**Online appendix to accompany:**

# Interdependence, complementarity, and ruggedness of performance landscapes

**Hazhir Rahmandad**
**MIT Sloan School of Management**
**Hazhir@mit.edu**

## Contents

## Appendix A- Why local maxima in PN landscape are corner solutions

Here we provide a sketch of proof for the proposition that all local peaks in the PN Landscape defined on a hypercube in $\mathbb{R}^N$are corner points.

The PN function is continuous with respect to choice vector ***s*** and is differentiable to any order. Therefore the necessary and sufficient conditions for interior peaks, i.e. those not constrained to the bounds of s, are the regular optimality conditions for continuous differentiable functions. $s*$ is a local maxima if and only if:

1) Zero gradient: $\frac{\partial \pi_{PN}(s^*)}{\partial s_i^*} = 0\ for\ 1 \leq i \leq N$
2) Negative definite Hessian matrix

All terms in the PN function are of orders 0 or 1 with respect to any single choice. Therefore the Hessian matrix for PN function is a symmetric matrix with a zero diagonal. It is a basic property of square matrices that the sum of their eigenvalues equals the trace of the matrix, which is zero in the case of PN Hessian matrix. Bringing these two observations together, we conclude that the Hessian for PN function cannot be Negative definite, and thus (barring singular cases such as zero Hessian) includes no internal peak.

The above argument rules out interior peaks, requiring that at least one of the choices be at its lower or upper bound, but it does not require all choices to be at their bounds (i.e. a corner solution). We next extend the argument to show that in any local peak all choices need to be on their bounds.

Starting with a PN function of order M, the argument above showed that at least one choice, call it $s_j$, should be at its boundary value in any local maxima. Fixing the value of $s_j$ at its boundary level, we would have a new function with its domain in $\mathbb{R}^{M-1}$ over which we search for local maxima. However, note that this new function is also a PN function: we simply absorbed $s_j$ in the vector ***c***. Therefore the resulting function is subject to the same argument, requiring yet another choice to be at its boundary, and leading us to only corner maxima by induction.

## Appendix B- Building NK landscapes using PN formulation

A given NK landscape can be replicated using a PN one with an appropriate coefficient vector, $\boldsymbol{c}$. To appreciate the correspondence between the two, consider an NK landscape with maximum interaction ($K=N-1$). This is the most flexible landscape for a given N. The contributions of each choice to performance is re-drawn for every one of the $2^N$ points on the strategy space, and averages of these independently drawn contribution factors become an independent random number representing the performance of each point. In this setting the PN landscape that corresponds to a given NK realization can be built by solving the system of equations for $c_i$ values that replicate $\pi_{NK}$ for the $2^N$ binary combinations of $\boldsymbol{s}$ (i.e. the full strategy space). This system of equations is linear in $c_i$ because every equation of the form $\pi_{PN}(\boldsymbol{c}|\boldsymbol{s})= \pi_{NK}(\boldsymbol{s})$ is linear with respect to $c_i$ values (see eq. 2). Therefore the system includes $2^N$ linear equations and $2^N$ unknowns, and (except for singular cases) has a unique solution. With the same method, one can find a unique PN landscape replicating any realization of NK landscapes for large values of K, where the underlying system of linear equations remain linearly independent. The linear independence condition holds as long as the heterogeneity in the underlying contribution factors of each choice ($C_i$ values in eq. 1) provides $2^N$ degrees of freedom to match. In these cases the simplest model for replicating an NK landscape *should* have $2^N$ free parameters, as is the case with PN, and the estimated PN landscape is unique.

For smaller K values, where the system of equations is not linearly independent, the following solution algorithm provides a unique solution with the fewest non-zero $c_i$'s required to replicate a given NK landscape. First, sort the system of equations based on the number of $c_i$'s in each equation, thus the first equation would be for the point with $s_i$=0 with a single unknown coefficient ($c_0= \pi_{NK}(\boldsymbol{s=0})$) and the last would include all $2^N$ coefficients. Then solve the successive equations in this order, plugging in the $c_i$ values found in the previous solutions as we move down the list. This process allocates the largest contributions to the least complex interactions and keeps the resulting PN landscape simple by zeroing out any interactions not necessary for replicating a given NK landscape. The process also works for solving the full system of linearly independent equations with higher values of K; and for large N is computationally more efficient than solving the system of equations using matrix inversion. Computational experiments confirm that the process above generates simple PN replications of any NK landscape. The resulting PN landscape not only is equivalent to the starting NK, but also is applicable in continuous strategy spaces and allows for a more direct exploration of the research questions at hand.

The actual code for building NK landscapes using PN is available in Appendix I.

## Appendix C- Comparing PN and NK landscapes

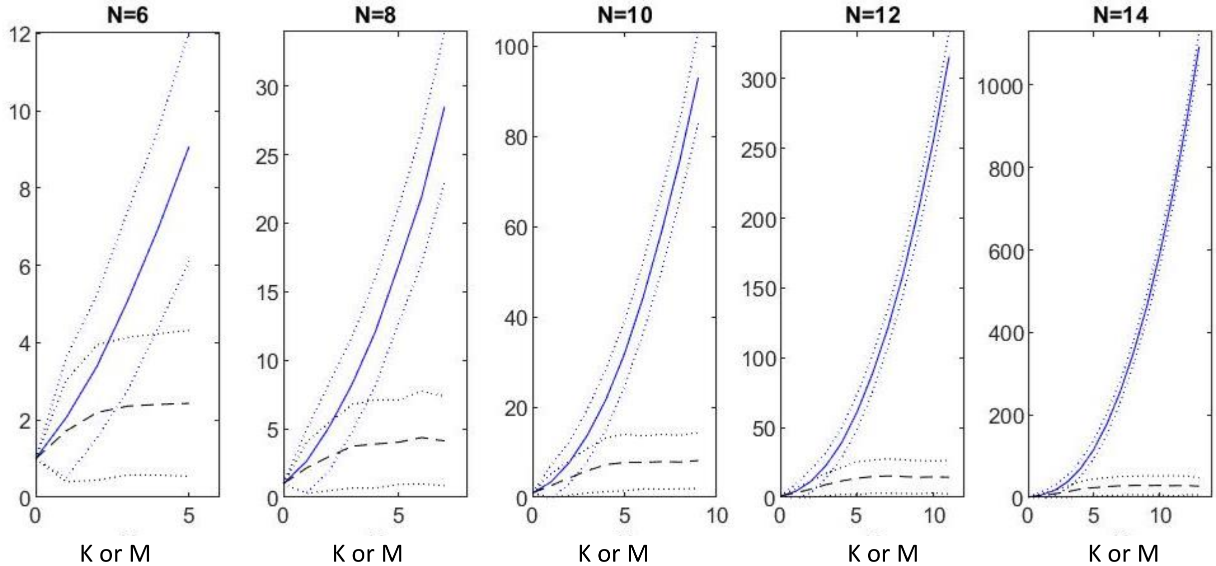


*Figure S 1-Average and 90% outcome range for the number of peaks with different N and K=M values under NK landscape (Solid blue) and PN generalization (dashed black).*

Figure S 1 reports the number of peaks grows with N (values between 6 and 14 are simulated) and M under PN formulation (dashed lines). Results are compared with number of peaks in NK (solid lines) landscapes with K=M. The graphs also include 90% confidence bands for each N and K (=M) combination. Each point on the graph is coming from a sample of 500 simulations. As already shown in the prior literature (Rivkin 2000, Rivkin and Siggelkow 2007), the number of local peaks in the NK landscape grows very rapidly with N and K, reaching an average of 1092 with N=14 and K=13. These results, however, are more subtle in the case of PN formulation: 1) For a given N, the number of peaks in the NK landscape is far larger than in the PN alternative and the gap between the two grows exponentially with N; for example average number of peaks in PN is only 27 with N=14 and M=K=13. 2) The growth in the number of peaks with M is much slower in the PN formulation than is the growth in peaks of NK landscapes with K. Of course one should be careful not to read too much into the direct comparison of M and K values because, while each represents a measure of frequency of interdependences, they are different concepts. Specifically, M values in PN specify the depth of interactions (i.e. up to how many choices may interact with each other), whereas in the NK definition K specifies the number of other choices on which the contribution of each choice depends. In the PN formulation interactions' impact on ruggedness saturates quickly and very deep interactions have little impact on the number of peaks. Therefore, even when all possible interactions are included, the gap between NK and PN formulations grows with N. Of course one can fine-tune the contributions of those interactions (ci's) so that they exactly cancel each other out making neighboring configurations completely independent (that is why PN can replicate any NK outcome). However, such combinations of coefficients are unlikely to be observed by chance in the baseline PN formulation. In replicating the NK landscape using PN formulation we sample heavily from a tiny region of parameter space not otherwise visited in typical PN realizations.

Previous results focused on the canonical NK formulation with random interactions, and the comparative results are qualitatively robust to considering other structured interaction patterns in the NK model. Table S 1 compares the number of peaks in PN formulation against various patterns of interaction previously studied for NK formulations (replicated from Table 3 in (Rivkin and Siggelkow 2007); available for the case of N=12 and 1≤K=M≤6). The highlighted columns report data from the current analysis. Replication of random NK closely matches prior results. While the gap between Dependent and Centralized interaction patterns with NK assumptions is large (others are in a closer range), the number of peaks in PN remains fewer than all others (except for low K values with Centralized NK), and grows much more slowly with M=K; in fact the gap would widen if larger K and N values were examined. In short, the impact of PN formulation on the number of local peaks exceeds the variance due to diverse interaction patterns within NK framework. Future analyses could consider different structures of interaction in PN formulation as well, though keeping it parallel to NK interaction networks may not be easy. For one thing conceptualization of interactions is not the same between PN and NK formulations; moreover, smooth and continuous functions exclude asymmetric interactions (because $\frac{\partial^2 \pi_{PN}}{\partial s_i s_j} = \frac{\partial^2 \pi_{PN}}{\partial s_j s_i}$).

*Table S 1- Comparison of number of peaks between PN and NK (results from (Rivkin and Siggelkow,2007)) with different interdependency patterns (N=12; 1≤K≤6)*

| | NK | | | | | | | | | PN |
|---|---|---|---|---|---|---|---|---|---|---|
| **K** | **Dependent** | **Diagonal** | **Block-diagonal** | **Local** | **Hierarchical** | **Random** | **Random (Replication)** | **Preferential attachment** | **Scale-free** | **Centralized** | |
| **1** | 55.7 | 7.6 | 5.6 | 5 | 2.3 | 4.8 | 4.3 | 5 | 4.5 | 1.9 | 3.3 |
| **2** | 133.3 | 21.6 | 16.5 | 14.2 | 4.9 | 11.7 | 11.6 | 12.1 | 10.3 | 3.4 | 6 |
| **3** | 177.1 | 43.6 | 33.6 | 27.5 | 9.7 | 23.1 | 23.2 | 23.4 | 20.5 | 6.2 | 9.2 |
| **4** | 209.5 | 74.2 | 40.2 | 48.5 | 24.3 | 39.3 | 39.1 | 36.9 | 35.1 | 11.1 | 11.8 |
| **5** | 238.2 | 114.4 | 82.9 | 71.7 | 64.3 | 60.8 | 61.1 | 57.7 | 56.3 | 18.8 | 13.8 |
| **6** | 248.7 | | 100.3 | 98.2 | | 88.7 | 88.5 | 80.7 | | 32.6 | 14.9 |

## Appendix D- Local search and basins of attraction in continuous strategy spaces

Expanding strategy spaces to include continuous choices has a few implications for search and its outcomes. These include implications for both the nature of search and the basins of attraction realized as a result of local adaptation. Whereas search on strategy spaces with binary choices is limited to random flipping of choices to find better alternatives, on a continuous strategy space the topology of local neighborhood, such as the gradient and curvature, can inform the directionality of search. In organizational settings both cognitive search (Gavetti and Porac 2018) and small daily variations experienced by diverse actors enable the formation of reasonable mental maps of local topology, even if actors do not have a global understanding of the landscape. That will allow them to focus on the more promising directions of local adaptation (though randomness and delays complicate this picture in practice (Levinthal and March 1993, March and Olsen 1975)), akin to gradient-descent optimization methods (Koziel and Yang 2011). On the other hand search on continuous spaces also requires the specification of step sizes, which is not trivial: small steps make search slow and inefficient, large steps can overshoot and be disruptive.

The introduction of PN landscape may also impact the size of basins of attractions, an important consideration in analyzing local adaptation. A landscape may have many local peaks but if one peak has a basin of attraction significantly larger than all other local peaks, then a large fraction of organizations are likely to start their exploration from within that peak's basin of attraction. Therefore local search will lead most firms to a single peak and reduce the heterogeneity in configurations; outcomes that are qualitatively similar to those on a unimodal landscape. On the other hand when local peaks have basins of attraction of comparable sizes (as in the NK formulation) the proliferation of peaks significantly increases heterogeneity. Using PN formulation I examine the sensitivity of size of basins of attraction to two mechanisms and compare those with the NK results as a well-known benchmark. First, given a discrete strategy space with binary choices, the size of basins of attraction for peaks in the PN formulation may follow a different distribution than NK. The lower the variance of that distribution, the higher is the relevance of local peaks to heterogeneity. Second, PN allows us to consider continuous strategy spaces, which may also change the size of basins of attraction. Discrete characterizations of basins of attraction focused on the number of discrete configurations from which local search takes the organization to a given peak. On a continuous strategy space the basins of attraction are specified as closed subsets of $\mathbb{R}N$, and their volume represents their size. The ability to generate PN landscapes that replicate any given NK one enables extending this comparison to NK-like landscapes with continuous strategy spaces. Given the same PN landscape (a standard one or one replicating an NK landscape) it is not clear whether the discrete and continuous basins of attraction would have similar size distributions. Therefore we don't know whether considering continuous choices in local adaptation increases, or decreases, the heterogeneity of local peaks discovered (compared to discrete strategy spaces). These two mechanisms are explored next.

For this analysis I track the local adaptation outcomes of a population of firms under four different experimental conditions for different values of 0≤M=K≤9 when N=10. The experimental conditions change the landscape definition (PN vs. NK) and continuity of choices (binary vs. continuous). For each experimental condition, and M or K, the experiment is replicated for 500 different landscapes (and populations of 500 firms each) to provide statistically reliable results. Note that in this analysis we compare landscapes with the same number of local peaks, so the differences in the definition of M and K are not relevant to the analysis. In each replication the 500 organizations are placed randomly on the strategy space, search locally (See below for technical details on the search process used), and find various local peaks with 'market shares' (hi for peak i) that in expectation are proportional to the size of basins of attraction for those peaks. The concentration of hi, formalized using Herfindahl index (Hirschman 1964), provides a simple measure to summarize the results (Figure S 2). For each frequency, F (>1), of discovered local peaks using local adaptation, we graph the Normalized Herfindahl index: $H_N = \frac{F\sum_{i=1}^{F} h_i^2 - 1}{F-1}$. This score varies between 0 (basins of equal size) and 1 (a single dominant basin). Comparing these concentration graphs across the four conditions (Figure S 2) reveals two features. First, for a given F, peaks in the PN formulation are more likely to have unequal basins of attraction. That is, for both discrete and continuous strategy spaces PN landscape (dotted line) has higher HN values than NK. The heterogeneity in size of basins of attraction in PN is mainly due to the higher correlation among neighboring points in this formulation, which boosts the size of larger basins (typically belonging to higher peaks). The effect weakens as the number of peaks grows, and for the largest number of peaks observed in PN (~F=101.2≈16) disappears.

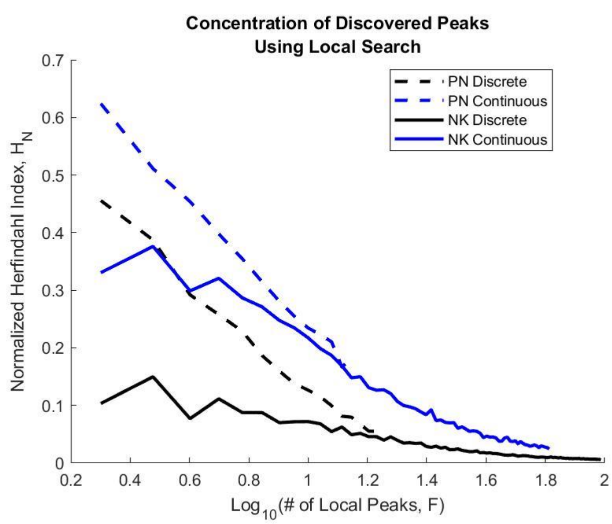


*Figure S 2- Normalized Herfindahl-Hirshman index as a function of number of local peaks for NK (solid) and PN (dashed) landscapes (N=10; 0≤K≤9) and with continuous (blue) vs. discrete (black) strategy spaces. For legibility logarithmic scale is used for number of local peaks (F). Only F values with more than 10 samples are shown.*

Second, compared to discrete spaces, a continuous strategy space increases heterogeneity in the size of basins of attraction. This effect is robust to the choice of PN vs. NK formulation. To understand this effect, consider a two dimensional strategy space (unit square between 0 and 1) with 4 corners and two local peaks at two opposite corners. The discrete search will give each peak a basin with 50% share. In the continuous space basins are two non-overlapping areas that together cover the unit square. Except for singular cases, the size of these basins will be different, giving one a larger share. More generally, whereas local search is limited to changing one dimension at a time in discrete choices, which guarantees each peak a basin of at least N points, continuous space expands the larger basins of attraction by allocating a bigger fraction of intermediate space to some (often taller) peaks and shrinks the size of smaller basins. This effect is pronounced for landscapes with modest ruggedness and attenuates with larger F values. Overall, given a ruggedness level, these two effects may reduce heterogeneity emerging from local adaptation, especially in landscapes with a handful of local peaks. For example, consider landscapes with F=4 peaks. The discrete NK has a normalized Herfindahl index of 0.08 (e.g. an index resulting from market shares h=[0.43, 0.28, 0.18, 0.11]) whereas continuous PN has HN=0.47 (e.g. resulting from h=[0.75, 0.19, 0.05, 0.01]); in the latter example the large majority of firms converge to a single peak (with its basin occupying 75% of strategy space) and very few (6%) land on the two smaller one. In contrast those peaks with smaller basins will have a notable 29% share in the discrete NK case.

**Measuring basins of attraction-** To measure the size of basins of attraction empirically I track the peaks to which a population of 'firms' converges to upon local adaptation. Specifically, for each replication final peaks are tracked for a population of 500 local searching firms. Local adaptation is straight forward in discrete space: a firm randomly flips one choice, tests the resulting payoff, and moves to the new position if that payoff is better than previous point. In the continuous space defining local adaptation is somewhat more complex. First, gradient and curvature of local neighborhood can (and for efficient convergence, should) inform the choice of next step. A purely random step direction would be too slow to converge when unlimited step directions are available for local adaptation. Responsiveness to gradient and curvature is indeed at the heart of most continuous optimization algorithms. A second challenge is related to the choice of step size, which could vary between 0 and 1: too small a step size and one would not converge in reasonable time, too large a step size and one may cross into a different basin of attraction thus missing potential local peaks. For this analysis I used Matlab™ local search algorithm using interior-point method and tested its ability to find the right number of local peaks with PN landscapes with a known number of local peaks. This process also requires a criteria to identify two peaks as identical if they are closer than a threshold. I code two peaks on the continuous strategy space as identical if they are closer than 0.1 in every one of the N=10 choices, an event that would happen only with probability of $10^{-10}$ if the two points were randomly chosen. In fact this process proves somewhat conservative, in that it occasionally identifies two solutions as distinct local peaks even though they would have been recognized as the same peak if the optimization was allowed to continue searching for much longer and with smaller step sizes. Running this algorithm from 500 random starting points (representing the 500 firms) provides a way to map the size of basins of attraction in continuous strategy spaces. The actual code for this analysis is available in Appendix I.

## Appendix E- Additional analysis on complementarity and ruggedness

**Generating supermodular landscapes**

Generation of regular PN landscapes is straight forward: we just sample $c_i$ values independently from a uniform distribution in the range [-$c_{max}$,$c_{max}$], with $c_{max}$ =20 in the analysis for this section. The resulting landscapes are however unlikely to be supermodular. I thus use the following sampling algorithm to conduct the analysis on supermodular PN landscapes. This algorithm, by sequantially assigning different coefficients and tracking a supermodularity budget along each of $N*(N-1)/2$ interactions, ensures that we do not violate the supermodularity constraint while sampling from a large parameter space. There is no unique solution for this problem and depending on the assumptions made one can get different densities of the sampling function on the feasible parameter space. The algorithm used below offers a good coverage of the feasible parameter space and the results remained robust in exploring a few variants of this algorithm.

1- Sort the $c_i$ parameters based on the number of interactions they represent. Call the set of parameters associated with j-fold interactions $c_{p(j)}$. Note that supermodularity does not depend on $c_{p(0)}$ and $c_{p(1)}$. So those parameters can be drawn from the same uniform distribution used for regular PN landscapes, *Uniform*(-$c_{max}$,$c_{max}$).
2- Note that $c_{p(2)}$ equals the mixed partial derivative when choices other than the two focal ones are at zero values, and thus should be non-negative. We can therefore draw these from the positive range of the uniform distributions generating regular PN, *Uniform*(0,$c_{max}$). The resulting $c_{p(2)}$ values define a budget, for each interaction, which later coefficients can consume without violating the supermodularity constraint.
3- Sequentially go over $c_{p(j)}$ for j>2. Set up the minimum ($c_{min}$) value for each coefficient $c_i$ before it may violate the various budgets it impacts, and draw that coefficient uniformly in the range [$c_{min}$,0]. This range keeps complementarity values small, but without violating supermodularity condition and as such may be a stronger test (results are robust to extending the range to [$c_{min}$,$c_{max}$]). Then update budgets for all affected binary interactions and go to the next coefficient.

The actual code for implementing this algorithm is available in Appendix I.

**Complementarity and ruggedness in different strategy spaces**

Results in the main analysis showed that number of local peaks are significantly smaller when strategy is $1 \le s_i \le 2$. Nevertheless, estimating the number of peaks as a function of complementarity fraction (L) in this strategy space further supports the strong effect of complementarity in reducing the number of local peaks. This regression is reported in Table S 2. Naturally the magnitude of effect is smaller than the base case results, because there is far less room to change the total number of local peaks in this strategy space, but the effect remains very significant statistically (t statistic of -73):

*Table S 2- Impact of complementarity on ruggedness. Linear regressions. *** p<0.001,**p<0.01,*p<0.05.*

| DV: | # of Peaks in PN |
|---|---|
| Constant | 1 (0.048) *** |
| Complementarity (*L*) | -0.02 (2.6e-4) *** |
| Dummies for *N* & *M* Combinations | YES |
| Observations | 25000 |
| Adjusted R-squared | 0.295 |

## Appendix F- Complementarity and costs of misperception

I follow (Siggelkow 2002b) to analyze the costs of misperceptions by introducing perceived coefficients for the landscape, $c' = c + r$, where $r$ is a random misperception vector distributed normally with mean zero and a standard deviation that is drawn from $\sigma_i$ ~Uniform[0,2$c_{max}$] for sample landscape $i$. Parameter σ varies the level of misperception in this analysis. I generated a sample of 900 PN landscapes under N=10 and M=9, introduced misperceptions using the above procedure, and augmented that sample with 100 supermodular PN landscapes with misperceptions (See appendix E for generation of supermodular PN landscapes). In each landscape I find the strategy appearing to be the global optima (based on misperceived coefficients, ***c'***) and record its performance gap (*G*) in comparison with the performance of the true global peak (based on ***c***), a measure of cost of misperceptions. I also record a binary variable informing whether the true global peak and the apparent one are the same (*T*). Using this sample of 1000 landscapes I run regressions predicting the G and T outcomes as a function of the fraction of complementarity (L; discussed in the main analysis) and its interaction with level of misperception (σ). The prediction based on prior theory is that the interaction of complementarity (L) with misperceptions (σ) should increase G (and reduce T). Results, reported in Table S 3 suggest otherwise. Specifically, the interaction effect is in the opposite direction (complementarity reduces the costs of misperceptions), and significant in the full sample. A large part of the reductions in the costs is due to supermodular landscapes, excluding that subsample from the analysis removes the statistical significance of the effect on the performance gap measure, though results for misperceived global peak matching the actual peak remain large even without supermodular landscapes. The code for this analysis is available in Appendix I.

*Table S 3- Impact of complementarity (L) and misperception magnitude (σ) on the performance gap between misperceived and actual peak (G) and identification of true global peak (T; binary outcome). Four regressions include full sample (first two models) and the sample limited to standard PN landscapes (excluding supermodular ones; columns 3 and 4).*

| | 1-G (linear) | 2-T (logistic) | 3-G (linear) | 4-T (logistic) |
|---|---|---|---|---|
| Intercept | -7.96 (13.09) | 0.86 (0.49) | 4.97 (20.22) | 0.63 (0.8) |
| L | 32.6 (23.1) | -0.26 (0.84) | 6.01 (39.4) | 0.46 (1.51) |
| σ | 108.41 (11.25)*** | -4.15 (0.62)*** | 73.24 (17.65)*** | -6.64 (1.16)*** |
| Lσ | -74.83 (19.42)*** | 2.65 (0.93)** | -2.4 (34.09) | 7.05 (2)*** |
| Sample | Full (1000) | Full (1000) | No supermodular (900) | No supermodular (900) |
| Adjusted $R^2$ | 0.287 | | 0.295 | |

All models are statistically significant

## Appendix G- Emergence of interior peaks

To explore the emergence of interior peaks, two alternatives with second order terms are added to the PN formulation. The $PN^2$ formulation adds a second order term without bounding the region where the interior peak may be realized. The $PN^{2b}$ formulation pushes for the interior peak(s) to be in the feasible region of unit hypercube, increasing the likelihood of finding interior peaks.

$$\pi_{PN^2}(s) = c_0 + \sum_{i=1}^{2^N-1} c_i \prod_{s_j \in P_i} s_j + \sum_{i=1}^{N} d_i s_i^2 \qquad \text{(s1)}$$

$$\pi_{PN^{2b}}(s) = c_0 + \sum_{i=1}^{2^N-1} c_i \prod_{s_j \in P_i} s_j + \sum_{i=1}^{N} d_i s_i (1 - s_i) \qquad \text{(s1)}$$

Then using N=10 and K=9, I generate PN landscapes under a few different experimental conditions for each of these two formulations. Specifically, drawing the parameter vectors randomly from the following distributions, ***c~Uniform(-$c_{max}$,$c_{max}$)*** and ***d~ Uniform(-$\alpha\beta c_{max}$,$\alpha c_{max}$)*** with $c_{max}$=10, I change α and β to get different experimental conditions. In each condition I conduct 500 replications for statistical reliability. Changing α between 0 and 16 we can vary the relative impact of second order terms compared to the original PN coefficients. Using values of β=0 and β=1 we can assess the impact of the sign of the interaction coefficients. Table S 4 reports the average total number of peaks (fraction of corner peaks) resulting from these experiments. In identifying local peaks I used the method discussed in appendix D, which being conservative, identifies a few more peaks than a perfect search with no computational restrictions would have found. Nevertheless, these results show that the number of local peaks do not change appreciably with the introduction of second order terms, unless those terms are specified to be very strong (e.g. α=16) and include not only positive terms but also negative coefficients that would diversify peaks by adding internal valleys.

*Table S 4- Number of local peaks and fraction of those that are interior (in parantheses) as a function of α (weight of second order effects compared to **c** vector), sign of second order terms (all positive for β=0 and both negative and positive for β=1), and two alternative formulations of second order terms.*

| **# PEAKS (FRACTION INTERIOR)** | ***$PN^2$*** | | ***$PN^{2B}$*** | |
|---|---|---|---|---|
| **A** **B** | 0 | 1 | 0 | 1 |
| **0** | 18.5 (1) | 18.4 (1) | 17.6 (1) | 17.9 (0.99) |
| **0.25** | 19.2 (1) | 19.3 (0.95) | 18.3 (0.91) | 17.4 (0.95) |
| **1** | 19.2 (1) | 18.2 (0.8) | 15.4 (0.7) | 19.5 (0.79) |
| **4** | 17.3 (1) | 17 (0.37) | 12.1 (0.33) | 21.6 (0.35) |
| **16** | 18.9 (1) | 15.9 (0.2) | 7.5 (0.17) | 54 (0.03) |

The actual code for this analysis is available in Appendix I.

## Appendix H- Landscape evolution under innovative shocks of different size

For this analysis impact of innovative shocks of different sizes are assessed in both PN and NK landscapes with N=10 and M=K=9. In PN setup shock size is operationalized by changing the number of coefficients in the parameter vector, ***c***, that are simultaneously redrawn before the innovation is assessed for performance enhancement. Environmental shocks are excluded (so *p*=1) due to their limited impact. Size of shocks is varied with a parameter *z*, with # of parameters redrawn=$2^{zN}$. So for z=0 a single element is changed, and with z=1 all parameters change. In the NK specification the contribution of each choice to payoff can take one of $2^N$ random values depending on that choice and the other N-1 choices. Redrawing a random subset of $2^{zN}$ of those choices will allow us to introduce potential innovations into the NK landscape in a manner consistent with the PN counterpart. Figure S 3 reports the number of peaks over time for different magnitudes of innovative shocks (z values), averaged over 500 replications. In both cases number of local peaks go down as a result of innovations, though the dynamics are faster and more pronounced with PN formulation. Also in both cases the intermediate shock sizes (z=0.8, i.e. 256/1024 parameters in PN; and z=0.4, i.e. 16/1024 contribution factors in NK) show the fastest decline rate in the number of local peaks. The actual code for this analysis is available in Appendix I.

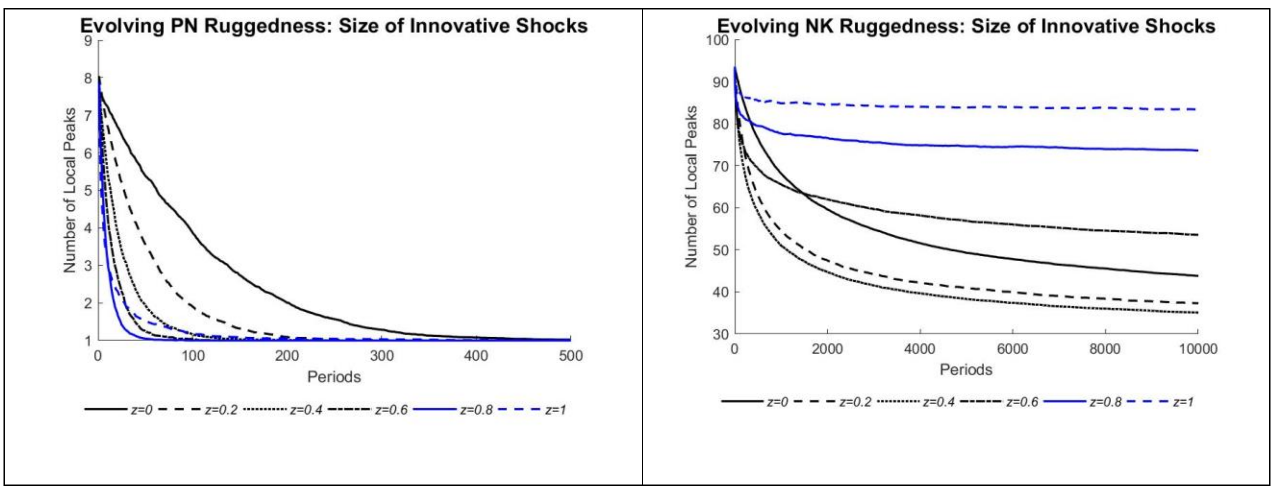


*Figure S 3- Number of local peaks over time as a function of innovative shocks of different sizes (fraction of landscape that changes=*$2^{N(z-1)}$), under A) PN and B) random NK landscapes. For each z value averages over 500 replications reported.

## Appendix I- Replication instructions

The analysis reported in this paper was conducting in Matlab software. For transparency and enhancing replicability, the code used for the analysis is provided in the following link:

https://www.dropbox.com/s/5sfenraltw0hh9m/AnalysisCodeDocumentation.zip?dl=0

To conduct various analysis you can open the main script, PolyNK_V12_Documented.m and choose the subset of analysis you want to conduct (the switches at the beginning of the code specify those subsets). The other functions included in the above archive are used within different segments of the main script or referenced by other functions in the list. Please note that conducting the full analysis, without parallelization, takes over a week on a typical desktop computer, so you may first reduce the replication numbers and separately replicate different subsets of analysis to make sure everything is working correctly in your setup. Also the full analysis may require some Matlab toolboxes (such as optimization toolbox).