to structural age-related inertia.*

$$\mathfrak{P} x \, [\Omega(x) \wedge \text{CP}(x) \rightarrow \text{SI}(x)].$$

4.8. **Discussion of the Constant-Change-Proportion Condition.** Theorem 1 invokes the constant-change-proportion condition as a sufficient condition for structural inertia but does not explicitly address what happens to organizations that fail to satisfy this condition. To understand what happens in this case, consider an organization x whose change proportion parameter is temporarily lower in one period than its baseline level c_x. This makes the organization "age" faster: the tenure distribution of architectural features becomes larger than that of a similar organization whose change proportion parameter would not have temporarily decreased. Interestingly, if in the following period the organization then changes as much as implied by its baseline change proportion c_x, it can to some extent "rejuvenate" itself: the tenure distribution of architectural features does not have to increase in that period. This possibly breaks the monotonicity of the stochastic process underlying the increase in the distribution of exposures and, in turn, the distribution of resistance to change. This implies that those organizations whose change proportion parameter varies over time are not necessarily subject to structural inertia, as per our (strict) definition that invokes a *monotonic* decline in potential rate of feature updating. Importantly, the lack of necessary monotonicity that happens when organizations temporarily decrease their change proportion parameter does not mean that those organizations are less subject to constraints on their potential speed of change than organizations whose change proportion parameter remains constant: the extent of the "rejuvenation" cannot fully compensate for the accelerated aging implied by the decrease in change proportion. Much to the contrary, the potential speed of an organization that temporarily reduces

its change proportion becomes lower than the potential speed of a similar organization with a constant change proportion parameter. This is the case in all the periods following the temporary reduction in change proportion. This result is formalized in Theorem 2A below.

Another way the constant-change-proportion condition can fail to be satisfied is when an organization whose c_x parameter is temporarily lower than 1 uses a larger proportion of its potential speed of change. In this case, changing more than the baseline proportion slows down the aging process: the potential speed of change becomes higher than it would have been, had the change proportion parameter remained constant. This result is formalized in Theorem 2B:

Theorem 2 (Varying Change Propensity and Aging). *Let $\overline{\mathrm{CP}}(x, t^*)$ be a predicate that is true when the change proportion is equal to c_x in all periods except in period t^* when it is equal to $c_x^* \neq c_x$. Consider an organization x which evolves according to the assumptions summarized in $\Omega(x)$ and for which $\overline{\mathrm{CP}}(x, t^*)$ is true. Let $\alpha(x)$ be an organization that is the same as x until the beginning of period t^* (it has the same architecture, the same distribution of tenures of memberships and the same distribution of tenures of architectural features and $c_{\alpha(x)} = c_x$). The only difference between $\alpha(x)$ and x is that $\alpha(x)$'s change proportion does not change in period t^*. The change proportion of $\alpha(x)$ is thus constant and equal to c_x in all periods.*

[A.] if $c_x^* < c_x$, then the actual rate of change of x is lower than the actual rate of change of $\alpha(x)$ in all the periods starting in t^*. This also applies to the potential speed of change.

$$\mathfrak{P} \, x, t^* \, \forall t \, [\Omega(x) \wedge \overline{\mathrm{CP}}(x, t^*) \wedge (t \geq t^* + 1) \rightarrow (z_{x,t} < z_{\alpha(x),t}) \wedge (\psi_{x,t} < \psi_{\alpha(x),t})].$$

[B.] if $c_x^* > c_x$, then the actual rate of change of x is higher than the actual rate of change of $\alpha(x)$ in all the periods starting in t^*. This also applies to the potential speed of change.

$$\mathfrak{P} \, x, t^* \, \forall t \, [\Omega(x) \wedge \overline{\mathrm{CP}}(x, t^*) \wedge (t \geq t^* + 1) \rightarrow (z_{x,t} > z_{\alpha(x),t}) \wedge (\psi_{x,t} > \psi_{\alpha(x),t})].$$

5. Inferences From Movement in the Space of Offer Portfolios

Our theoretical developments have lead to theorems that make claims about the dynamics of architectural feature updating. It is often difficult, however, to observe architectural features reliably, not least how they are updated over time. But researchers frequently have good access to what organizations offer in terms of product or patent portfolios. Besides, they can use archival sources to retrieve how these portfolios evolved over time (e.g. Carroll et al. (2010); Sørensen and Stuart

(2000)). Therefore, they might want to use what they observe in terms of offer portfolios in order to make inferences about the operation of age-related structural inertia: if firms alter their offer portfolios with decreasing speed, this should suggest that age-related inertia is operating. While making such inference is attractive, it is fraught with potential pitfalls. In this section, we emphasize that making such inference requires some (strong) assumption regarding the relation between organizational architecture and offer portfolios.

In representing offers we distinguish between a particular offer (e.g., an auto model, a restaurant menu item) and an organization's portfolio of offers (e.g., an auto firm's array of models or a restaurant's menu). Let $\mathcal{O}$ denote the space of all possible offers. A particular offer of the focal organization at time t is denoted by $o_{x,t}$. An organization's portfolio of offers, in notation $\mathbf{p}_{x,t} = \{\mathbf{o}_{1,x,t}, \ldots, \mathbf{o}_{M,x,t}\}$, is an element in the powerset of the positions in offer space, denoted by $\mathcal{P}(O)$. Let $D^{\mathcal{O}}(\mathbf{p}_1, \mathbf{p}_2)$ denote the distance between two offer portfolios $\mathbf{p}_1$ and $\mathbf{p}_2$. $D^{\mathcal{O}}$ is a set-to-set distance, since the portfolios are sets of offers. A simple example of such distance is the Jaccard distance, but other distance choices might be possible depending on how the analysts chooses to represent offers and offer portfolios. When information about the distance or similarity of individual offers is available, it is possible to use some generalization of the Jaccard distance (Kovács and Hannan 2013), the Hausdorff distance (Burago et al. 2001; Le Mens et al. 2014), or a modified Hausdorff distance (Dubuisson and Jain 1994).

To relate the dynamics of architectural updating with what happens in the space of offer portfolios, we need to posit the existence of a systematic relation between the distances in both spaces. We define a predicate $\text{ISO}_1(\langle \mathcal{A}, d \rangle, \langle \mathcal{P}(O), D^{\mathcal{O}} \rangle)$ that is true when:

1. there is an isomorphic mapping from architecture space to the space of offer portfolios:
$$\omega : \mathcal{A} \longrightarrow \mathcal{P}(\mathcal{O});$$

2. distances ordering in the space of offer portfolios are preserved in architectural space.

$$\forall \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4 \in \mathcal{P}(\mathcal{O}) \left[D^{\mathcal{O}}(\mathbf{p}_1, \mathbf{p}_2) < D^{\mathcal{O}}(\mathbf{p}_3, \mathbf{p}_4) \rightarrow d(\omega^{-1}(\mathbf{p}_1), \omega^{-1}(\mathbf{p}_2)) < d(\omega^{-1}(\mathbf{p}_3), \omega^{-1}(\mathbf{p}_4)) \right].$$

This predicates means that a portfolio of offers corresponds to a unique architectural configuration and that distance ordering are the same in both spaces: very different offer portfolios can only be produced by very different architectural configurations. When this predicate holds, we can make inferences from speed of movement in the space offer portfolios about the presence or absence

of structural inertia. We define the speed of movement of an organization x in the space of offer portfolios as the ratio of the distance traveled to the time elapsed, as we did in architectural space.

Definition 7 (Speed in space of offer portfolios)**.** $z_{x,t}^{\mathcal{O}} = D^{\mathcal{O}}(\mathbf{p}_{x,t}, \mathbf{p}_{x,t+1})$

The following proposition provides a sufficient set of conditions for making inferences about structural inertia on the basis of observations of declining speed in offer space.

Proposition 2 (Inferences about structural inertia)**.** *Consider an organization that has a constant change proportion. If the mapping from architecture space and the space of offer portfolios satisfies the condition* $\text{ISO}_1(\langle \mathcal{A}, d \rangle, \langle \mathcal{P}(O), D^{\mathcal{O}} \rangle)$*, observing that the speed of change in offer space declines with age implies that the organization is subject to structural inertia:*

$$\text{ISO}_1(\langle \mathcal{A}, d \rangle, \langle \mathcal{P}(O), D^{\mathcal{O}} \rangle) \to \mathfrak{P}\, x\, \forall t, t'\, [\Omega(x) \wedge \text{CP}(x) \to ((z_{x,t}^{\mathcal{O}} > z_{x,t'}^{\mathcal{O}}) \to (\psi_{x,t} > \psi_{x,t'}))].$$

The isomorphic mapping condition is strong, but we do not see how to guarantee the correctness of inferences from the space of offer portfolios to architectural space without it. Oftentimes, the only way to proceed will be to make a qualitative assessment of the extent to which studying what happens in offer space is an appropriate proxy for what happens in architectural space.

Part of the assessment requires an analysis of the distance measures used in both spaces. A particular distance measure in the space of offer portfolios might reflect architectural changes more accurately than others. For example, a measure that is subject to strong influence by just one offer in the portfolio (such as the standard Hausdorff distance) might be less likely to reflect an architectural change than a measure that reflects global changes to the portfolio rather than local changes such as modified Hausdorff distance due to Dubuisson and Jain (1994) and used in tour empirical analyses. But there are settings where the existence of an appropriate choice of distance measure is unlikely. For example, in settings where organizations can change their architectures without substantially altering their offer portfolios, observing a slowing down in offer space does not necessarily imply that age-related structural inertia is operating. This can happen, for example, when organizations decide to externalize or internalize part of the production process while not altering the design of the product that is put on the market. Another example consists in settings where organizations externalize customer service to an independent call center or to a different country with cheaper labor costs while (hoping) it does not change the quality of the support provided to customers.

Finally, it is worth noting that inferences from offer space about age-dependent structural inertia can also break down if the constant-change-proportion condition is not satisfied. Then firms might decide to reconfigure their architectures more slowly than they could, which implies they move more slowly than what could in offer space. So, although a slowing down in offer space might reflect a slowing down in architecture space, it might be due to a voluntary slow-down. Similarly, if an organization decides to use more of its change potential than it did in the past, this can lead to an increase of the actual speed of change while the potential speed is decreasing.

6. Empirical Illustration: Declining Speed in the Space of Offer Portfolios

We reanalyze one of the datasets used by Sørensen and Stuart (2000) and show how we can do an analysis to document declining speed in the space of offer portfolios.[7] We intend this analysis to serve mainly as a proof of concept that empirical estimation that builds on our approach can yield sensible patterns.

The earlier study finds that firms in semiconductor manufacturing and biotechnology actually patented at higher rates at older ages, which appears to contradict at least the spirit of the inertia argument. However, Sørensen and Stuart (2000) argue that older firms produce more incremental than and radical innovation. They find that the patents of older firms had less impact on future discovery. Moreover, older firms tended to cite older patents and to remain close to the firm's previous work in the sense of having more citations to their own existing patents.

Taken together, these findings seem to cast some doubts on the possibility of coming up with an unequivocal answer to the question of whether older firms innovate more quickly than younger firm. Sørensen and Stuart (2000)'s findings go both ways: older firms innovate more (they create more patents) and less (their individual patents lie closer to existing knowledge). Although these results provide a refined understanding of organizational aging processes, they have limited utility to a policy maker who tries to decide whether to allocate resources to older or younger firms. A unidimensional ranking of organizations in terms of innovativeness would greatly facilitate this kind of decision. Our conception of speed of change provides a measure on which to base such a ranking.

We analyze only the biotechnology firms. We think that this setting fits the constant-change-proportion condition that is relied upon in the formulation of the structural inertia theorem. In science and drug discovery, being first to publish a result or patent a drug makes an enormous

---

[7] We thank Jesper Sørensen for sharing the data.

difference. The agents involved are highly motivated to move to the research frontier at their full potential, therefore we believe that the change proportion for firms in this industry is constant equal to very close to 1 ($c_x = 1$). In addition, we believe that this setting also satisfies the isomorphic mapping condition $\text{ISO}_1(\langle \mathcal{A}, d \rangle, \langle \mathcal{P}(O), D^{\mathcal{O}} \rangle)$. Most biotechnology firms are highly specialized organizations structured with the sole purpose of producing innovations at the boundary of the technological frontier. Because patented innovations are rarely the outcome of chance but instead the result of purposeful and long-term research efforts, distances in patent space likely reflect distance in architectural space.

6.1. **Data and Measurement.** Sørensen and Stuart (2000) chose a random sample of 237 dedicated biotechnology firms listed in *The Corporate Technology Directory* in 1988. They analyzed patents granted to these firms during 1987–1994. We analyze those firms and time period, but we use the data differently: we see the patent portfolio of a firm as its *offer portfolio* and we compute the speed at which firms reconfigure their portfolios, as specified in the previous section. There are 61 firms which own at least one patent for which they applied before 1987. We restrict our analysis to these firms because these are the firms for which we are able to compute their initial position in the space of patent portfolios at the beginning of the observation period. If we find that speed of change declines with age, it will suggest that firms are subject to structural inertia (assuming that the conditions for making such inference hold—see Proposition 2 above).

Space of Offer Portfolios. The space of offer portfolios is the powerset of the set of patents.

Distance Measure between Patent Portfolios. We associate patents with subclasses of prior art. We use the NBER patent database (Hall et al. 2001) which contains 2458 subclasses, each with a subclass number. We associate an index with the full set of subclasses and keep this index fixed for all patents and we associate a vector of cited subclasses with each patent. For example, patent 4854172 cites two patents: patents 4766647 and 4292847. The first of these patents is in subclass 0.22 (as primary subclass), which is the smallest subclass number in our data and therefore has index equal to one. The second of these cited patents is in subclass 587, which is the 1974th largest subclass number. Therefore, the focal patent (4854172) is represented by a 2458-long vector of zeroes almost everywhere, and ones at slots 1 and 1974. More formally, the offer space is thus the space of patents, characterized in terms of the 2458 subclasses of the database we use: $\mathcal{O} = \{0, 1\}^{2458}$.

With this representation, we measure the distance between two patents as the number of "slots" that differ between the vectors that represent these two patents. Note that if a patent cites two patents in the same subclass, we put a one in the corresponding slot (citing more than one patent in a given subclass does not change the representation of the patent). More formally, the distance between two patents p_a and $p_{a'}$ whose representations in $\mathcal{O}$ are vectors v and v' is denoted by $d^{\mathcal{O}}(v, v')$. It is the Hamming distance between v and v'.

On the basis of the distance between offers, $d^{\mathcal{O}}$, we construct a distance between offer portfolios. We use the modified Hausdorff distance, sometimes used in computer graphics to measure the similarity between objects (Dubuisson and Jain 1994). This distance measure is interesting, because it quantifies the difference between the patent *portfolios* by building on our measure of distance between *patents*, $d^{\mathcal{O}}$. It is defined as follows. We first construct a directed distance from one portfolio to the other, denoted by $\vec{h}(\mathbf{p}, \mathbf{p'})$, that records the average of the minimum distances of the objects (offers) in $\mathbf{p}$ to the objects in $\mathbf{p'}$: $\vec{h}(\mathbf{p}, \mathbf{p'}) = \frac{1}{|\mathbf{p}|} \sum_{\mathbf{o} \in \mathbf{p}} \inf_{\mathbf{o'} \in \mathbf{p'}} d^{\mathcal{O}}(\mathbf{o}, \mathbf{o'})$. The modified Hausdorff distance between the portfolios is the maximum of the two directed distances: $D^{\mathcal{O}}(\mathbf{p}, \mathbf{p'}) = \max(\vec{h}(\mathbf{p}, \mathbf{p'}), \vec{h}(\mathbf{p'}, \mathbf{p}))$.

Speed of Change. Consider a firm x at the beginning of year t. It is characterized by its patent portfolio $\mathbf{p}_{x,t}$, the set of granted patents with application dates until the end of year t. Its speed of change in that year is defined on the basis the distance between the patent portfolios at times t and t':[8] $z_{x,t}^{\mathcal{O}} = D^{\mathcal{O}}(\mathbf{p}_{x,t}, \mathbf{p}_{x,t+1})$. Intuitively, it measures the average distance of new patents from the existing patent portfolio. If the focal firm does not patent in year t, then the patent portfolio does not change and its speed of change between t and $t+1$ is 0.

6.2. **Descriptive Graph.** We can get a sense of the overall relationship between speed in patent space and age (including the between-firm variation) from the plot in Figure 1. It shows the fit of a local polynomial smoothed function that relates speed of change and age with 95% confidence intervals. We see a negative relation of age and speed. Speed is high at young ages and goes down to relatively close to 0 at old ages. This pattern suggests that the overall effect of age on speed is negative. Next, we evaluate this claim using econometric estimations.

6.3. **Results.** We adopt a conservative approach in testing the theory by fitting fixed effects for firm with OLS. In other words, we use only *within-firm* variation in age and speed to test the theory.

---

[8]In order to do so, we had to supplement Sørensen-Stuart's data to characterize all of the patents held by these firms – not only patents in the biotechnology area. We did this on the basis of the NBER patent database.
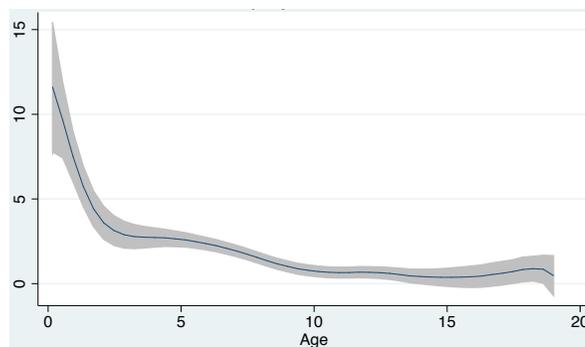
FIGURE 1. Smoothed local polynomial fit of the speed in the space of patent portfolios as a function of organizational age.

In this kind of estimation, we can learn whether the same firm, on average, moves more slowly at later ages than at earlier ages. Following Sørensen and Stuart (2000) we control for the size of firm patent portfolios and sizes (number of employees), both measured in the prior year.

Estimates of these firm-fixed-effect regressions of the speed on age and other covariates can be found in column 1 in Table 1. The first-order effect of age is negative, and the second-order effect is positive. According to these estimates, the relation between age and speed turns from negative to positive when age reaches 20 years. Nearly all of the observation (98%) pertain to ages in the range $[0, 20]$. So the implied upturn in the relationship at older ages does not depend on much empirical support. In other words, the data fit a pattern in which the speed of movement in technical space declines with age at a decreasing rate—the same pattern we see in the plot reported in Figure 1.

We also tried replacing the firm effects with year-fixed-effects. The results in column 2 show that the size of the effect of age is lower, but it remains negative and significant. Even with the positive square term, the relation between age and speed turns from negative to positive when age reaches 19 years. Nearly all of the observation (98%) pertain to ages in the range $[0, 19]$. So the implied upturn in the relationship at older ages does not depend on much empirical support. Therefore the implied pattern of age dependence is very similar to that implied by the results in column 1.

Additional analyses reported in Appendix C demonstrate that our main result holds with alternative ways of computing the speed of change. We also use a standard Heckman (1979) two-step selection specification but without firm fixed-effects (because the two-step estimator is inconsistent for the combination of fixed effects and the correction for sample selection). The reason we use such a model is that we face the complication of right-censoring of spells between patents. That is, some records show a firm has patented at some time during the study interval but then did not do so for

TABLE 1. Effects of age on speed in the space of patent portfolios, OLS estimates with fixed effects for firm (column 1) and for year (in column 2).

	(1)	(2)
Age	-1.07^{***}	-0.36^{***}
	(0.20)	(0.055)
Age squared/100	2.67^{*}	0.957^{***}
	(1.07)	(0.199)
Cumulative patents/100	0.009	-0.009
	(0.006)	(0.007)
Size/1000	0.34	0.11
	(0.23)	(0.15)
Constant	8.46^{***}	3.87^{***}
	(0.80)	(0.32)
Observations	440	440
Number of firms	61	61
$F(4,60)$	48.1	
$F(4,428)$		85.8

Standard errors (clustered by firm for Model 1) are in parentheses.
$^{***}\ p < 0.001$; $^{**}\ p < 0.01$; $^{*}\ p < 0.05$.

the rest of the study period. These observations contain information relevant for our argument—not moving at all obviously means that speed is zero.

6.4. **Conclusion.** Our empirical analyses demonstrate that actual speed in the space of patent portfolios declines with age. But what about the evolution of *potential* speed of architectural change? As stated in Prop. 2, observing declining speed of change in the space of offer portfolios implies that age-related structural inertia is operating when the isomorphic mapping assumption and the 'constant-change-proportion' conditions hold. Although we have good reasons to believe these assumptions hold in this setting, we have not tested them. Thus our results can only suggest that age-related structural inertia operates here to the extent that these conditions hold.

7. DISCUSSION

7.1. **Relation to Existing Explanations for Age-Dependent Structural Inertia.** Our model proposes an organizational-level mechanism to explain why structural inertia frequently increases with organizational age. We invoke the strong tendency for individuals to like the familiar to postulate that organizational members will tend to oppose alterations to organizational features to which they have been long exposed. Our model shows that this individual-level assumption and the fact that the distribution of exposure to architectural features by organizational members increases

with time leads to an organizational level outcome: the emergence of age-related structural inertia. Our explanation thus most closely resembles explanations that focus on routines and organizational memory (see Section 3). However, the postulated process is different: our model relies on the idea that organizational members will oppose change when management tries to implement it. Other explanations emphasize that managers solve problems as they arise — they fall in the tradition of the behavioral theory of the firm (Cyert and March 1963). When organizational memory is rich, organizational members likely see new problems as similar to previously encountered ones and apply existing routines instead of devising new solutions. These explanations complement each other, and integrating them in a unifying model is an interesting avenue for future research.

7.2. **Relation to Prior Research on Age-Dependence in Failure Hazards.** Our model focuses on the dynamics of architectural feature updating. To derive the implications for organizational performance, specifically resource accumulation or organizational survival, architectural features and processes must be tied to resource accumulation. In the theoretical program on which we build (Le Mens et al. 2011; Le Mens et al. 2014), organizations accumulate resources when their offers match the relevant audience's tastes. Structural inertia will thus affect the flow of resource accumulation to the extent that its architecture constrains what an organization can do in terms of offers. Perhaps the simplest way to draw a connection between architectural dynamics and offers on the market is to invoke a variant of the isomorphism condition introduced in Section 5. This condition is directional: it is defined in terms of implications of what happens in offer space for what we can infer in architecture space. Here, we need a condition that goes in the other direction:. Let us define another predicate, $\text{ISO}_2(\langle \mathcal{A}, d \rangle, \langle \mathcal{P}(O), D^{\mathcal{O}} \rangle)$ that is true when:

(1) there is an isomorphic mapping from architecture space to the space of offer portfolios: $\omega : \mathcal{A} \longrightarrow \mathcal{P}(\mathcal{O})$;

(2) distances ordering in architectural spaces are preserved in the space of offer portfolios.

$$\forall \, \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4 \in \mathcal{P}(\mathcal{O}) \, [d(\mathbf{a}_1, \mathbf{a}_2) < d(\mathbf{a}_3, \mathbf{a}_4) \rightarrow \, D^{\mathcal{O}}(\omega(\mathbf{a}_1), \omega(\mathbf{a}_2)) < D^{\mathcal{O}}(\omega(\mathbf{a}_3), \omega(\mathbf{a}_4))].$$

When this predicate holds, what the dynamics that occur in architecture space essentially translate in offer space. An interesting implication of Theorem 1 is that when the 'constant-change-proportion' condition holds, the actual speed of change in architectural space is a decreasing function of organizational age. If, in addition, the $\text{ISO}_2(\langle \mathcal{A}, d \rangle, \langle \mathcal{P}(O), D^{\mathcal{O}} \rangle)$ condition holds, this implies

that the actual speed of change in the space of offer portfolios declines with age. This, in turn, implies that it converges downward to a limiting value (because that is bounded by zero).

This implication bears on an argument delineated by Le Mens et al. (2014) which explains the increase in failure hazard for old organizations as a consequence of drift in audience tastes. Given such drift, there is a point in time after which the producer cannot move as fast in offer space as do the audience tastes. This argument postulated that the limiting potential speed in offer space was zero (this was part of Post. 2 in that paper).

The results of the current paper amend this claim: the limiting potential speed in architectural space does not have to be zero. If the limiting potential speed exceeds the speed of drift in audience tastes, then obsolescence might never happen. Obsolescence will occur for sure only if the limiting potential speed in the space of offer portfolios falls below the speed of the drift in audience tastes. According to the argument developed in this paper, the first scenario (where obsolescence never happens) is more likely to occur when turnover of members remains high at all ages. In this case, exposure to features remains relatively low, and so does the resistance to change.

When is the $\text{ISO}_2(\langle \mathcal{A}, d \rangle, \langle \mathcal{P}(O), D^{\mathcal{O}} \rangle)$ likely to hold? Intuitively, when there is a tight coupling between architecture and offers. Such tight coupling is likely when there is little possibility of externalization of the design and production process. This applies, for example, to high tech firms that commercialize products that are close to the technological frontier. On the other hand, this condition is less likely to hold in settings where firms can move substantial parts of the design or production process across firm boundaries with little impact on the perception by customers.

7.3. **Choice of Distance Measure.** In Section 2, we defined the distance between architectures as a Hamming distance and we used this measure for the distance between patents in knowledge space, in Section 6. This is one of the simplest distance measures, because it gives equal weight to all the attributes and the contribution of each attribute is equal to 0 or 1. We made this assumption to ensure tractability of our dynamic model and to ensure we could derive formal predictions. But arguments like ours can use distance measures that give different weights to different architectural features or to additions and eliminations features, e.g., the differential costs of adding or eliminating organizational units. And, instead of seeing organizations as points in a metric multidimensional space of architectural features, one can regard organizations as trees or networks of inter-connected units and characterize the difference between architectures by relying on distance over spaces of trees or networks. We believe that the choice of distance measure for

empirical investigations should be guided by the richness and quality of available data. Yet, more complexity and specificity of the chosen distance measure complicates cross-study comparisons. Moreover, it might be difficult to decide, a priori, on a choice of attribute weights. And implementing complex estimation procedures that estimate the weights of the attributes might lead to overfitting. There is ample evidence from research on forecasting that prediction algorithms that give weights of zero or one to available attributes tend to make better our-of-sample predictions than more complex weighting rules (Dawes 1979; Kahneman 2011). Applied to the context of measure distance between organizations, this suggests it might not be easy to come up with the distance measures that lead to more accurate predictions than a simple Hamming distance. Systematically analyzing how results of empirical estimation of organizational speed change with the choice of distance measure would be an interesting follow up to this paper.

7.4. **Discussion of the Assumptions of the Dynamic Model.** Our dynamic model relies on three postulates and a set of assumptions denoted by $\Omega(x)$. Post. 1 and 2 deal with the relation between exposure by an agent to an organizational feature and the resistance to changing that feature. For simplicity, we assumed that resistance to change could take only two levels. This restriction is not crucial to our main results but keeps the proofs relatively simple. Post. 3 connects resistance to change in an organization and the potential speed of change. Although we can think of situations where a particular leader would react to a *perceived* high level of resistance to change by trying to impose change, maybe in settings where the leader is an outsider hired to "turn around" the organization, these seem to be the exception rather than the rule. These exceptions are allowed for by the use of the "normally" quantification in the formal statement of the postulate.

A potentially limiting aspect of our model is its reliance on a set of homogeneity assumptions. According to the assumptions in $\Omega(x)$, all members and all features are equally likely to be replaced. In particular, the likelihood that a feature is replaced does not depend on the average exposure of organizational members to it. Instead, it depends only on the rate of feature turnover, which depends on the whole distribution of resistance to change (see Post. 3). We made this assumption to ensure mathematical tractability. Without it, equation 1 (in Appendix B) would have to be modified, because it relies on the assumption that the likelihood that a feature is replaced does not depend on the the tenure of that particular feature. Intuition suggests that our result will still hold if we allow the probability that a particular feature gets replaced to depend on its tenure (because average exposure to this feature is higher and thus average resistance to change to that feature is

higher). But verifying this formally would require an entirely new analysis that goes beyond the scope of the current paper.

Our main theoretical result (Theorem 1) relies on the constant change proportion condition (CP(x)). As discussed in Section 4.8, this condition ensures that the speed of change declines monotonically with age. When this condition is not satisfied, the speed of change will be higher or lower than when it is satisfied (see Theorem 2), but it does not have to decline monotonically with age. When are organizations likely to satisfy this condition? A special case of interests concerns organizations that change as much as they can ($c_x = 1$). This condition is likely satisfied when the inflow of resources from the environment dries up if they do not change. This encompasses at least two (non-exclusive) sets of situations. The first set concerns drift in audience tastes (see Le Mens et al. (2014) for a formal definition of drifting tastes). Tastes tend to drift when there are broad cultural evolutions (such as an increased tendency for individualism) that drives the dynamics of particular tastes. They also tend to drift in settings where the audience is made of different groups with some striving for distinctiveness (Brewer 1991; Bourdieu 1984). Finally, tastes might drift in response to innovation by one producer which creates expectations of a constant stream of new product features, as seen in the smartphone realm, for example. The second set of situations concerns high competitive intensity, with pervasive imitation, new entries, or the threat of new entries. In such settings, firms that do not innovate see their profit margins erode quickly because competitors quickly release products that are very similar to the products of the innovators and are thus similarly appealing to audience members. Firms are pushed to the limit of their adaptive potential, and this is where the operation of age-related structural inertia will be most visible.

7.5. **The Consequences of Restraining Architectural Changes.** Theorem 2 suggests that organizations that do not change as much as they could, even during a short period, will bear long-term consequences of this choice: their subsequent potential speed of change will be lower than it would have been had they changed to the maximum of their potential. Taken at face value, this claim might be read as a suggestion that firms should change as much as they can. We do not take this stance. Our model is about the likelihood that change will be implemented, not about the likelihood that change will lead to higher organizational performance. A given architectural change could be misdirected in the sense that the consequential change in offer portfolio might move away the offers from the audience's tastes, which reduces their appeals and thus the ability of the organization to extract resources from its environment.

Nevertheless, our model has implications for change management. Theorem 2A suggests that organizations that want to preserve their adaptive potential should keep changing some of their architectural features, even if these changes are inconsequential from an organizational performance perspective. This implies that in settings where change does not have a systematic effect on performance, firms should engage in architectural changes to preserve their adaptive potential. Such attempt at preserving adaptive potential will be particularly valuable in settings where tastes might drift (even if they do not change for some time) or when competitors might engage in adaptation efforts to improve the fit of their offers to audiences' tastes (even if they do not move for some time). Settings where maintaining speed is crucial to 'remain in the game' as competitors also engage in adaptive efforts will lead to dynamics akin to the Red Queen competition (Barnett and Hansen 1996) where firms engage in successions of exploitive moves that make them better on an absolute level but not better relative to their competitors (because competitors also improve).

We want to emphasize, however, that in settings where changes do not systematically lead to performance increase (e.g. when changes concern the core features of the organizations, or have an explorative nature), the prescriptions of our model are less clear cut. In such settings, firms have to balance the preservation of their adaptive potential with the risks associated to change. Pushing the parallel with research on Red Queen competition further, it is worth noting that there is evidence that those firms which have an history of exploitive changes tend to be particularly poor at core changes (Barnett and Pontikes 2008). This suggests that preserving adaptive potential by engaging in exploitive moves with the expectation it will prove useful for deeper organizational change later on might be a wrong idea.

7.6. Increasing Turnover, Overcoming Resistance to Change, and Organizational Performance.

Our model also has implications for what happens when an organization manages to overcome resistance to change and move over a path of changes in the architectural space (against Post. 3) or has a wide scale replacement of members (against the assumption of constant turnover that is part of $\Omega(x)$). In our model, either scenario sets exposure to zero for each of the features on the path. This means that resistance to further changes in those features is low and thus the distribution of resistance to change prevailing in the organization decreases. This implies, in turn, that the producer's speed of change rises. So the inertial force weakens in a fashion similar to that of Theorem 2B. What does this mean for organizational performance?

The theory we propose does not provide a coherent answer to this question. Considerations of organizational learning (as in the competency-trap story) and stability of organizational routines (Hannan and Freeman 1984; Barnett and Carroll 1995) suggest that performance is degraded when organizations change their structures or their memberships. The formulation on which this paper builds (Le Mens et al. 2011; Le Mens et al. 2014) ties the quality/quantity of engagement to organizational age, not to the duration of the pairs of structural arrangements and members. This part of the argument does not allow performance to degrade with change. Here is an opportunity for a deeper integration. Recognizing that an organization that changes all of its features or replaces all of its members is essentially a new organization, it would make sense to "reset the clock" (Amburgey et al. 1993). That is, the theory could replace organizational age, as measured by an external clock that marks the passage of time since founding, with an internal clock that tracks total experience of the membership with the organization's architecture.

Another potential consequence of turnover in membership is an increase in explorative behavior and, in turn, in the variability of organizational performance. In terms of the constructs of the theoretical framework on which we build, the infusion of new members that have not been socialized to the organization (or organizational code—see March (1991)) can lead to an increase in the variability of the intrinsic appeal of the organizational offer (i.e., how close the organizational offer is from the audience's tastes—Le Mens et al. (2014)). For example, consider engineers who join an existing design team. These engineers bring with them their own ideas about audience's tastes and distinct skills that can lead the team to come up with different designs. Because these engineers are not yet socialized to the organizational culture and interpretive frames, this added variability in inputs will likely imply increased variability in organizational output. This in turn, implies an increased variability in the match between organizational offer and audience's tastes. In summary, although increased turnover can lead to a higher potential speed of change, this can also imply that actual organizational change attempts will have more uncertain outcomes. That is, a higher potential speed likely comes with a higher likelihood of mis-directed organizational changes.

Integrating these ideas about the effect of increased turnover—a systematic effect negative effect on organizational performance together with a positive effect on the variability of organizational performance—together with our ideas of the effect of increased turnover on potential speed of change seems essential to addressing whether organizations that undertake extensive change might catch up with drifting tastes. This is an exciting avenue for future research.

APPENDIX A. QUANTIFIERS

We use two quantifiers $\mathfrak{N}$, and $\mathfrak{P}$, to be read as "normally" and "presumably," respectively. Both quantifiers express the level of generality of the formulas within the scope of quantification—they are rules with exceptions, generic sentences. They place different roles in the theory. The "normally" prefix indicates that what follows is meant to be a substantive principle of the theory, while the "presumably" prefix signals that (the most current version of) the theory implies that what follows is the case, though future refinements that make the theory more specific might not support the claim. Athough the truth and falsity of the propositions prefixed with these quantifiers is not definable in terms of the number or proportion of the positive and negative instances of the generic formulas, a truth-conditional rendering of the meanings of these quantifiers is still possible. It relies on a somewhat enriched model-theoretic rendering of information states associated to theory stages (Veltman 1996). In this kind of formalization, information states might contain both factual information (expressible in first-order logic) and regularities, or generics (Pólos and Hannan 2004). A proposition of the form "Normally if x is a ϕ then x is a ψ," i.e., $\mathfrak{N}\,x\,[\phi(x) \to \psi(x)]$ is true in the context of a theory stage $\mathcal{T}$ if the pair $< [\![\phi(x)]\!], [\![\psi(x)]\!] >$ representing the regularity is present in the theory stage $\mathcal{T}$. Here $[\![\phi(x)]\!]$ and $[\![\psi(x)]\!]$ stand for truth values of $\phi(s)$ and $\psi(x)$ (which, in technical terms, are open-formula intensions) (Pólos and Hannan 2004). A proposition of the form "Presumably if x is a ϕ then x is a ψ," i.e., $\mathfrak{P}\,x\,[\phi(x) \to \psi(x)]$ is true in the context of a theory stage $\mathcal{T}$ if the most specific chains of regularities in $\mathcal{T}$ connects $[\![\phi(x)]\!]$ to $[\![\psi(x)]\!]$ but does not connect it to $[\![\neg\psi(x)]\!]$ (Pólos and Hannan 2004).

Both the "normally" and "presumably" sentences refer to regularities, defaults, rules with expectations. They differ on their inferential roles though. What is "normally" the case are those regularities, in the context of a theory, the theory is firmly committed to. What is "presumably" the case are those predictions the current phase theory is ready to make, however future extensions of the theory might withdraw these predictions.

In the papers on which we build, nonmonotonic inference played a key role: a central assumption in the first stage of the theory was over-ruled by a more specific assumption in the second stage. This is not an issue in the current paper. So all of the nonmonotonic quantifiers could be replaced by a first-order universal quantification for checking inferences. We do not do so, however, because this would break the continuity of the larger theory project and make some postulates and claims stronger than we think justified (make them universally true).

APPENDIX B. PROOFS

In the following proofs we assume we are in the "normal" case.

Proof of Selznick's Law, Proposition 1.

Suppose that x, t, and t' are such that $E_{x,t} \succ E_{x,t'}$. Let $Pr\{[\![\,\boxed{\mathrm{D}}{}_x^t\,]\!]\}$ denote the proportion of member-feature pairs that are taken as defaults in organization x at time t. We have, by the law of total probability, $Pr\{[\![\,\boxed{\mathrm{D}}{}_x^t\,]\!]\} = \sum_{e=0}^{\infty} Pr\{E_{x,t} = e\} D_x(e)$. According to Post. 1, D_x is an increasing function. This and the fact that $E_{x,t} \prec E_{x,t'}$ imply that $\Pr\{[\![\,\boxed{\mathrm{D}}{}_x^t\,]\!]\} < \Pr\{[\![\,\boxed{\mathrm{D}}{}_x^{t'}\,]\!]\}$. Post. 2 implies that the resistance to a change for a randomly selected agent-feature pair in organization x at time t equals $\hat{\nu}_x$ with probability $\Pr\{[\![\,\boxed{\mathrm{D}}{}_x^t\,]\!]\}$ and equals $\check{\nu}_x$ with probability $1 - \Pr\{[\![\,\boxed{\mathrm{D}}{}_x^t\,]\!]\}$. Similarly the resistance to such a change by a randomly selected agent in producer x at time t', equals $\hat{\nu}_x$ with probability $\Pr\{[\![\,\boxed{\mathrm{D}}{}_x^{t'}\,]\!]\}$ and equals $\check{\nu}_x$ with probability $1 - \Pr\{[\![\,\boxed{\mathrm{D}}{}_x^{t'}\,]\!]\}$. Because $\hat{\nu} > \check{\nu}$ and $\Pr\{[\![\,\boxed{\mathrm{D}}{}_x^t\,]\!]\} < \Pr\{[\![\,\boxed{\mathrm{D}}{}_x^{t'}\,]\!]\}$, it follows that $R_{x,t} \prec R_{x,t'}$. $\qquad\square$

Proof of Lemma 1.

We proceed by recursion (i.e., proof by induction). First we demonstrate that $W_{x,1} \prec W_{x,2}$. To do so, we need to show that, for all $w \geq 0$, $\Pr(W_{x,1} \leq w) \geq \Pr(W_{x,2} \leq w)$ and that there exists some w' such that $\Pr(W_{x,1} \leq w') > \Pr(W_{x,2} \leq w')$. Let $w \geq 0$. By construction $W_{x,1} = 0$. Therefore $\Pr(W_{x,1} \leq w) = 1$. Also, $\Pr(W_{x,2} \leq w) = (1 - m)\Pr(W_{x,1} \leq w - 1) + m$. There are two cases to consider. First, suppose $0 \leq w < 1$. We have $\Pr(W_{x,1} \leq w - 1) = 0$, because tenures cannot be negative. Therefore $\Pr(W_{x,2} \leq w) = m < \Pr(W_{x,1} \leq w)$. Second, suppose that $w \geq 1$. We have $\Pr(W_{x,1} \leq w - 1) = 1$. Therefore $\Pr(W_{x,2} \leq w) = 1 = \Pr(W_{x,1} \leq w)$. It follows that $W_{x,1} \prec W_{x,2}$.

Next let us suppose that $t > 1$ and $W_{x,t-1} \prec W_{x,t}$. We want to show that $W_{x,t} \prec W_{x,t+1}$. Let $w \geq 0$. We can write $\Pr(W_{x,t+1} \leq w) = (1 - m)\Pr(W_{x,t} \leq w - 1) + m \leq (1 - m)\Pr(W_{x,t-1} \leq w - 1) + m = \Pr(W_{x,t} \leq w - 1) \leq \Pr(W_{x,t} \leq w)$.

The inequality in the second row of the foregoing formula follows from the direct application of the recursion hypothesis that $W_{x,t-1} \prec W_{x,t}$. The recursion hypothesis also implies that there exists some ω' such that $\Pr(W_{x,t} \leq w') < \Pr(W_{x,t-1} \leq w')$. Algebraic manipulations like those just above imply that $\Pr(W_{x,t+1} \leq w') = (1 - m)\Pr(W_{x,t} \leq w' - 1) + m < (1 - m)\Pr(W_{x,t-1} \leq w' - 1) + m = \Pr(W_{x,t} \leq w' - 1) \leq \Pr(W_{x,t} \leq w')$. Therefore, $W_{x,t} \prec W_{x,t+1}$.

By application of the principle of induction, we have $\forall\, t, t'\, [\Omega(x) \wedge (t < t') \to W_{x,t} \prec W_{x,t'}]$. $\quad\square$

Proof of Lemma 2.

We demonstrate that $U_{x,t}$ is stochastically increasing in t by using a proof by induction in settings

where $\mathrm{CP}(x)$ holds. The first step is to show that $U_{x,t_1} \prec U_{x,t_1+1}$. To do so, we show that, for all $u \geq 0$, $\Pr(U_{x,2} \leq u) \leq \Pr(U_{x,1} \leq u)$ and that there exists some u' such that $\Pr(U_{x,2} \leq u') < \Pr(U_{x,1} \leq u')$. Let $u \geq 0$. We have $P(U_{x,1} \leq u) = 1$ because $U_{x,1}$ is constant and equal to 0 (no feature value can be in place before the organization begins operations). Because $\Pr(U_{x,2} \leq u)$ is a probability, we have $\Pr(U_{x,2} \leq u) \leq 1$. This implies $\Pr(U_{x,2} \leq u) \leq P(U_{x,1} \leq u)$. Let u' be such that $0 \leq u' < 1$. We have $\Pr(U_{x,2} \leq u') = (1 - z_{x,2})\Pr(U_{x,1} \leq u' - 1) + z_{x,2}$. The first term of the RHS of the foregoing formula pertains to the features that have not been updated in period 1, and the second term pertains to those that have been updated in period 1. Because the tenure of any newly updated feature is 0, it cannot exceed u'. Because $u' - 1 < 0$, $\Pr(U_{x,1} \leq u' - 1) = 0$ and $\Pr(U_{x,2} \leq u') = z_{x,1}$. But, we assumed that $z_{x,t} < 1$ for all t in the definition of Ω. Therefore, $\Pr(U_{x,2} \leq u') < 1 = P(U_{x,1} \leq u')$. In summary, $U_{x,1} \prec U_{x,2}$.

The second step is to show that, if $t > 1$ and the inductive hypothesis holds (i.e., if $U_{x,t-1} \prec U_{x,t}$), then $U_{x,t} \prec U_{x,t+1}$. Let $t > 1$ and assume that $U_{x,t-1} \prec U_{x,t}$. From Lemma 1, we have $W_{x,t-1} \prec W_{x,t}$. Def. 5 implies that for every organizational member y and every feature value v, the exposure of y to v is such that $e_{y,v,t} = g_x(\min(w_{y,t}, u_{v,t})$. This implies that at the level of the organization, $E_{x,t} = g_x(\min(W_{x,t}, U_{x,t})$. All-in-all, this implies that $E_{x,t-1} \prec E_{x,t}$. From Prop. 1 (Selznick's Law), it follows that $R_{x,t-1} \prec R_{x,t}$. This and Post. 3 imply that, presumably, the potential speed in architectural space is decreasing: $\psi_{x,t-1} > \psi_{x,t}$. This, the fact we assumed that the focal organization has a constant change proportion ($\mathrm{CP}(x)$ holds), and Def. 6 imply that the same presumably holds for actual speed: $z_{x,t-1} > z_{x,t}$.

We now show that $U_{x,t} \prec U_{x,t+1}$. Let $u \geq 0$. First note that

$$(1) \qquad\qquad \Pr(U_{x,t+1} \leq u) = (1 - z_{x,t})\Pr(U_{x,t} \leq u - 1) + z_{x,t}.$$

This equation implies that $\Pr(U_{x,t+1} \leq u)$ is non-decreasing in $z_{x,t}$, which together with the fact that $z_{x,t-1} > z_{x,t}$, implies:

$$(2) \qquad\qquad \Pr(U_{x,t+1} \leq u) \leq (1 - z_{x,t-1})\Pr(U_{x,t} \leq u - 1) + z_{x,t-1}.$$

The induction hypothesis, $U_{x,t-1} \prec U_{x,t}$, implies that $\Pr(U_{x,t} \leq u - 1) \leq \Pr(U_{x,t-1} \leq u - 1)$. Together with eq. 2, this gives: $\Pr(U_{x,t+1} \leq u) \leq (1 - z_{x,t-1})\Pr(U_{x,t-1} \leq u - 1) + z_{x,t-1}$. The RHS of this equation is exactly $\Pr(U_{x,t} \leq u)$. We have thus shown that $\Pr(U_{x,t+1} \leq u) \leq \Pr(U_{x,t} \leq u)$.

We now show that there exists u' such that $\Pr\left(U_{x,t+1} \leq u'\right) < \Pr\left(U_{x,t} \leq u'\right)$. The induction hypothesis implies that there exists u'' such that $\Pr\left(U_{x,t} \leq u''\right) < \Pr\left(U_{x,t-1} \leq u''\right)$. This, equation 2, and the fact that $z_{x,t-1} < 1$ imply $\Pr\left(U_{x,t+1} \leq u'' + 1\right) < \left(1 - z_{x,t-1}\right)\Pr\left(U_{x,t-1} \leq u''\right) + z_{x,t-1}$. The RHS of this equation is exactly $\Pr\left(U_{x,t} \leq u'' + 1\right)$. We set $u' = u'' + 1$. We have $\Pr\left(U_{x,t+1} \leq u'\right) < \Pr\left(U_{x,t} \leq u'\right)$. This concludes the proof of the inductive step: $U_{x,t} \prec U_{x,t+1}$.

The principle of mathematical induction implies that $U_{x,t} \prec U_{x,t+1}$ for all t. This, in turn, implies that $\forall t, t'\left[\Omega(x) \wedge \mathrm{CP}(x) \wedge (t < t') \rightarrow U_{x,t} \prec U_{x,t'}\right]$. □

Proof of Lemma 3.

Lemmas 1 and 2 and Def. 5 imply: $\mathfrak{P}\, x \,\forall t, t'\left[\Omega(x) \wedge \mathrm{CP}(x) \wedge (t < t') \rightarrow E_{x,t} \prec E_{x,t'}\right]$. This formula and Prop. 1 imply: $\mathfrak{P}\, x \,\forall t, t'\left[\Omega(x) \wedge \mathrm{CP}(x) \wedge (t < t') \rightarrow R_{x,t} \prec R_{x,t'}\right]$. □

Proof of Theorem 1. This theorem holds that the model assumptions and $\mathrm{CP}(x)$ condition yield structural inertia. A cut rule on the (rule chain warranting) Lemma 3 and Post. 3 gives $\mathfrak{P}\, x \,\forall t, t'\left[\Omega(x) \wedge \mathrm{CP}(x) \wedge (t < t') \rightarrow \psi_{x,t} > \psi_{x,t'}\right]$, which satisfies the definition of structural inertia.

□

Proof of Theorem 2.

We detail the proof for (1). The proof for (2) follows easily by changing the signs of the inequalities in the proof for (1). We will use the following Lemma (a version of Selznick's law):

Lemma 4. *The following holds* $\mathfrak{P}\, x, t^*, t\left[\Omega(x) \wedge \overline{\mathrm{CP}}(x, t^*) \wedge \left(E_{x,t} \succ E_{\alpha(x),t}\right) \rightarrow R_{x,t} \succ R_{\alpha(x),t}\right]$.

Proof. Suppose that x, t, and t^* are such that $E_{x,t} \succ E_{\alpha(x),t}$. Let $Pr\{[\![\,\boxed{\mathrm{D}}\,]_x^t\,]\!]\}$ denote the proportion of member-feature pairs that are taken as defaults in organization x at time t. Similarly, $Pr\{[\![\,\boxed{\mathrm{D}}\,]_{\alpha(x)}^t\,]\!]\}$ denote the proportion of member-feature pairs that are taken as defaults in organization x at time t. We have, by the law of total probability, $Pr\{[\![\,\boxed{\mathrm{D}}\,]_x^t\,]\!]\} = \sum_{e=0}^{\infty} Pr\{E_{x,t} = e\}D_x(e)$. We have $D_x = D_{\alpha(x),t}$, thus $Pr\{[\![\,\boxed{\mathrm{D}}\,]_{\alpha(x)}^t\,]\!]\} = \sum_{e=0}^{\infty} Pr\{E_{\alpha(x),t} = e\}D_x(e)$. According to Post. 1, D_x is an increasing function. This and the fact that $E_{x,t} \succ E_{\alpha(x),t}$ imply that $Pr\{[\![\,\boxed{\mathrm{D}}\,]_x^t\,]\!]\} > Pr\{[\![\,\boxed{\mathrm{D}}\,]_{\alpha(x)}^t\,]\!]\}$. Post. 2 implies that the resistance to a change for a randomly selected agent-feature pair in organization x at time t equals $\hat{\nu}_x$ with probability $Pr\{[\![\,\boxed{\mathrm{D}}\,]_x^t\,]\!]\}$ and equals $\check{\nu}_x$ with probability $1 - Pr\{[\![\,\boxed{\mathrm{D}}\,]_x^t\,]\!]\}$. Similarly, because $\hat{\nu}_{\alpha(x)} = \hat{\nu}_x$ and $\check{\nu}_{\alpha(x)} = \check{\nu}_x$ the resistance to such a change by a randomly selected agent in producer $\alpha(x)$ at time t, equals $\hat{\nu}_x$ with probability $Pr\{[\![\,\boxed{\mathrm{D}}\,]_{\alpha(x)}^t\,]\!]\}$ and equals $\check{\nu}_x$ with probability $1 - Pr\{[\![\,\boxed{\mathrm{D}}\,]_{\alpha(x)}^t\,]\!]\}$. Because $\hat{\nu}_x > \check{\nu}_x$ and $Pr\{[\![\,\boxed{\mathrm{D}}\,]_x^t\,]\!]\} > Pr\{[\![\,\boxed{\mathrm{D}}\,]_{\alpha(x)}^t\,]\!]\}$, it follows that (presumably) $R_{x,t} \succ R_{\alpha(x),t}$. □

We prove the theorem by induction. Let x and $\alpha(x)$ and t^* are such that $\Omega(x)$ and $\overline{\mathrm{CP}}(x, t^*)$ hold. To prove the base step, we first show that $\psi_{\alpha(x),t^*+1} > \psi_{x,t^*+1}$ and $z_{\alpha(x),t^*+1} > z_{x,t^*+1}$. We have

$$(3) \qquad \Pr\left(U_{x,t^*+1} \leq u\right) = \Pr\left(U_{x,t^*} \leq u - 1\right) + \left(1 - \Pr\left(U_{x,t^*} \leq u - 1\right)\right) z_{x,t^*},$$

$$(4) \qquad \Pr\left(U_{\alpha(x),t^*+1} \leq u\right) = \Pr\left(U_{\alpha(x),t^*} \leq u - 1\right) + \left(1 - \Pr\left(U_{\alpha(x),t^*} \leq u - 1\right)\right) z_{\alpha(x),t^*}.$$

By construction, $\alpha(x)$ is the same organization as x until the beginning of period t^*. This implies, in particular, $U_{\alpha(x),t^*} = U_{x,t^*}$ and $\psi_{\alpha(x),t^*} = \psi_{x,t^*}$. We assumed $c_x^* < c_x$. Therefore, $z_{x,t^*} < z_{\alpha(x),t^*}$. This, and equations 3 and 4 imply $\Pr\left(U_{x,t^*+1} \leq u\right) < \Pr\left(U_{\alpha(x),t^*+1} \leq u\right)$. Because this holds for all u, we have $U_{\alpha(x),t^*+1} \prec U_{x,t^*+1}$. We also have $W_{\alpha(x),t^*+1} = W_{x,t^*+1}$, because the turnover of organizational members does not depend on architectural changes (this equally holds for all periods). This, the fact that $U_{\alpha(x),t^*+1} \prec U_{x,t^*+1}$, and Definition 5 imply that $E_{\alpha(x),t^*+1} \prec E_{x,t^*+1}$. Lemma 4 implies $R_{\alpha(x),t^*+1} \succ R_{x,t^*+1}$. Postulate 3 then implies $\psi_{\alpha(x),t^*+1} > \psi_{x,t^*+1}$. This and the assumption that in all periods after t^* both x and $\alpha(x)$ satisfy the constant-change-proportion condition imply that $z_{\alpha(x),t^*+1} > z_{x,t^*+1}$.

The second step is to show that, if $t > t^* + 1$ and the inductive hypothesis holds (i.e., $(z_{x,t-1} < z_{\alpha(x),t-1}) \wedge (\psi_{x,t-1} < \psi_{\alpha(x),t-1})$), then we have $(z_{x,t} < z_{\alpha(x),t}) \wedge (\psi_{x,t} < \psi_{\alpha(x),t})$. This follows easily by adapting the proof of the base step outlined above (we do not provide the details in the interest of space). The principle of induction implies that $(z_{x,t} < z_{\alpha(x),t}) \wedge (\psi_{x,t} < \psi_{\alpha(x),t})$ for all $t \geq t^* + 1$. $\square$

Proof of Proposition 2

Consider an organization x at time t. Suppose that $z_{x,t}^{\mathcal{O}} > z_x^{\mathcal{O}}(t+1)$ and $\mathrm{CP}(x)$ hold. The isomorphic mapping predicate implies $z_{x,t} > z_{x,t+1}$, which under the constant-change-proportion condition implies $\psi_{x,t} > \psi_{x,t+1}$ (see Def. 6). According to Def. 3, $\mathrm{SI}(x)$ holds. $\square$

Appendix C. Empirical Analyses: Robustness Checks

This appendix reports the results of some explorations of the robustness of the empirical findings reported in the body of the paper.

C.1. **Selectivity.** There are two forms of selectivity here. First, among those who do patent, not all firms patent in every year in the observation window. We set the speeds for such years to zero.

TABLE 2. Effects of age on speed of change in the space of patent portfolios with controls for selection: ML estimates of Heckman specifications (columns 1 & 2). Linear regressions with alternative measures of distance between patent (columns 3 & 4).

	(1)	(2)	(3)	(4)
Age	-0.51***	-0.18***	-0.14***	-0.46***
	(0.077)	(0.053)	(0.028)	(0.089)
Age squared/1000	14.6***		3.98***	11.6***
	(3.16)		(1.02)	(3.20)
Cumulative patents/100	-0.0057	0.020	-0.0037	-0.017
	(0.0085)	(0.022)	(0.0035)	(0.011)
Size/1000	0.29	1.42**	0.034	0.23
	(0.24)	(0.53)	(0.078)	(0.24)
y87	8.21***	1.85*	3.27***	12.7***
	(0.41)	(0.92)	(0.16)	(0.52)
y88	3.78***	0.69	1.26***	4.57***
	(0.45)	(0.85)	(0.17)	(0.54)
y89	4.16***	-0.12	1.38***	4.91***
	(0.48)	(0.85)	(0.18)	(0.58)
y90	3.86***	-0.19	1.32***	4.85***
	(0.51)	(0.86)	(0.20)	(0.61)
y91	3.88***	-0.088	1.27***	4.61***
	(0.56)	(0.90)	(0.20)	(0.64)
y92	4.07***	-0.19	1.36***	5.04***
	(0.57)	(0.89)	(0.21)	(0.66)
y93	3.74***	0.071	1.27***	4.58***
	(0.61)	(0.93)	(0.22)	(0.69)
y94	3.93***	-0.058	1.29***	4.56***
	(0.65)	(0.96)	(0.23)	(0.72)
Selection Equation				
Age	-0.090	0.025		
	(0.051)	(0.025)		
Age squared/1000	1.17	-2.26		
	(1.78)	(1.16)		
Cumulative patents/100	0.56*	2.92***		
	(0.25)	(0.28)		
Size/1000	3.08***	1.81***		
	(0.72)	(0.26)		
Year (1987=0)	-0.13***	-0.0087		
	(0.039)	(0.018)		
Constant	1.40***	-0.80***		
	(0.27)	(0.11)		
Mills ratio (λ)	1.25*	4.52***		
	(0.63)	(0.59)		
Observations	440	1595	440	440
Number of firms	61	237	61	61
F(12,428)			57.4	94.6
Wald chi square	679.7	47.3		

Standard errors are in parentheses; *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Second, some firms do not patent at all. In a real sense these firms are not moving at all in this offer space. We deal with each case in turn.

We start with the set of firms with some observed patents, the ones used in the analyses just discussed. We estimate standard Heckman (1979) two-step selection specification but without firm fixed effects (because the two-step estimator is inconsistent for the combination of fixed effects and the correction for sample selection). We set the probit model for the probability of inclusion to depend on all covariates in the main model. We also include a time trend (set equal to zero in 1987, the beginning of the observation window) because firm-years closer to the end of the observation window should be less likely to be followed by an observed patent so that a speed can be calculated.

The estimates are reported in column 1 of Table 2. The effect of age on the speed of change is again significantly negative, and the coefficient estimate is of a similar magnitude to what we obtained with the models without Heckman correction for selection. It turns out that the correction makes a difference. Note in column 1 that the estimate of λ (the hazard of having a patent based on an application in the focal year) is significant. This means that we can reject the null hypothesis of lack of correlation between the disturbances of the selection equation and of the equation for speed. The probability of being selected—that is, of having an observed measure of speed—declines with age. So, adjusting for speed of patents granted, the firms in this sample have longer spells between patents. Unsurprisingly the probability of inclusion drops sharply with the proximity of the end of the observation window. Larger firms and those with larger portfolios of patents are significantly more likely to be included.

Column 2 reports estimates of a Heckman model using the full original dataset with the 237 firms, including those that never patented during the observation period and those that were not granted patents they applied for before the observation period. With this sample, we also find a negative main effect of age on speed (in an additional estimation not reported here, we also included a squared age term, and both coefficients were negative, although not statistically significant).

C.2. **Alternative Measure of Distance.** We also wondered about the impact of the decisions we made in definition on distance measure between patents. Column 3 in Table 2 reports estimates of a model similar to the model used for column 2 in Table 1, but using the classes of the cited patents (instead of the primary subclasses) to compute the distance between patents. In this case, each patent is represented by a 353-long vector of cited classes. The results are similar to what we obtained in the base case: the effect of age on speed remains negative. In column 4, we report

estimates that use a patent-to-patent distance measure which includes the subclass of the focal patents as a feature value with a large weight in addition to the cited subclasses. The weight is computed as follows. When computing the distance of a new patent p_a from another patent p'_a in the current patent portfolio, the weight of each cited subclass is 1 and the weight of the difference between the subclasses of p_a and $p_{a'}$ is equal to the number of distinct subclasses cited by patent p_a. The results are similar to the base case: the effect of age on speed remains negative.

References

Amburgey, Terry L., Dawn Kelly, William P. Barnett. 1993. Resetting the clock: The dynamics of organizational change and failure. *Administrative Science Quarterly* **38**(1) 51–73.

Aronson, Elliot. 2007. *The Social Animal*. 10th ed. Worth Publishers, New York.

Barnett, William P., Glenn R. Carroll. 1995. Modeling internal organizational change. *Annual Review of Sociology* **21** 217–236.

Barnett, William P., Morten T. Hansen. 1996. The red queen in organizational evolution. *Strategic Management Journal* **17**(S1) 139–157.

Barnett, William P., Elizabeth G. Pontikes. 2008. The red queen, success bias, and organizational inertia. *Management Science* **54**(7) 1237–1251.

Bourdieu, Pierre. 1984. *Distinction: A Social Critique of the Judgement of Taste*. Harvard University Press, Cambridge, Mass.

Brewer, Marilyn. 1991. The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin* **17** 475–482.

Burago, Dimitri, Yuri Burago, Sergei Ivanov. 2001. *A Course in Metric Geometry*, *Graduate Studies in Mathematics*, vol. 33. American Mathematical Society, Providence, R.I.

Carroll, Glenn R., Mi Feng, Gaël Le Mens, David G. McKendrick. 2010. Organizational evolution with fuzzy technological formats: Tape drive producers in the world market, 1951–1998. *Research in the Sociology of Organizations* **31** 203–234.

Carroll, Glenn R., Michael T. Hannan. 2000. *The Demography of Corporations and Industries*. Princeton University Press, Princeton.

Christensen, C M, J L Bower. 1996. Customer power, strategic investment, and the failure of leading firms. *Strategic Management Journal* **17**(3) 197–218.

Cyert, Richard M, James G March. 1963. A behavioral theory of the firm. *Englewood Cliffs, NJ* **2**.

Dawes, R M. 1979. The robust beauty of improper linear models in decision making. *American Psychologist* **34**(7) 571–582.

Dubuisson, Marie-Pierre, Anil K. Jain. 1994. A modified hausdorff distance for object matching. *Proc. of the 12th Int. Conf. on Pattern Recognition*, vol. 1. IEEE Comput. Soc. Press, 566–568.

Festinger, Leon. 1957. *A Theory of Cognitive Dissonance*. Stanford University Press, Stanford.

Gilbert, C G. 2005. Unbundling the structure of inertia: Resource versus routine rigidity. *Academy of Management Journal* **48**(5) 741–763.

Gilbert, R J, DMG Newbery. 1982. Preemptive patenting and the persistence of monopoly. *The American Economic Review* .

Hall, Bronwyn H, Adam B Jaffe, Manuel Trajtenberg. 2001. The NBER patent citation data file: Lessons, insights and methodological tools. Tech. rep., National Bureau of Economic Research.

Hannan, Michael T., John Freeman. 1984. Structural inertia and organizational change. *American Sociological Review* **49**(2) 149–165.

Hannan, Michael T., László Pólos, Glenn R. Carroll. 2007. *Logics of Organization Theory: Audiences, Codes, and Ecologies*. Princeton University Press, Princeton, N.J.

Heckman, James J. 1979. Sample selection bias as specification error. *Econometrica* **47**(1) 153–161.

Kahneman, Daniel. 2011. *Thinking Fast and Slow*. Farr, Strauss, and Giroux, New York.

Kovács, Balázs, Michael T. Hannan. 2013. Category spanning, distance, and appeal. Research Report 2081R, Stanford Graduate School of Business.

Le Mens, Gaël, Michael T. Hannan, László Pólos. 2011. Founding conditions, learning, and organizational life chances: Age dependence revisited. *Administrative Science Quarterly* **56**(1) 95–126.

Le Mens, Gaël, Michael T. Hannan, László Pólos. 2014. Organizational obsolescence, drifting tastes, and age-dependence in organizational life chances. *Organization Science* **in press**.

Levinthal, Daniel A. 1997. Adaptation on rugged lansdcapes. *Management Science* **43**(7) 934–950.

Levinthal, Daniel A., Massiomo Warglien. 1999. Landscape design: Designing for local action in complex worlds. *Organization Science* **3** 342–357.

Levitt, Barbara, James G. March. 1988. Organizational learning. *Annual Review of Sociology* **14** 319–340.

March, James G. 1991. Exploration and exploitation in organizational learning. *Organization Science* **2**(1) 71–87.

March, James G., Martin Schulz, Xueguang Zhou. 2002. *The Dynamics of Rules: Change in Written Organizational Codes*. Stanford University Press, Stanford.

Naslund, L, F Pemer. 2012. The appropriated language: Dominant stories as a source of organizational inertia. *Human Relations* **65**(1) 89–110.

Nelson, Richard R, Sidney G. Winter. 1982. *An Evolutionary Theory of Economic Change*. Harvard University Press.

Péli, Gabor, László Pólos, Michael T. Hannan. 2000. Back to inertia: Theoretical implications of alternative styles of logical formalization. *Sociological Theory* **18**(2) 193–213.

Pólos, László, Michael T. Hannan. 2004. A logic for theories in flux: A model-theoretic approach. *Logique et Analyse* **47** 85–121.

Polós, László, Michael T. Hannan, Greta Hsu. 2010. Modalities in sociological arguments. *Journal of Mathematical Sociology* **34**(3) 201–238. doi:DOI10.1080/0022250X.2010.485708.

Reinganum, J F. 1983. Uncertain innovation and the persistence of monopoly. *The American Economic Review* .

Samuelson, William, Richard Zeckhauser. 1988. Status quo bias in decision making. *Journal of risk and uncertainty* **1**(1) 7–59.

Schumpeter, Joseph A. 1934. *The Theory of Economic Development*. Harvard University Press, Cambridge.

Selznick, Philip. 1957. *Leadership in Administration: A Sociological Interpretation*. University of California Press, Berkeley and Los Angeles.

Sørensen, Jesper B., Toby E. Stuart. 2000. Aging, obsolescence, and organizational innovation. *Administrative Science Quarterly* **45**(1) 81–112.

Tversky, Amos, Daniel Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* **211** 453–458.

van der Steen, M. 2009. Inertia and management accounting change: The role of ambiguity and contradiction between formal rules and routines. *Accounting, Auditing and Accountability Journal* **22**(5) 736–761.

Veltman, Frank. 1996. Defaults in update semantics. *Journal of Philosophical Logic* **25** 221–261.

Zajonc, Robert B. 1968. Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology* **92** 1–27.