Supplemental Information for

"Predicting evolutionary rescue via evolving plasticity in stochastic environments"

Jaime Ashander, Luis-Miguel Chevin, Marissa L. Baskett

Proceedings of the Royal Society B

http://dx.doi.org/10.1098/rspb.2016.1690

Appendix A. Approximate dynamics of growth rate in terms of maladaptation $x(t)$

Our equation (2) shows the Malthusian growth rate $r(t)$ is the sum of the maximum Malthusian growth rate $r_{\max}$, the variance load L_G, and the lag load L_L with $r_{\max} = \log W_{\max}$, $L_G = \log\left(\sqrt{S(\varepsilon_c(t))\omega^2}\right)$, and $L_L = \frac{S(\varepsilon_c(t))}{2} x(t)^2$,

$$\log \bar{W}(t) = r_{\max} + L_G(t) + L_L(t). \tag{A.1}$$

Below, we derive an approximation to this where the only time dependence is through the squared maladaptation $x(t)^2 = (\bar{z}(t) - \theta(\varepsilon_s(t)))^2$.

Appendix A.1. Strength of selection

The strength of selection, $S(\varepsilon_c(t))$, appears in both the variance load L_G and lag load L_L but depends inversely on the phenotypic variance, which changes due to stochastic variation in the environmental cue, which affects the phenotype. Therefore, we approximate the expectation of the strength of selection using a Taylor series:

$$\begin{aligned}\mathbb{E}[S(\varepsilon_c(t))] &= \left(\mathbb{E}\left[\sigma_z^2(\varepsilon_c(t))\right] + \omega^2\right)^{-1} - \frac{\mathrm{Var}(\sigma_z^2)}{2\mathbb{E}[\sigma_z^2]^3} + O(\frac{\mathrm{Var}(\sigma_z^2)^3}{\mathbb{E}[\sigma_z^2]^4}) \\ &\approx \left(\mathbb{E}\left[\sigma_z^2(\varepsilon_c(t))\right] + \omega^2\right)^{-1}.\end{aligned}$$

Expanding σ_z^2 using equation (1) of the main text and taking expectations, we get the approximate expected strength of selection S_δ, assuming $\delta \ll \sigma^2$, which we denote

$$S_\delta = \mathbb{E}[S(\varepsilon_c(t))] \approx \left(\sigma_a^2 + \sigma_b^2(\delta^2 + \sigma_c^2) + \sigma_e^2 + \omega^2\right)^{-1}. \tag{A.2}$$

Where we use the $\mathbb{E}[\varepsilon_c(t)^2] = \mathrm{Var}[\varepsilon_c(t)] + \mathbb{E}[\varepsilon_c]^2$ and the mean environmental shift δ.

Appendix A.2. Variance load

The first term of equation (A.1) is the phenotypic variance load at time t

$$L_G(t) = 1/2 \log S(\varepsilon_c(t)) + 1/2 \log \omega^2.$$

Taking expectations, and using Taylor series:

$$\begin{aligned}
\mathbb{E}[L_G] &= 1/2 \log \omega^2 + 1/2 \mathbb{E}\left[\log S(\varepsilon_c(t))\right] \\
&\approx 1/2 \log \omega^2 + 1/2 \log\left(\mathbb{E}\left[S(\varepsilon_c(t))\right]\right) - \frac{\mathrm{Var}(\sigma_z^2)}{2\mathbb{E}[\sigma_z^2]^2} \\
&\approx 1/2 \log \omega^2 + 1/2 \log\left(\mathbb{E}\left[S(\varepsilon_c(t))\right]\right) \\
&\approx 1/2 \log \omega^2 + 1/2 \log S_\delta
\end{aligned}$$

where the right hand side of the last line is the approximate expected variance load after using $\mathbb{E}\left[S(\varepsilon_c(t))\right] \approx S_\delta$ from (A.2).

Appendix A.3. Lag load

The second term of equation (A.1) is the lag load at time t, $L_L(t)$. Here we show that in expectation, this load has two components, a stochastic load and a shift load. To derive expressions for these, we take the expectation of the entire second term:

$$\begin{aligned}
\mathbb{E}[L_L(t)] =& \mathbb{E}\left[\frac{S(\varepsilon_c(t))}{2} x(t)^2\right], \\
=& \mathbb{E}\left[\frac{S(\varepsilon_c(t))}{2}\right] \mathbb{E}\left[x(t)^2\right] + \mathrm{Cov}\left[\frac{S(\varepsilon_c(t))}{2}, x(t)^2\right],
\end{aligned}$$

where we used the identity $\mathrm{Cov}\,[ab] = \mathbb{E}[ab] - \mathbb{E}[a]\mathbb{E}[b]$. The covariance in the final line represents how stochastic changes in environment (and optimum) that cause maladaptation also cause either larger or smaller phenotypic variance, depending on the direction, due to our assumption that genetic variance in plasticity increases with δ. Furthermore, under weak stabilizing selection, variance in z contributes little to the strength of selection S. For both these reasons, we expect this covariance to be small.

If we neglect the covariance term and use the approximate variance load from Appendix A.2 (S_δ in main text) the expectation is $\mathbb{E}[L_L(t)] \approx \frac{S_\delta}{2}\mathbb{E}\left[x(t)^2\right]$. Removing the expectation, we use this as to approximate the dynamics of the lag load

$$L_L(t) = \frac{S_\delta}{2} x(t)^2.$$

Appendix A.4. Full dynamics

Bringing together the approximations developed above, we have an expression for the growth rate, where all components except the maladaptation are averaged over the fluctuations,

$$r(t) \approx r_{\max} + 1/2 \log(\omega^2 S_\delta) - \frac{S_\delta}{2} x(t)^2. \tag{A.3}$$

From this, we can obtain both the average long-run growth rate, by taking an expectation to obtain $\bar{r}(t) \approx r_{\max} + 1/2 \log(\omega^2 S_\delta) - \frac{S_\delta}{2}(\bar{x}(t)^2 + \sigma_x^2(t))$ (because $\mathbb{E}[x^2] = \bar{x}^2 + \sigma_x^2$). This shows the lag load consists of two loads. The first, expected load from reduction in log mean fitness due to maladaptation of the mean trait relative to the mean optimum in the new environment, is "shift load" $-\frac{S_\delta}{2}\bar{x}(t)^2$. The second, expected load due to reduction in log mean fitness due to random fluctuations in the environment, is "stochastic load" $-\frac{S_\delta}{2}\sigma_x^2(t)$. We can also analyse the variance in trajectories of $r(t)$, and thus population growth because $\log N(t+1) = r(t) + \log N(t)$.

Both of our subsequent analyses require computing the dynamics of mean squared maladaptation $\bar{x}(t)^2$ and the variance in maladaptation $\sigma_x^2(t)$.

Appendix B. Dynamics of shift load depend on mean maladaptation $\bar{x}(t)^2$

In this section, we derive the dynamics of the shift load $-\frac{S_\delta}{2}\bar{x}(t)^2$ under the approximation introduced by Lande [9] that separates adaptation into a fast Phase 1 and a slow Phase 2. We demonstrate that the population is approximately perfectly adapted in mean trait value by the end of Phase 1. We first write down the mean trait dynamics without the approximation, then describe the timescales of the two phase approximation, and derive approximate dynamics of the shift load owing to maladaptation in the mean trait.

Appendix B.1. Mean trait

Trait dynamics follow the standard equation $\Delta\mathbf{y} = \mathbf{G}\boldsymbol{\beta}$, where $\mathbf{y} = (\bar{a}, \bar{b})^T$, $\mathbf{G}$ is the additive genetic variance-covariance matrix, and $\boldsymbol{\beta}$ is the selection gradient. The selection gradient on reaction norm height and slope obtained by taking the log-gradient of $\bar{W}$ [from the equation for fitness above eq. (2) of main text; 9] is

$$\boldsymbol{\beta} = -S(\varepsilon_c(t)) \begin{pmatrix} \bar{a}(t) - A + \bar{b}(t)\varepsilon_c - B\epsilon_s \\ \left(\bar{a}(t) - A + \bar{b}(t)\varepsilon_c - B\epsilon_s\right)\varepsilon_c \end{pmatrix}. \tag{B.1}$$

With a constant additive genetic variance-covariance matrix ($\mathbf{G}$ matrix), the change per generation in $\bar{a}(t)$ and $\bar{b}(t)$ is given by

$$\Delta \begin{pmatrix} \bar{a}(t) \\ \bar{b}(t) \end{pmatrix} = \begin{pmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{pmatrix} \boldsymbol{\beta}.$$

In the new environment, the expectation of the change per generation conditional on $\bar{a}$ and $\bar{b}$ is

$$\begin{aligned} \Delta \begin{pmatrix} \mathbb{E}(\bar{a}(t)) \\ \mathbb{E}(\bar{b}(t)) \end{pmatrix} &= \mathbb{E}[\mathbf{G}\boldsymbol{\beta}] \\ &\approx -S_\delta \mathbf{G} \begin{pmatrix} \bar{a} - A + \bar{b}\delta - B\delta \\ \mathbb{E}[(\bar{a} - A)\varepsilon_c(t)] + \mathbb{E}[\bar{b}\varepsilon_c^2(t)] - B\mathbb{E}[\varepsilon_c \varepsilon_s] \end{pmatrix}, \end{aligned}$$

where the approximation comes from the treating $\mathbb{E}[S(\varepsilon_c(t))]$ as a constant. We assume no environmental tracking by phenotypic plasticity, i.e., $\mathrm{Cov}[\bar{b}_t, \varepsilon_c(t)^2(t)] = 0$ so $\mathbb{E}[\bar{b}_t \varepsilon_c^2] = \mathbb{E}[\bar{b}_t]\mathbb{E}[\varepsilon_c(t)^2(t)]$, or by reaction norm elevation, i.e., $\mathrm{Cov}[\bar{a}_t, \varepsilon_c(t)] = 0$ so $\mathbb{E}[\bar{a}_t, \varepsilon_c(t)] = \mathbb{E}[\bar{a}_t]\mathbb{E}[\varepsilon_c(t)]$. Note that the tracking of the environment by the reaction norm elevation could be included, reducing the expected mean plasticity [28], but we neglect it here. However, we do allow for environmental tracking when computing the lag load (below). Then, using $\mathbb{E}[\varepsilon^2(t)] = \delta^2 + \sigma^2$ and $\mathbb{E}[\varepsilon_c(t)\varepsilon_s] = \delta^2 + \rho\sigma^2$, the expectation of the change is approximately

$$\mathbb{E}\left[\Delta \begin{pmatrix} \bar{a}(t) \\ \bar{b}(t) \end{pmatrix}\right] \approx -S_\delta \mathbf{G} \left[\begin{pmatrix} 1 & \delta \\ \delta & \delta^2 \end{pmatrix} \begin{pmatrix} \bar{a}(t) - A \\ \bar{b}(t) - B \end{pmatrix} + \begin{pmatrix} 0 \\ (\bar{b}(t) - \rho B)\sigma^2 \end{pmatrix} \right]. \tag{B.2}$$

The approximation is exact if $\mathrm{Cov}[\bar{b}_t, \varepsilon_c(t)^2(t)] = \mathrm{Cov}[\bar{a}_t, \varepsilon_c] = 0$. Note that this differs from Lande [9], where this relation was treated as exact [28].

We solve for the explicit trait dynamics relative to the long-run equilibrium state [9]. Setting selection gradient $\boldsymbol{\beta}$ in equation (B.1) to zero, solve for long run trait values

$$\begin{pmatrix} \bar{a}_\infty \\ \bar{b}_\infty \end{pmatrix} = \begin{pmatrix} A + B\delta(1 - \rho_\delta) \\ B\rho_\delta \end{pmatrix} \tag{B.3}$$

One can show, with some algebra [9], the one-generation change (B.2) is the product -$S_\delta \tilde{\mathbf{G}} \mathbf{z}(t)$, where $\mathbf{z}(t)$ is the difference between mean trait values and their long-run equilibrium values computed above and

$$\tilde{\mathbf{G}} = \begin{pmatrix} {\sigma_a}^2 & {\sigma_a}^2\delta \\ {\sigma_b}^2\delta & {\sigma_b}^2\delta^2 \left(1 + \frac{\sigma^2}{\delta^2}\right) \end{pmatrix}.$$

The expected dynamics can then be expressed in terms of the eigenvectors and eigenvalues of the matrix $\tilde{\mathbf{G}}$ [9, 16],

$$\mathbf{z}(t) = c_1\mathbf{e}_1\left(1 - S_\delta\lambda_1\right)^t + c_2\mathbf{e}_2\left(1 - S_\delta\lambda_2\right)^t, \tag{B.4}$$

where $\mathbf{e}_i$ and λ_i are eigenvectors and eigenvalues of $\tilde{\mathbf{G}}$ respectively and c_i terms are constants determined by initial conditions.

Appendix B.2. Approximation for a large environmental shift

As in earlier work, we consider the case where the shift in the mean environment is very large relative to background noise, $\frac{\sigma^2}{\delta^2} \ll 1$. Here, we initially write down the eigenvalues and eigenvectors of $\tilde{\mathbf{G}}$ to first order in this small term, but thereafter follow Chevin and Lande [16] in deriving approximate dynamics to leading order. To first order in $\frac{\sigma^2}{\delta^2}$, the eigenvalues are

$$\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix} = \begin{pmatrix}(\sigma_a^2 + \delta^2\sigma_b^2) + \delta^2\sigma_b^2\frac{\sigma^2}{\delta^2}\phi\\ \sigma_a^2\frac{\sigma^2}{\delta^2}\phi\end{pmatrix},$$

and the eigenvectors are

$$\mathbf{e}_1 = \begin{pmatrix}\delta\left(1-\phi\right)\left(1-\frac{\sigma^2}{\delta^2}\phi\right)\\ \phi\end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix}\delta\left(1+\frac{\sigma^2}{\delta^2}\phi\right)\\ -1\end{pmatrix}.$$

These match the calculations of [9] up to a constant of S_δ (equivalent to γ in Lande [9]).

Assuming the population has long evolved in an environment with predictability ρ, the initial trait values are $(\bar{a}_0, \bar{b}_0) = (A, \rho B)$. Using (B.3), the initial conditions in the re-centred trait x are

$$\mathbf{z}_0 = \begin{pmatrix}-B\delta(1-\rho_\delta)\\ B(\rho-\rho_\delta)\end{pmatrix}$$

To leading order, the constants are

$$\begin{pmatrix}\mathrm{c1}\\ \mathrm{c2}\end{pmatrix} = \begin{pmatrix}-B(1-\rho)\\ -B(\rho(1-\phi)-\rho_\delta+\phi)\end{pmatrix}$$

Appendix B.2.1. Timescales of phases 1 and 2

If most phenotypic variation in the new environment is due to variance in plasticity, $\phi \approx 1$, and the shift in the mean environment is large (relative to background variability as in our approximation above), the trait change takes place in two phases that occur at very different timescales [9]. When selection is

weak, geometric terms in (B.4) can be replaced by exponential terms $e^{-tS_\delta\lambda_i}$, indicating the relative timescales of change along the eigenvectors e_i are given by $t_i \approx \frac{1}{\lambda_i}$. The ratio $t_1/t_2 \approx \phi(1-\phi)\frac{\sigma^2}{\delta^2}$, and when much of the additive genetic variation is due to variation in plasticity so $\phi \approx 1$, then $t_1/t_2 \approx \frac{\sigma^2}{\delta^2}$, which is small in the approximate case we treat. Then, change along $\mathbf{e}_1$ occurs very fast relative to change along $\mathbf{e}_2$ [9]. At the end of Phase 1, $e^{-tS_\delta\lambda_1} \approx 0$ while $e^{-tS_\delta\lambda_2} \approx 1$ Thus, the approximate state of the system relative to its final state, i.e., (B.3), is $c_2 e_2$. The trait values, at the end of Phase 1 are, to leading order,

$$\mathbb{E}\left[\begin{pmatrix}\bar{a}_{O(t_1)}\\ \bar{b}_{O(t_1)}\end{pmatrix}\right] \approx \begin{pmatrix} A + B\delta(1-\rho)(1-\phi) \\ B(\rho+\phi(1-\rho)) \end{pmatrix} \tag{B.5}$$

The effect of the initial environment occurs through predictability ρ, which under our assumption that the population is adapted initially also determines the initial mean plasticity $\bar{b}_0 = \rho B$. We see the initial plasticity has a strong influence at the end of Phase 1 only if ϕ is small. When ϕ is large, the plasticity at the end of Phase 1 is close to "perfect" i.e. $b_{O(t_1)} \approx B$. Note also that to first order, the mean phenotype is perfectly adapted $\bar{z}_{O(t_1)} \approx A + B\delta$.

Where the extinction risk is calculated at half of the characteristic timescale of Phase 2, i.e., $t_{\text{bef}} = \frac{\phi\delta^2}{2\sigma_a^2\sigma^2}$

Appendix B.2.2. Trait dynamics during Phase 1

Throughout Phase 1, the term $(1 - S_\delta\lambda_2)^t \approx 1$, so the dynamics are given by

$$\mathbf{z}(t) = c_2\mathbf{e}_2 + c_1\mathbf{e}_1\left(1 - S_\delta\lambda_1\right)^t.$$

We again replace the geometric term with an exponential (valid for weak selection) and re-normalize. To leading order, after some rearranging, the right hand side of the expected dynamics is

$$\mathbb{E}\left[\begin{pmatrix}\bar{a}(t)\\ \bar{b}(t)\end{pmatrix}\right] = \left(1 - e^{-tS_\delta\lambda_1}\right) B(1-\rho)\begin{pmatrix}\delta(1-\phi)\\ \phi\end{pmatrix} + \begin{pmatrix}A\\ B\rho\end{pmatrix}, \tag{B.6}$$

which is analogous to the result of Chevin and Lande [Supporting Information, eq A6; 16]. As in that paper, we compute the eigenvalue only to leading order in $\frac{\sigma^2}{\delta^2}$ so $\lambda_1 \approx \sigma_a^2 + \delta^2\sigma_b^2$ which is equivalent to the expression $\frac{\sigma_a^2}{1-\phi}$ used in Chevin and Lande [16].

Appendix B.2.3. Trait dynamics during Phase 2

During Phase 2, the term $(1 - S_\delta\lambda_1)^t \approx 0$, so the dynamics are given by

$$\mathbf{z}(t) = c_2\mathbf{e}_2\left(1 - S_\delta\lambda_2\right)^t.$$

Then, using (B.3) and again replacing the geometric term with an exponential (valid for weak selection) and re-normalizing, the expected dynamics to leading order in $\frac{\sigma^2}{\delta^2}$ during Phase 2 are

$$\mathbb{E}\left[\begin{pmatrix}\bar{a}(t)\\ \bar{b}(t)\end{pmatrix}\right] = \begin{pmatrix}A + B\delta(1-\rho_\delta)\\ B\rho_\delta\end{pmatrix} - B(\rho - \rho_\delta + \phi(1-\rho))\begin{pmatrix}\delta\\ -1\end{pmatrix} e^{-tS_\delta \sigma_a^2 \frac{\sigma^2}{\delta^2}\phi}, \tag{B.7}$$

where we have used $\lambda_2 \approx \sigma_a^2 \frac{\sigma^2}{\delta^2}\phi$. Note that for $t = O(t_1)$, this exponential term equals 1 and this equation agrees with (B.5).

Appendix B.3. Dynamics of expected maladaptation during Phase 1

We derive the expected maladaptation of the mean trait during an initial phase of evolutionary rescue, focusing on a case where the size of the environmental shift is large and much of the additive genetic variance in the new environment owes to genetic variance in reaction norm slope. The shift load (computed in Appendix A) is $\frac{S_\delta}{2}\bar{x}(t)^2$. After we compute the dynamics of the mean maladaptation $\bar{x}(t)^2$, we will have an approximation for dynamics of the shift load,

$$\begin{aligned}\bar{x}(t)^2 = \mathbb{E}[x(t)]^2 &= \Big(\mathbb{E}[\bar{a}(t)] - A + \mathbb{E}[\bar{b}(t)\varepsilon_c(t)] - B\mathbb{E}[\varepsilon_s]\Big)^2\\ &\approx \Big(\mathbb{E}[\bar{a}(t)] - A + \delta(\mathbb{E}[\bar{b}(t)] - B)\Big)^2.\end{aligned}$$

The last equation comes from assuming the covariance between reaction norm slope and the cuing environment is small relative to the mean value of the new environment. Then, $\mathbb{E}[\bar{b}(t)\varepsilon_c(t)] \approx \mathbb{E}[\bar{b}(t)]\mathbb{E}[\varepsilon_c(t)] = \delta\mathbb{E}[\bar{b}(t)]$ in the new environment. This is reasonable when $\frac{\sigma^2}{\delta^2} \ll 1$. After using (B.6) and some algebra, we obtain

$$\bar{x}(t)^2 \approx B^2\delta^2(1-\rho)^2 e^{-2tS_\delta \frac{\sigma_a^2}{1-\phi}}. \tag{B.8}$$

Where we use ϕ to represent the proportion of additive genetic variation in the new environment due to variation in plasticity,

$$\phi = \frac{\delta^2\sigma_b^2}{\sigma_a^2 + \delta^2\sigma_b^2}.$$

Equation (B.8) indicates the shift load goes to zero as t increases.

Appendix C. Stochastic load: variance in maladaptation σ_x^2 at stationarity

Appendix C.1. Perceived environment with fixed plasticity

A tactic from Michel *et al.* [35] aids in calculating the variance as a function of fixed plasticity. We define the perceived optimum $\psi(t)$ as the difference between the optimum and the mean trait after accounting

for the plastic response $\psi(t) = B\varepsilon_s(t) - \bar{b}^*\varepsilon_c(t)$ so that $x(t) = \bar{a}(t) - \psi(t)$. Then, the perceived variance in the optimum is

$$\sigma_\psi^2(\bar{b}^*, \rho_\delta) = \sigma^2(B^2 + \bar{b}^*(\bar{b}^* - 2B\rho_\delta)), \tag{C.1}$$

and autocorrelation in the perceived optimum is

$$T_\psi(\bar{b}^*, \rho_\delta) = -\log\left[\rho_\delta^{1/\tau}\left(1 - \frac{B\bar{b}^*(\rho_\delta^{-1} - \rho_\delta)}{B^2 + \bar{b}^*(\bar{b}^* - 2B\rho_\delta)}\right)\right]^{-1} \tag{C.2}$$

[35]. We can then express maladaptation in terms of the intercept and perceived environment as $x(t) = \bar{a}(t) - \psi(t)$.

Appendix C.2. Variance at stationarity

We derive an approximation for variance in maladaptation under stationarity, which we denote σ_x^2. In practice, this means we solve for the effect of fluctuations on maladaptation after a long time, we assume the mean maladaptation is zero, and we also assume fixed mean plasticity $\bar{b}^*$. We are interested in finding an asymptotic expression for the variance of this term.

Assuming fixed plasticity, all change in the trait occurs through change in the reaction norm height,

$$\Delta\bar{z} = \Delta\bar{a}(t) = -S_\delta\sigma_a^2 x(t).$$

When selection is weak relative to genetic variance in reaction norm height, and the fluctuations in the perceived environment are not large, evolution can be approximated in continuous time [22, 35] as

$$\frac{dx}{dt} + S_\delta\sigma_a^2 x = -\frac{d\psi}{dt},$$

where $x = \bar{a}(t) - \psi$. For $t \gg t_1$ and constant genetic variance σ_a^2, the solution to this differential equation is

$$\bar{a}(t) = S_\delta\sigma_a^2\int_0^\infty \exp\left(-S_\delta\sigma_a^2\tau\right)\psi(t-\tau)\mathrm{d}\tau. \tag{C.3}$$

What remains is to compute $\mathrm{Var}[x(t)]$. Using our quasi-stationarity assumption, we need only compute $\mathbb{E}[x(t)^2] = -2\mathbb{E}[\bar{a}(t)\psi(t)] + \mathbb{E}[\bar{a}(t)^2] + \mathbb{E}[\psi(t)^2]$. The last of these expectations is simply the variance of the perceived environment σ_ψ^2. The first and second expectations integrate over time (from eq. C.3),

$$\begin{aligned}
-2\mathbb{E}[\bar{a}(t)\psi(t)] &= -2S_\delta\sigma_a^2\int_0^\infty \exp\left(-S_\delta\sigma_a^2\tau\right)\mathbb{E}[\psi(t)\psi(t-\tau)]\mathrm{d}\tau \\
\mathbb{E}[\bar{a}(t)^2] &= S_\delta^2\sigma_a^4\int_0^\infty\int_0^\infty \exp\left(-S_\delta\sigma_a^2(\tau_1+\tau_2)\right)\mathbb{E}[\psi(t-\tau_1)\psi(t-\tau_2)]\mathrm{d}\tau_1\mathrm{d}\tau_2.
\end{aligned}$$

Because ψ is a linear combination of autoregressive Gaussian processes $\varepsilon_c(t)$ and ε_s, we can express the expectations involving ψ in terms of autocovariance $\mathbb{E}[\psi(t), \psi(t-\tau)] = \sigma_\psi^2 \exp(-\tau/T_\psi)$ and $\mathbb{E}[\psi(t-\tau_1), \psi(t-\tau_2)] = \sigma_\psi^2 \exp(-|\tau_1 - \tau_2|/T_\psi)$. In both of these expressions, T_ψ is the characteristic autocorrelation time of the perceived environment ψ.

The first expectation is

$$\begin{aligned} -2\mathbb{E}[\bar{a}(t)\psi(t)] &= -2S_\delta \sigma_a^2 \sigma_\psi^2 \int_0^\infty \exp\left(-\tau(S_\delta \sigma_a^2 + 1/T_\psi)\right) \mathrm{d}\tau, \\ &= -2\frac{T_\psi S_\delta \sigma_a^2 \sigma_\psi^2}{(T_\psi S_\delta \sigma_a^2 + 1)}. \end{aligned}$$

The second expectation involves an absolute value term, meaning the integral must be taken in two parts, and evaluates to

$$\begin{aligned} \mathbb{E}[\bar{a}(t)^2] =& S_\delta^2 \sigma_a^4 \int_0^\infty \int_0^\infty \exp\left(-S_\delta \sigma_a^2(\tau_1 + \tau_2)\right) \sigma_\psi^2 \exp(-|\tau_1 - \tau_2|/T_\psi) \mathrm{d}\tau_1 \mathrm{d}\tau_2, \\ =& S_\delta \sigma_a^2 \sigma_\psi^2 \frac{T_\psi}{(T_\psi S_\delta \sigma_a^2 + 1)} \end{aligned}$$

Combining these expressions,

$$\begin{aligned} \mathbb{E}[(\bar{a}(t) - \psi(t))^2] &= \sigma_\psi^2 \left(-2\frac{T_\psi S_\delta \sigma_a^2}{(T_\psi S_\delta \sigma_a^2 + 1)} + S_\delta \sigma_a^2 \frac{T_\psi}{(T_\psi S_\delta \sigma_a^2 + 1)} + 1 \right) \\ &= \frac{\sigma_\psi^2}{(T_\psi S_\delta \sigma_a^2 + 1)} \left(T_\psi S_\delta \sigma_a^2 + 1 - 2T_\psi S_\delta \sigma_a^2 + T_\psi S_\delta \sigma_a^2 \right) \\ \sigma_x^2(\bar{b}^*, \delta, \rho_\delta) &:= \frac{\sigma_\psi^2}{(T_\psi S_\delta \sigma_a^2 + 1)}, \end{aligned}$$

where the last line defines the variance we sought to calculate. As our notation emphasizes, this load depends primarily on the level of plasticity $\bar{b}^*$ and the size of the shift δ. Note that this formula matches Lande and Shannon [22].

Appendix C.3. Quasi-stationary variance over Phase 1

The variance in maladaptation σ_x^2 depends on variance $\sigma_\psi^2(\bar{b}^*, \rho_\delta)$ from eq. (C.1) and characteristic timescale $T_\psi(\bar{b}^*, \rho_\delta)$ from eq. (C.2) of perceived fluctuations in the perceived optimum (given autocorrelation ρ_δ at timescale τ in the true optimum), as

$$\sigma_x^2(\bar{b}^*, \delta, \rho_\delta) = \frac{\sigma_\psi^2(\bar{b}^*, \rho_\delta)}{\left(S_\delta \sigma_a^2 T_\psi(\bar{b}^*, \rho_\delta) + 1\right)}. \tag{C.4}$$

After a long time, on the timescale of Phase 1, t_1, the mean maladaptation is zero, as shown in eq. (B.8), and the variance in maladaptation is stationary. At this point, as shown in Appendix B, plasticity is at its approximate maximum $b_{\max} = B(\rho + \phi(1-\rho))$. It remains as this value for a long time (on the scale of Phase 1; although eventually decays to the predictability ρ_δ according to the dynamics in Appendix B.2.3) and so we evaluate (C.4) at $b_{\max}$ to determine persistence in (3). Using the "quasi-stationary" approximation that fluctuations achieve stationarity while $\bar{b}^* = b_{\max}$, we evaluate $\sigma_x^2(b_{\max}, \delta, \rho_\delta)$.

Note that we could also evaluate (C.4) at any time t during Phase 1,

$$\sigma_x^2(t) = \sigma_x^2(\bar{b}(t), \delta, \rho_\delta), \tag{C.5}$$

by using dynamics for reaction norm slope $\bar{b}(t)$ from eq. (B.6). Doing so assumes for each change in $\bar{b}$ over Phase 1 the stochastic variance in maladaptation achieves stationarity. No doubt this is inaccurate in some cases, but it is a tractable analytical approximation for the variance. Using this approximation over the parameter ranges we examine in this paper, the stochastic load does not change much with changes in plasticity (Figure S5). Accordingly, we use the simpler approximation above, and evaluate the variance at $b_{\max}$ for all time.

Appendix D. Simulation

In simulation, we implement the autocorrelated environment as an autoregressive function with correlation κ on a timescale with $n = \text{ceiling}[\frac{1}{\tau}]$ time units in a generation. The process simulated is

$$x_i = \kappa x_{i-1} + \sqrt{1-\kappa^2}\xi\sigma,$$

where ξ is a unit normal random variable. If x and ξ are independent (as generally assumed in autoregressive functions) and the process is stationary, it has variance σ^2 (and mean 0). Further, the covariance of observations one time unit apart is $\kappa\sigma^2$ and the correlation of such pairs of observations is κ. For observations n time units apart, the correlation becomes κ^n. Thus, using the n time steps per generation, the simulated process relates to the exponential autocovariance function given above as $\kappa^n = \rho^{1/\tau}$. Accordingly we set the correlation within simulations equal to that of the environmental predictability at timescale τ, i.e., $\kappa = \rho$.

Appendix E. Alternative assumptions of reaction norm shape

If, instead of assuming phenotypic variance in the reference environment is minimal, we assume the environment is shifted to an environment where the variance is minimal, then the additive genetic variance of the expressed trait $z(t)$ increases quadratically away from the novel environment $\varepsilon_c(t) = \delta$. Without the assumption that variance is minimized in the reference environment, the additive genetic covariance between a and b is non-zero in the reference environment [9]. From that paper, the full form of (1b) is $\sigma_z^2(\varepsilon_c(t)) = \sigma_a^2 + 2\sigma_{ab}\varepsilon_c(t) + \sigma_b^2\varepsilon_c^2(t) + \sigma_e^2$, where σ_{ab} is the additive genetic covariance between reaction norm slope and variance in the reference environment. (With the assumption of variance minimized at $\varepsilon_c = 0$, $\sigma_{ab} = 0$, and the covariance in any other environment is $\mathrm{Cov}(a, b) = \sigma_b^2\varepsilon_c$.) Then $\sigma_z^2(\varepsilon_c(t))$ is minimized in the environment $\varepsilon_c^* = -\sigma_{ab}/\sigma_b^2$ [9].

Thus, assuming minimal variance in the new environment implies $\sigma_{ab} = -\delta\sigma_b^2$. The mean of the expressed trait value $z(t)$ before selection is the same as in (1), but the variance differs, giving

$$\bar{z}(t) = \bar{a}(t) + \bar{b}(t)\varepsilon_c(t) \tag{E.1a}$$

$$\sigma_z^2(\varepsilon_c(t)) = \sigma_a^2 - 2\sigma_b^2\delta\varepsilon_c(t) + \sigma_b^2\varepsilon_c^2(t) + \sigma_e^2, \tag{E.1b}$$

which assumes the additive genetic variances are constant in time.

In this case, the expressed trait z and the slope b have covariance $\mathrm{Cov}(z, b) = (\varepsilon_c(t) - \delta)\sigma_b^2$ and so, with $\varepsilon_c(t) \approx \delta$ there is approximately zero covariance between the trait and reaction norm slope, and so direct selection on the trait results in very weak selection on reaction norm slope.

Consequently, the transient increase in plasticity should not be expected without the assumed increase of genetic variance in novel environments. Given the uncertainty described by McGuigan and Sgro [12] concerning the effects of stress (i.e., novelty) on additive genetic variance, theory could usefully outline empirical possibilities. The derivation of (E.1) is just the beginning. It does show, however, that the theory presented in the main text is not completely general. No theory is unless it thoroughly considers the possible relationships between additive genetic variance and environmental shifts. In the main text, we analyse only one set—where many are possible—of assumptions on how plasticity, demography, and evolution interact during evolutionary rescue. However, this set of assumptions appears met by empirical reality in at least some cases (see main text, "Assumptions and Caveats").

Appendix F. A small shift and large additive variance in plasticity σ_b^2

We expect our approximation to perform best when a large proportion of additive variance in the novel environment is due to variance plasticity (i.e., $\phi \approx 1$). This occurs in our model either when σ_b^2 (GxE) is large in the reference environment, or when a large shift in the mean environment causes quadratic increases in additive genetic variance (as $\delta^2\sigma_b^2$).

In the former case, the population has inherently high genetic variability in plasticity, and our assumptions imply a relatively large changes in additive genetic variance for even small shifts in the mean environment. In the latter case, increased additive genetic variance is driven by the novelty of the environment.

Simulations reveal that despite both cases increasing ϕ, the approximation does not perform equivalently well. If the mean shift is small then the approximation for the threshold between decline and persistence performs poorly as additive variance in plasticity increases ($\delta = 1.5$, Figure S4). In fact, the approximation appears to perform better for lower values of ϕ.

Appendix G. Stochastic House of Cards Approximation

Figure S3 shows effects of genetic variance changing with population size according to a modified Stochastic House of Cards (SHC) approximation: $\sigma_{SHC}(\sigma_g^2) = \sigma_g^2/(1+\frac{\omega^2+\sigma_e^2}{\mu^2 N_e})$ and $N_e \approx 2R_0N/(2R_0-1)$. Parameters and panels are as in Figure 2 with the additional parameter μ^2, which is the variance of the effect of new mutations, fixed at $\mu^2 = 0.005$. For the un-modified SHC the numerator, σ_g^2, is replaced by a term that includes the per-generation total mutation rate V_m and the strength of selection: $2V_m(\omega^2 + \sigma_e^2)$ [see 23, who used α^2 instead of μ^2]. Thus, the modified version we use effectively assumes that larger initial values of σ_a^2 and σ_b^2 reflect populations with larger mutation rates.

Appendix H. Supplementary Figures

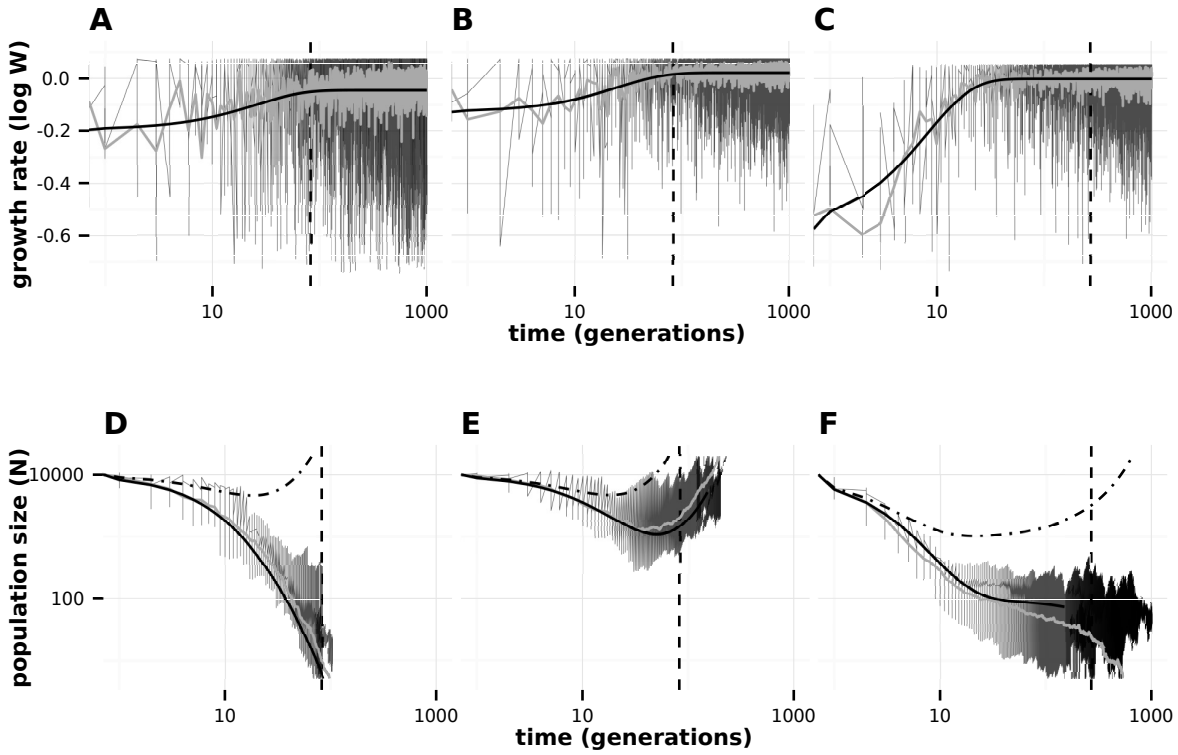

Figure S1: Dynamics of growth rate ($\log \bar{W}$, **A-C**), and population size ($\log N$, **D-F**) versus time (in generations, log scale) under evolutionary rescue for three scenarios of environmental shift δ and predictability ρ_δ following the shift: a modest shift and low predictability (δ =2.5, ρ_δ =0.3: **A,D**), a modest shift and high predictability (δ =2.5, ρ_δ =0.7: **B,E**), and a large shift and high predictability (δ =5, ρ_δ =0.7: **C,F**). Each panel shows 10 replicate simulations of 1500 generations (thin black lines), the mean of these simulations (thick grey line), and predicted trajectories of the mean (solid black line); also shown for comparison are predictions without amplifying effect of plasticity on stochastic fluctuations (dash-dot line). The dashed vertical lines indicate the time during Phase 1 at which we compute quasi-extinction before rescue (t_{bef}; see Figure 1). Parameters: initial predictability $\rho = 0.5$, additive genetic variance in plasticity $\sigma_b^2 = 0.05$; other parameters as in Figure 2. Greater shift size implies larger increase of additive genetic variance in the new environment; for a modest shift (**A,B, D,E**) our model assumes additive genetic variance increases by a factor of 4, for a large shift, (**C,F**) the increase is by a factor of 14.

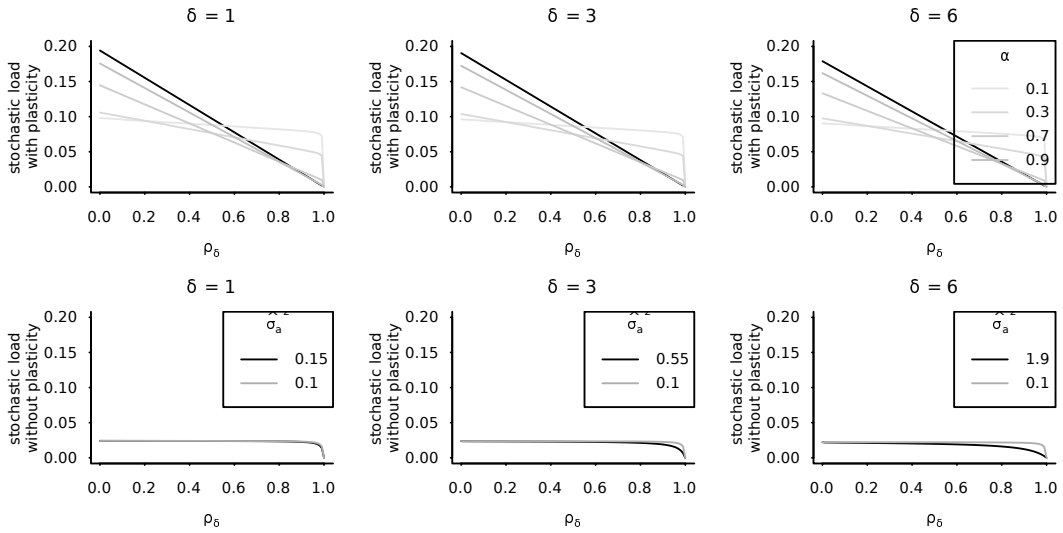

Figure S2: Top row: Stochastic load with plasticity plotted against ρ_δ for various values of relative plasticity α. Bottom row: Stochastic load without plasticity (i.e., load component for "positively auto-correlated fluctuations" from Table 1 of Lande and Shannon 1996) plotted against ρ_δ for equivalent values of total additive genetic variance. Specifically, the additive genetic variance in reaction norm elevation is adjusted to value $\hat{\sigma}_a^2$, set either to its intial value ($\hat{\sigma}_a^2 = \sigma_a^2$, grey line) or to the total additive genetic variance in the new environment with plasticity in the top row ($\hat{\sigma}_a^2 = \delta^2\sigma_b^2 + \sigma_a^2$, black line). Also, note that relative plasticity α has no effect in the bottom row). Other parameters are σ_b^2 =0.05, $\tilde{S}(\delta)$ =0.0446429, B =2, $\sigma^2 = 1$, σ_a^2 =0.1.

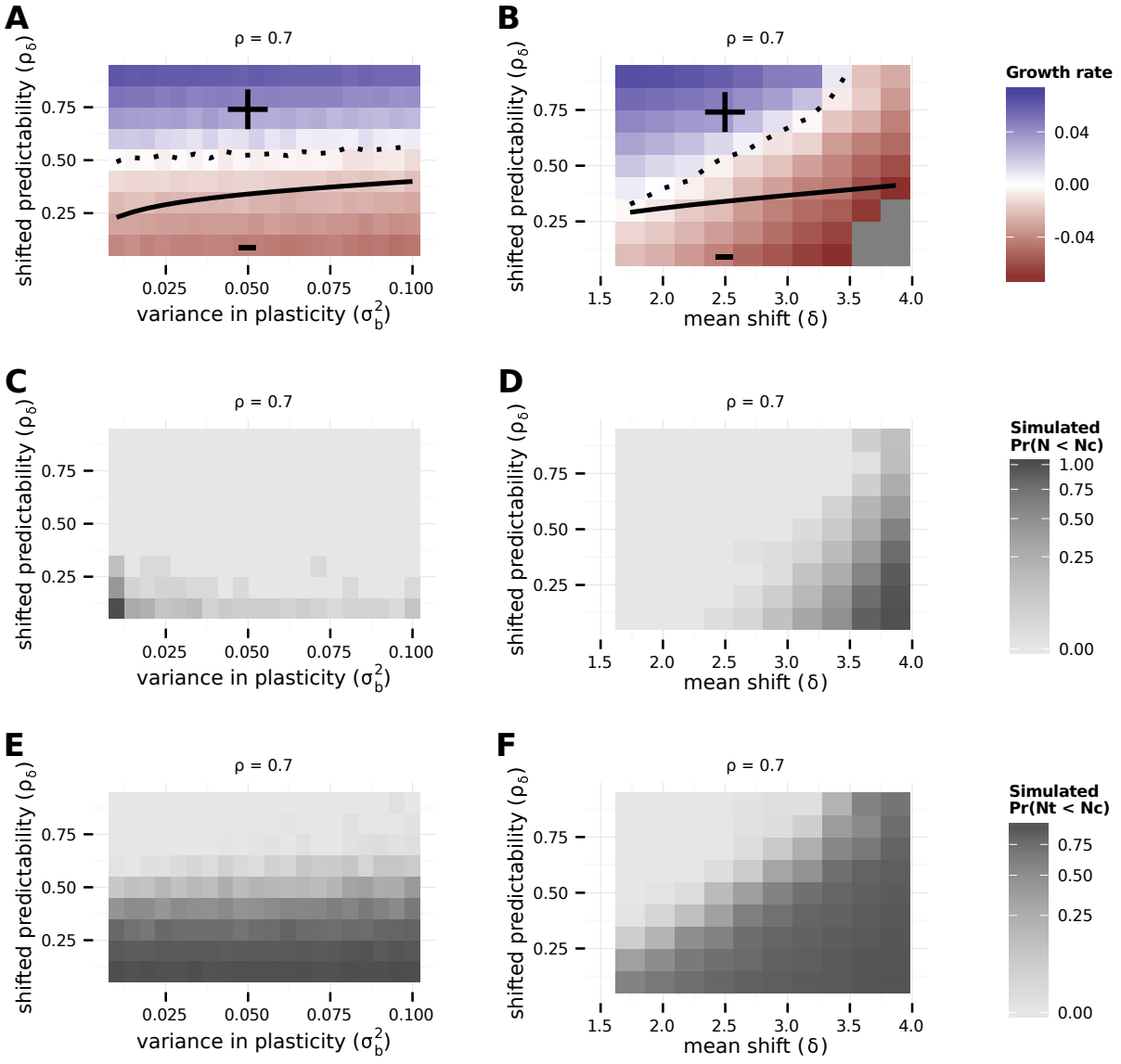

Figure S3: Figure 2 but with genetic variance changing with population size according to a modified Stochastic House of Cards (SHC) approximation: $\sigma_{SHC}(\sigma_g^2) = \sigma_g^2/(1 + \frac{\omega^2+\sigma_e^2}{\mu^2 N_e})$ and $N_e \approx 2R_0N/(2R_0 - 1)$. Parameters and panels are as in Figure 2 with the additional parameter μ^2, which is the variance of the effect of new mutations, fixed at $\mu^2 = 0.005$. For the un-modified SHC the numerator, σ_g^2, is replaced by a term that includes the per-generation total mutation rate V_m and the strength of selection: $2V_m(\omega^2 + \sigma_e^2)$ [see 23, who used α^2 instead of μ^2]. Thus, the modified version we use effectively assumes that larger initial values of σ_a^2 and σ_b^2 reflect populations with larger mutation rates.

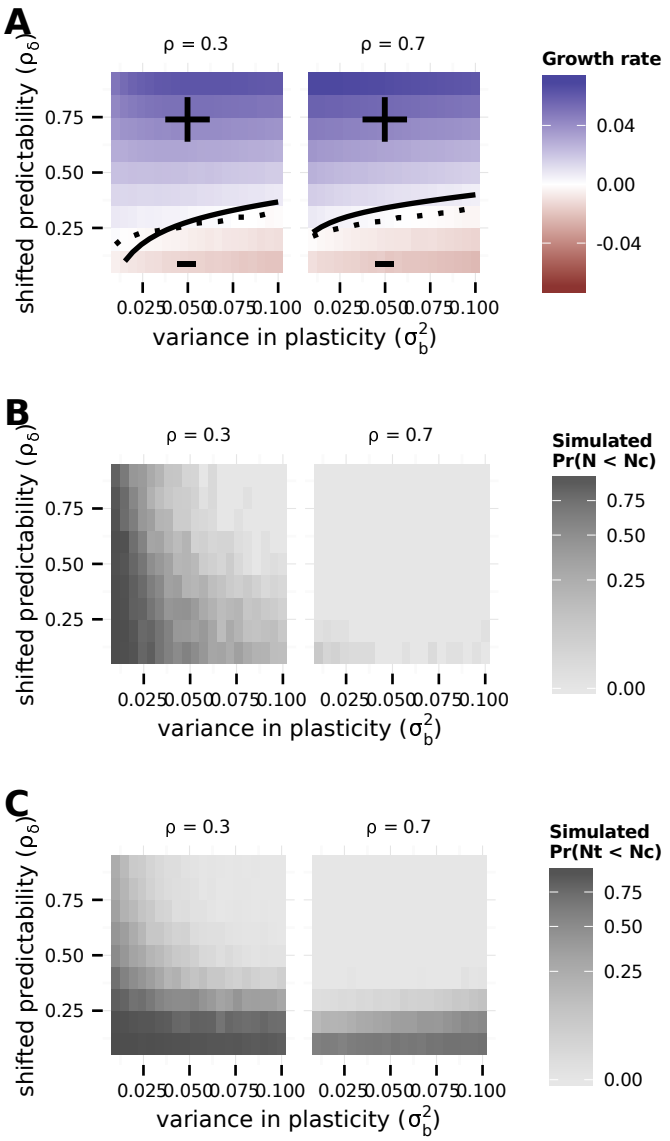

Figure S4: Potential for evolutionary rescue over a range of values for post-shift predictability ρ_δ versus genetic variance in plasticity the reference environment σ_b^2 (**A,B**). Within panels, columns show low ($\rho = 0.3$) and high ($\rho = 0.7$) initial predictability. **A** Growth rates at the end of rescue are computed from numerical simulations as the stochastic growth rate λ_s between t_{bef} and t_{aft} spanning Phase 1 and 2 (diverging heatmap: white 0, blue positive, and red negative). Black lines indicate the threshold between decline (-) and persistence (+) based on the analytical approximation ($\bar{r}_1 = 0$, eqn 3 ; solid line) and stochastic simulations ($\lambda_s = 0$, dotted black line). **B** Simulated probability of quasi-extinction before rescue. Quasi-extinction is defined at t_{bef} as illustrated in Figure 1. Shift size is set to $\delta = 2.5$ (so that additive variance increases by a factor of 4 in the new environment when σ_b^2 =0.05). Other parameters: initial population size $N(0) = 10^4$, selection strength $\omega^2 = 20$, developmental delay $\tau = 0.2$, additive genetic $\sigma_a^2 = 0.1$ and environmental $\sigma_e^2 = 0.5$ variances, and maximum fitness $e^{r_{\max}} = 1.1$.

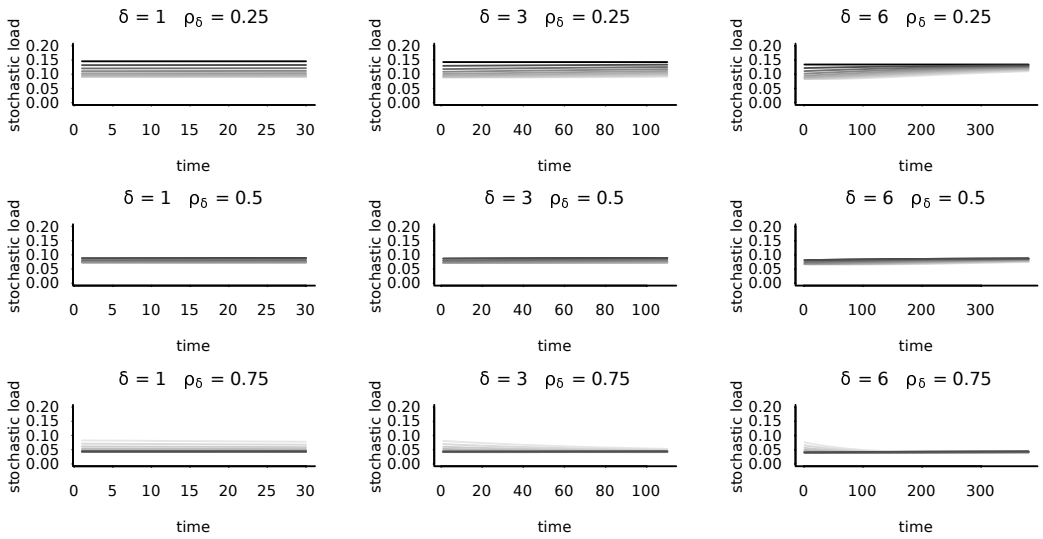

Figure S5: Quasi-stationary stochastic load plotted against time during Phase 1 (up to the characteristic timescale of Phase 2 t_2) for the same values of relative plasticity α as Figure S2. While changes in the magnitude of stochastic load over Phase 1 are not large, and in some cases not apparent, when the change reduces (or increases) the mismatch the stochastic load follows (e.g., bottom right or top right panel). Other parameters are σ_b^2 =0.05, $\tilde{S}(\delta)$ =0.0446429, B =2, $\sigma^2 = 1$, σ_a^2 =0.1.

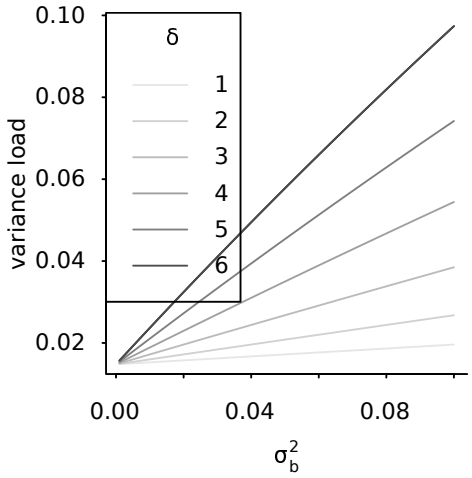

Figure S6: Variance load plotted against σ_b^2 for various values of δ. Other parameters are B =2, $\sigma^2 = 1$, σ_a^2 =0.1. Note predictability after shift ρ_δ does not affect variance load.