Natural selection as the process of accumulating adaptive information.

Ryan Seamus McGee

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2021

Reading Committee:
Carl Bergstrom, Chair
Benjamin Kerr
Thomas Daniel

Program Authorized to Offer Degree:
Department of Biology

© Copyright 2021

Ryan Seamus McGee

University of Washington

Abstract

Natural selection as the process of accumulating adaptive information.

Ryan Seamus McGee

Chair of the Supervisory Committee:

Carl T. Bergstrom

Department of Biology

A defining feature of life is its ability to respond to complex conditions by leveraging information about the environment. In many respects, evolutionary biology is interested in understanding how genetic information changes over time to produce organisms that are exquisitely adapted to their environments. The core insight of the modern synthesis is that this adaptive information is acquired by the process of natural selection acting on genetic variation. While population genetics has developed a large body of theory regarding how genetic variance changes in the process of evolution, we lack correspondingly rich theory describing how adaptive information changes as a consequence of these same dynamics.

In Chapter 1 I aim to clarify the sense in which the genome encodes adaptive information about particular environments, establish and interpret Shannon's mutual information as the appropriate measure of adaptive genetic information, and illustrate how different evolutionary forces shape this information. In addition, I demonstrate that this view of genetic information is not only of theoretical interest, but also practical and useful as a measure of adaptive differentiation.

Results from classical population genetics suppose that there is a cost of selection, measured in terms of substitutional load, that limits how much information can be gained by this process. In Chapter 2 I test, extend, and reinterpret this theory to illuminate a rigorous and meaningful relationship between information gain and fitness under very general conditions. By reframing the process of natural selection in learning theoretic terms, I clarify and formalize the cost of selection in terms of regret—a relative measure of load. I then establish general bounds on this cost and show that there is indeed a fundamental fitness cost associated with information acquisition by selection.

These results highlight the centrality of information acquisition to natural selection and the value that information-theoretic perspectives have in evolutionary biology. Overall, this work contributes to an understanding of selection that advances evolutionary theory by synthesizing population genetics, information theory, and learning theory in modeling the acquisition of adaptive information.

In the final year of my Ph.D., the world witnessed the outbreak of the COVID-19 pandemic. In response to this crisis, I shifted much of my effort to modeling the impact of non-pharmaceutical interventions for mitigating disease transmission. Infectious diseases spread within a social contact network, and many strategies for mitigating spread can be thought of in terms of this network (e.g., social distancing and contact tracing). A framework that incorporates realistic contact networks is essential to explore the efficacy of such interventions. I developed a flexible stochastic network model that incorporates the structure of data-driven contact networks and other extensions that are critical for studying COVID-19 (e.g., pre- and asymptomatic transmission, hospitalization, testing, tracing, isolation). This work led to collaborations with researchers and stakeholders across disciplines. In Part 2, I describe how we used the model framework I developed to inform model-driven mitigations for schools and workplaces.

Everything sufficiently beautiful is connected to all other beautiful things.

John Baez

For my parents, Mitch and Mary Ann.

Acknowledgements

You keep on learning and learning, and pretty soon you learn something no one has learned before.

Richard Feynman

As with all learning, the doctoral thesis is a process of ups and downs, gains and losses, regrets and revelations. I am indebted to numerous people for the wisdom they passed on, the impressions they left, and the support that allowed me to bear the load and gain this insight.

First, I would like to thank my thesis advisor, Carl Bergstrom. Carl's vast breadth of knowledge provided inspiration at every turn, and his unique expertise in evolution, information, and epidemiology seeded the questions that I pursued in this work. Beyond this, I learned a great deal from Carl about the integrity of science as a process, the value of effective communication, and the art of calling bullshit. I have great appreciation for the incredible patience and freedom Carl offered me in pursuing my wide range of interests and for the many doors his support and advocacy opened for me, both personally and professionally.

I would also like to thank my *de facto* co-advisor, Ben Kerr. Ben graciously took me on as one of his own, and his thoughtful mentorship led me to grow as a scientist and academic in countless ways. Ben fostered in me an appreciation for the beauty in conducting and communicating science from first principles, and his guidance during my journey into experimental biology gave me invaluable new perspectives on the rich interplay of theory and experiment. Ben is a model of selfless mentorship, and my academic career is immeasurably greater for the support he has extended me.

Thank you to the Kerr Lab for welcoming into your community with open arms and for all of the collaboration and camaraderie I shared with you all. I want to extend a special thank you to Katie Dickinson, whose dedication and consideration far above and beyond the duty of lab manager made possible my development as an experimentalist, as a mentor, and more.

This work would not have been possible without the contributions of a fabulous community of co-authors and collaborators. In particular, I wish to extend a special thank you to Olivia Kosterlitz for being a steadfast companion at the bench, at the whiteboard, and at every twist and turn of life in graduate school. Thank you to Artem Kaznatcheev for opening engaging lines of dialog and for pushing my work in compelling new directions. Thank you to Alicia Zhou and the Color Health Scientific Affairs team for supporting my work on modeling interventions for COVID-19 and for helping to bring this work to a wider audience. I would also like to thank a host of collaborators whose interactions and mentorship have given me an invaluable bank

of knowledge, experience, and questions to explore—Eli Shlizerman, Adrienne Fairhall, Michael Lachmann, Noa Pinter-Wollman, Julie Miller, Jevin West, Tom Daniel, and many others.

I am very thankful to my wife Sydney, who has graciously shaped her own life and career to accompany me on this journey and for the love and support she has given me along the way.

Finally, I am thankful to my parents who instilled in me a lifelong passion for learning about the world, and who have encouraged and supported my curiosity and education at every stage. It is to them that this thesis is dedicated.

I was supported throughout my graduate studies by a NSF Integrative Graduate Education and Research Traineeship (IGERT) in Big Data and Data Science, a NSF Graduate Research Fellowship, and grants from the NSF BEACON Center for the Study of Evolution in Action. I have also received financial awards from the University of Washington Department of Biology, the Seattle Chapter of the ARCS Foundation, and from the Rockefeller Foundation. I would like to express my sincere gratitude to each of these institutions for the support that made this work possible.

Contents

Part 1. Natural selection as the process of accumulating adaptive information. 1

Chapter 1. On the meaning and measurement of adaptive genetic information. 3

The meaning of adaptive genetic information 4

Quantifying adaptive genetic information 7

The effects of evolutionary forces on adaptive genetic information 10

Adaptive genetic information as a measure of adaptive differentiation 18

Summary 19

Chapter 2. The cost of natural selection as an information acquisition process. 21

Natural selection as an information acquisition process 22

Fitness loss associated with information gain 23

Measuring load in realistic environments 25

The cost of selection as a learning process 28

Natural selection as a no-regret learning process. 32

Learning problems faced by evolving populations. 33

The cost of information acquisition by natural selection. 39

Discussion 41

Methods 42

Appendix 2.A. Modeling natural selection. 44

Appendix 2.B. Natural selection as a learning process. 48

Appendix 2.C. The cost of natural selection. 60

Appendix 2.D. Selection experiments 90

Part 2. Stochastic dynamical network models of interventions for the COVID-19 pandemic. 99

Chapter 3. Model-driven mitigation measures for reopening schools during the COVID-19 pandemic. 101

Model and methods 102

Results 105

Limitations 110

Summary 113

Appendix 3.A. SEIRS+ Extended SEIR Network Model 115

Appendix 3.B. School Models 118
Appendix 3.C. Sensitivity of Outcomes to Model Parameters 135

Chapter 4. Proactive COVID-19 testing in a partially vaccinated population. 155
Methods 155
Results 159
Discussion 162
Appendix 4.A. An analytic approximation for the effects of testing and vaccination .. 164
Appendix 4.B. Overview of parameter values 165
Appendix 4.C. Lower vaccine efficacy 166

References 171

Part 1

Natural selection as the process of accumulating adaptive information.

CHAPTER 1

On the meaning and measurement of adaptive genetic information.

Take this. It's part of me
and everything I know
about this emergent art
of getting by.

Since what I am survived
this long, this place,
my information may enable you
to live a little.

Jean Lenski, *The First Gift*

A defining feature of life is its ability to adapt to complex conditions by leveraging information about the environment. Much of the information for generating and maintaining an organism that is well-suited to its environment is encoded in the genome. Not surprisingly, the concept of information is central to our understanding of the genome. We say that the genetic code is transcribed, translated, copied, proofread, corrected, and transmitted across generations. In many respects, evolutionary biology is interested in understanding how genetic information changes over time to produce organisms that are exquisitely adapted to their environments. The field of population genetics has developed a rich set of tools for describing how genetic variation changes over time, but an understanding of how these changes relate to the information content of genomes is lacking.

Shannon information theory provides a framework for formalizing and quantifying concepts of information (Shannon 1948). Information theory has been used extensively in other fields of biology where information plays a major role, such as neuroscience and systems biology. This has led to major advances and has fundamentally changed the way that both theorists and experimentalists approach these systems (Agüera y Arcas et al. 2001, Borst and Theunissen 1999, Cheong et al. 2011, Dubuis et al. 2013, Levchenko and Nemenman 2014, MacKay and McCulloch 1952, Niven and Laughlin 2008, Palmer et al. 2015, Strong et al. 1998, Tikhonov and Bialek 2013, Tkacik and Bialek 2016). While some theorists have studied the genome and population genetics in terms of Shannon theory (Adami 2012, Felsenstein 1971, Frank 2012, Harper 2009a, Kimura 1961a, Kolchinsky and Wolpert 2018), information-theoretic thinking has not taken hold, opened new lines of inquiry, or found practical use in evolutionary biology to the same extent it has in other fields. This is due in part to the fact that the mapping of Shannon's transmission sense of information onto evolutionary systems is not as immediate as it is for other systems where signals and transmission channels are more obvious (Bergstrom and Rosvall 2011). This has led to a number of misconceptions about the meaning

of genetic information and an ongoing philosophical debate about what, if anything, the genome represents semantically (Godfrey-Smith 2007, Griffiths 2001, Maynard Smith 2000, Shea 2007a, Sterelny 2000).

The aims of this chapter are to clarify the sense in which the genome encodes adaptive information about particular environments, establish and interpret Shannon's mutual information as the appropriate measure of adaptive genetic information, and illustrate how different evolutionary forces shape this information. In addition, we aim to demonstrate that this view of genetic information is not only of theoretical interest, but also practical and useful as a measure of adaptive differentiation.

The meaning of adaptive genetic information

Evolutionary biologists are interested in identifying genetic instructions that inform adaptations to particular environments and tracking how this adaptive information changes over time. For example, how much genetic information is associated with a particular adaptive trait? How quickly is information about a new niche encoded in the genome following an environmental shift? How much information have two divergent populations acquired about their respective environments since their split? To answer these questions, we seek a definition that distinguishes *adaptive* information from other forms of information carried by the genome. In order for genetic information to be adaptive, it must 1) inform the development of a phenotype that is adaptive for a particular environment, 2) be accumulated by the process of natural selection in that environment, and 3) be informative at the level of the individual organism.

We implicitly understand the genome as a store of information, but what exactly is meant by "genetic information?" Broadly speaking, information can be understood as "that which reduces uncertainty." In other words, an object Y provides information about a system X if observing Y reduces uncertainty about the state of X to some extent. In this general sense, there are multiple ways in which the genome carries information.

When the states of two systems are correlated, knowing the state of one system allows us to narrow down the state of the other. The sequence of a gene is correlated with the structure of the protein it encodes. The ensemble of regulatory sequences in the genome correlate with the transcriptome of the cell. Certain SNPs may be correlated with geography, and different genotypes are often associated with particular environmental conditions. In all of these cases, knowing a particular genetic sequence allows us to predict other variables better than chance. Thus the genome carries *correlational* information about these systems. The fact that DNA encodes correlational information about other features is the basis for many fields of biology in the age of sequencing—from molecular phylogenetics, to bioinformatics, personalized medicine, and other areas.

Of course, the genome does not just encode barcodes for the convenience of biologists. Genes code *for* the structure of particular proteins. Regulatory regions code *for* the binding of particular transcription factors. That is, the genetic material biochemically determines the states of other systems and influences phenotype to some extent. Genomes are correlated with protein structures, regulatory states, and phenotypes because these features have some degree of dependency on the DNA. Genetic sequences specify RNAs and polypeptides with relative certainty (errors and redundancy in transcription notwithstanding) because the genetic code is a specific and consistent mapping between chemical symbols. In this sense, the genome is said to carry *causal* information. Cracking the genetic code is one of the greatest breakthroughs in biology to date, and our understanding of causal genetic information has enabled advances in genetic engineering, synthetic biology, and mRNA vaccine technology.

While these notions of genetic information pervade modern biology, neither correlational nor causal information refer to the information that specifically encodes adaptations for an organism's environment. Genetic sequences can just as easily correlate with maladaptive phenotypes as adaptive ones, and the same amount of causal sequence information can specify either a highly adapted or nonfunctional protein. As a result, quantifying correlational or causal genetic information does not necessarily measure the degree to which an organism's genome informs its fit to a particular environment. Adaptations evolve over time as genomes are passed from one generation to the next and natural selection acts on the reproductive success of the information they contain. The correlational and causal notions of genetic information fail to capture the heritability and evolvability essential to genetic information.

The genome is not merely a template for transcribing and binding proteins. Rather, a critical function of the genome is to convey information from one generation to the next. DNA has a number of features that make it remarkably effective as a medium for storing and transmitting information. As Bergstrom and Rosvall (2011) point out, DNA is "a molecule that is exquisitely fashioned so as to (1) encode lots of sequence information in a small space, (2) be incredibly easy to replicate, (3) be arbitrarily and infinitely extensible in what it can say, and (4) be structurally very stable and inert." While one can imagine other paradigms that could fill the role of templating and replicating enzymes, such as the polypeptide template thought experiment entertained by Godfrey-Smith (2000a) or the RNA world hypothesis (Rich 1962), it is hard to imagine a system that is fundamentally different from DNA yet possesses the same information storage and transmission capabilities. Not even today's computer hardware can match the space and replication efficiency of DNA. These properties support a teleofunctional view that the primary function of the genome is to transmit information over generations (Bergstrom and Rosvall 2011). This implies that the contents of the genome have an inherently and intentionally communicative nature, which Bergstrom and Rosvall (2011) refer to as the *transmission* sense of genetic information.

A transmission sense of information falls under the domain of Shannon information theory (Shannon 1948). The basic conceptualization of Shannon theory considers an agent that receives a signal that has been transmitted through a noisy channel. The amount of information conveyed by the signal is the amount by which its contents reduce the recipient's uncertainty about some other system (as formalized in the following section). Shannon theory measures uncertainty and information in terms of the conditional statistical properties of signals while remaining deliberately agnostic to their semantic meaning. However, it is still possible to speak to the referent of Shannon information: the system that a signal conveys information "about." If an object Y reduces an observer's uncertainty regarding a system X, then Y carries information *about* X. In other words, Y encodes a *representation* of X that is predictive to some degree.

What does the genome transmit information about? What does it represent? A conventional view supposes that genotype encodes a representation of phenotype (Maynard Smith 2000). This view is in line with the central dogma of molecular biology, but it breaks down when considering the genome as an informative representation that is to be transmitted across generations. First, genetic sequences represent the phenotypic properties of an organism as far as the primary structures of proteins, but not much further. Genetic determinism is limited, and higher order traits are shaped by emergent, contingent, and environmental interactions that are not represented in the genome (Godfrey-Smith 2000b, 2007, Griffiths 2006, 2001, Kitcher 2001). If the genome is a representation of phenotype, it is not a very informative one. Furthermore, viewing the genome as a representation of phenotype puts causality at odds with the directionality of transmission. An organism's

phenotype is caused by the genetic material (to the extent that it is) after it is inherited. Having no access to the traits of its offspring before the genetic material is replicated and passed, the parent cannot encode a representation of its offspring's particular state before it is transmitted. Any system that is represented in the genome in a transmission sense must be external to and prior to the genetic material and the properties that follow from it. The genome does carry correlational and causal information about the offspring's phenotype, but this is incidental to whatever representation it encodes.

At this point, it is useful to further formalize what a genetic "representation" entails. In a Shannon transmission sense, Y is a representation of X if the function of Y is to reduce uncertainty about X on the part of an agent that observes Y (Bergstrom and Rosvall 2011). Shea (2007b,a) introduced criteria for identifying representational information by relating the process by which a representation is established (encoded) with the means by which that representation is interpreted (decoded). Similar to Shannon, Shea's framework involves a producer that generates a range of alternative intermediates (e.g., signals, symbols) and a consumer that "responds systematically to each different intermediate." If there is a unique "success condition" associated with the output of the consumer system in response to each intermediate, and if the producer generates intermediates such that intermediates are correlated with the corresponding success conditions, then the intermediates are thereby representations of the associated success conditions.

Applying Shea's framework to the genome, genetic sequences are intermediates that are "consumed" by the cellular machinery and developmental system to produce the phenotypic traits of an organism. The success condition is for the organism to occupy an environment where these traits confer high fitness. If the genetic sequences that organisms possess are correlated with the environments they experience, then the genome encodes a representation of the environment that reduces uncertainty about prevailing conditions (Shea 2007a). Natural selection establishes such a correlation between genetic sequences and environmental conditions, and therefore acts as the "producer" in Shea's framework. Due to the variation in traits conferred, not all genes are equally likely to be transmitted to the next generation. The sequences that are associated with phenotypes that have high relative fitness in an environment increase in frequency, and an adaptive matching between genotypes and environments develops. Thus selection results in an accumulation of representational information in the genome about the environmental history.

An example borrowed from Shea (2007a) helps to make this clear. Suppose there are two alleles in a population: one that leads the organism to develop thick fur, and another that leads to thin fur. The thick fur allele confers high fitness in cold environments, but is disadvantageous in warm conditions. Conversely, the thin fur allele has a fitness advantage in warm environments but not in cold (i.e., there is a unique success condition associated with each variant). If a population experiences an environment that is predominantly cold over evolutionary time, then natural selection increases the frequency of thick-furred organisms, and the corresponding allele comes to be correlated with the population's environmental history. As a result, the frequency of the thick fur allele in the gene pool allows us to infer that the environment has been cold, and the genome encodes representational information about the environmental history of the population. Transmission of this information from parent to offspring can be interpreted as mother communicating to daughter, "*The environment is cold (conducive to thick fur), so grow thick fur*" (Shea 2007a).

This perspective integrates the correlational, causal, and transmission senses of information. Genetic sequences transmit representational information because they are (to some degree) correlative with respect to the environment ("*the environment is cold*") and causal with respect to phenotype ("*grow thick fur*"). Critically,

the representational information in the genome is established by the environment-dependent process of selection and is associated with high relative fitness in the particular conditions in which it accumulated. Previous work has shown that this kind of information has fitness value; that is, the long-term fitness of a lineage is proportional to the amount of information it has about the environment (Bergstrom and Lachmann 2004, Donaldson-Matasci et al. 2010, Hilbert 2017, Rivoire 2016, Rivoire and Leibler 2011).

This is precisely the genetic information that we wish to track in evolutionarily biology. Having established these concepts, we can now offer a formal definition of *adaptive* genetic information.

> **Adaptive genetic information** refers to the contents of the inherited material that encode a representation of the population's environmental history and thereby reduce the organism's uncertainty about the current environment such that its expected fitness is greater than it would be by chance.

Quantifying adaptive genetic information

Now we turn the matter of measuring *how much* adaptive information is encoded in a genetic sequence. Adaptive information refers to the contents of the genome that represent—and thereby reduce uncertainty about—the environment. This notion conforms precisely to the quantitative definition of *mutual information* from Shannon information theory, as we summarize here.

Consider a system X that exists in a particular state x with probability $P(x)$. If one state has high probability and all other states have low probabilities, then one can be fairly certain that the system exists in the high-probability state. On the other hand, if the states are more equiprobable then the system is more uncertain. This notion of uncertainty is quantified by the **entropy**

$$H(X) = -\sum_x P(x) \log P(x) \,. \tag{1.1}$$

The more uncertain the system the more uniform the probability distribution and the higher its entropy.

We are interested in quantifying how informative a representation is about the state of a system. In other words, how much does observing a representation Y reduce uncertainty about the state of the system X? The representation can only carry information about the system if the system and the representation are somehow correlated. The more correlated they are, the more precisely the representation specifies the state of the system. The **mutual information** $I(Y;X)$ quantifies how much the entropy of (uncertainty about) the system X is reduced when given the representation Y

$$I(Y;X) = H(X) - H(X|Y) \,. \tag{1.2}$$

Therefore, the amount of **adaptive genetic information** that a genotype G provides about the environment E of the population is quantified by the mutual information

$$I(G;E) = H(E) - H(E|G) \,. \tag{1.3}$$

The set of genotypes that G refers to may be defined as any set of alternative genetic sequences (e.g., bases at a single nucleotide position, alleles of a gene, collections of SNPs, k-word sequences, whole-genome labels) or

other inherited components that carry information across generations in a transmission sense (e.g., methylation tags, chromatin structure). The set of environments that E refers to can be defined as any set of mutually exclusive (non-overlapping) conditions.

Organisms draw on their individual genome in order to develop in accordance with the environment. Therefore, it is critical for a notion of adaptive information to reduce uncertainty at the level of the individual, rather than exclusively at the population level. We can think of an individual organism inheriting a particular genome g as the individual drawing a random genotype from the gametic gene pool. Upon observing that the genotype g was drawn, the individual might reasonably infer that that the environment is one where g is present at high frequency because the traits it causally encodes are favorable. The amount of information that this particular genome provides to the individual about its environment is quantified by the reduction in entropy conditioned on this observation

$$H(E) - H(E|g)\,. \tag{1.4}$$

The mutual information $I(G;E)$ (Equation 1.3) is equivalent to the expected reduction in uncertainty across all individuals in the population

$$I(G;E) = H(E) - \mathbb{E}_G\left[H(E|g)\right]\,. \tag{1.5}$$

Mutual information is a general measure of correlation that can be equivalently expressed and understood as the divergence from independence of two random variables. Imagine that a population is subdivided and introduced to a set of novel environments (such as in an isolation event, or an experimental evolution study). Initially, the genotype frequencies of each population are equal and arbitrary with respect to the features of the respective environments. At this point, G and E are independent, and their joint distribution is equal to the product of their marginal distributions $P(G)P(E)$. Over time, natural selection in each environment creates a covariance between genotypes and environments due to environment-dependent changes in genotype frequencies. The adaptive genetic information $I(G;E)$ is equivalent to the Kullback-Leibler (KL) divergence between the actual joint distribution $P(G,E)$ that results from selection and the hypothetical product distribution $P(G)P(E)$ that assumes independence

$$I(G;E) = D_{KL}(P(G,E)\;||\;P(G)P(E)) \tag{1.6}$$

$$= \sum_{g,e} P(g,e)\log\frac{P(g,e)}{P(g)P(e)}\,. \tag{1.7}$$

This interpretation reinforces that adaptive genetic information reflects the correlations between genetic sequences and environmental histories that are established by selection.

It is important to note that the calculation and interpretation of mutual information depends on how alternative genotypes and environments are delineated. In particular, the maximum amount of adaptive genetic information that can be encoded is proportional to the logarithm of the number of genotypes or the logarithm of the number of environments, whichever is smaller. This may seem to introduce an arbitrarity to the measurement of adaptive genetic information, but this is in fact a meaningful and conceptually necessary feature of this quantity.

The set of environments that E can refer to defines the set of counterfactual conditions that adaptation is considered with respect to. Indeed, the concept of adaptation is inherently counterfactual. Identifying a trait

as an adaptation for a given context is to say that this trait confers high fitness for that context *in particular*, as opposed to all other contexts where it is less fit. For example, we say that the morphology of a penguin's wing is an adaptation for an aquatic environment, because it serves as a highly-effective flipper in this setting, as opposed to all other contexts where it is more or less nonfunctional. If all birds had essentially the same wing morphology, we would not say that different wings were adapted to locomotion in different environments. Similarly, streamlined body shapes are considered a convergent adaptation for an aquatic lifestyle because multiple species that all occupy aquatic niches exhibit this trait, while no exclusively terrestrial species do so. If the frequency of torpedo-like bodies were as frequent on land as they are at sea, then it would not be clear that penguins and dolphins share a similar body plan because they have adapted to the same environment.

Just as it is impossible to identify adaptation to a particular environment without comparison to evolutionary outcomes from at least one control condition, it is impossible to measure mutual information (i.e., it is undefined) without realized distributions of genotypes from at least two counterfactual environments. For example, there is presumably some amount of information encoded in the genome about the rate of gravitational acceleration on Earth, since bone density, blood pressure, and other traits are likely adapted to the strength of gravity on this planet in particular. However, the lack of genetic material from populations evolved under counterfactual gravitational constants makes it impossible to measure how much genetic information pertains specifically to this parameter. Only by evolving a population elsewhere in the universe and comparing the derived genomes to those evolved on Earth can this be determined.

The number and scope of the counterfactual environment classes define the specificity of adaptation that is to be considered and the resolution of genetic representations that can be encoded. If the space of environmental conditions is divided into 2 classes—hot and cold, say—then the genome can encode up to $\log_2(2) = 1$ bit about which of these conditions a population has experienced. On the other hand, if the range of possible conditions is quantized into 16 temperature ranges, then the genome can encode up to $\log_2(16) = 4$ bits about which temperate range it has evolved in. Why does the maximum amount of attainable information increase just by increasing the number of environment bins? If there are unique genotypes that are favored in many or all of the 16 temperature ranges, then selection can establish finer-grained correlations between genotypes and environmental conditions, and observing a particular genotype decreases uncertainty about the environment to a greater extent. However, alleles that are optimal in 'hot' and 'very hot' conditions, respectively, are redundant when it comes to representing a 'hot' environment as opposed to a 'cold' environment and do not provide any additional information about this binary delineation of environments.

The amount of adaptive genetic information that can be encoded also depends on the number of distinct genotypes that selection can act upon. In Shea's framework, the number of alternative intermediates corresponds to the number of conditions that can possibly be represented. In other words, the number of distinct genotypes defines the resolution of a genetic sequence's representational ability and thus its information capacity. If there are only two alleles A and a, then this genetic material can only represent two classes of environments: those where A is more fit than a, and those where a is superior to A. The maximum amount of adaptive information that can be encoded by this locus is 1 bit, regardless of how many counterfactual environments might be defined. If more alternative genotypes become available, such as by mutation, then selection may establish representations that resolve between environments more finely and thus convey more information.

The definition of mutual information $I(G;E)$ incorporates all of these intuitions about the representational resolution of adaptive genetic information out of the box.

The effects of evolutionary forces on adaptive genetic information

In order for genetic information to be truly adaptive, it should be accumulated by the process of natural selection and not by other non-adaptive processes. In this section, we show that our definition of adaptive genetic information is indeed accumulated by selection but not by drift, mutation, or migration in expectation.

To aid in this analysis, we introduce the following expression that tracks the mutual information between genotypes and environments over time, which we interpret below.

$$I^t(G;E) = \underbrace{\mathbb{E}_E\left[D_{KL}\left(P^t(G|e)||P^0(G|e)\right)\right]}_{\text{average change in each environment}} - \underbrace{D_{KL}\left(P^t(G)||P^0(G)\right)}_{\text{change common to all environments}} \tag{1.8}$$

Here $P^t(G|e)$ denotes the genotype frequency distribution in a particular environment e at time t, and $P^t(G)$ is the overall distribution of genotypes when populations from all environments are pooled together. The corresponding initial genotype distributions are given by $P^0(G|e)$ and $P^0(G)$, respectively. The derivation of this expression assumes that the initial genotype distribution is the same in all environments. The KL divergence $D_{KL}\left(P^t(G|e)\ ||\ P^0(G|e)\right)$ measures how much allele frequencies have changed in environment e by time t, and the divergence $D_{KL}\left(P^t(G)\ ||\ P^0(G)\right)$ measures how much the composition of the pooled metapopulation changes in this time. When genotypes are advantageous in one environment but delterious in another, environment-specific genotype frequency changes will tend to cancel out in the pooled metapopulation. Therefore this latter divergence term quantifies the central tendency of population change common to all environments.

Equation 1.8 tells us that the mutual information $I^t(G;E)$ at time t can be decomposed as the average amount of change in genotype frequencies in each environment minus the change that is redundant across environments. Adaptive genetic information refers to what is left: the amount of genetic change that is *specific* to *particular* environments. Adaptive information accumulates when populations that experience different conditions evolve in distinct, environment-dependent ways.

In the following sections, we discuss the effects of selection, drift, mutation, and migration on adaptive genetic information. Changes in adaptive information as a result of these processes are illustrated using a simple model of haploid evolution with discrete generations .

Selection. Natural selection updates the composition of a population such that genotypes with greater than average relative fitness in the current conditions are enriched. Selection enacts changes that are inherently environment-dependent and establishes an adaptive matching between genotypes and environments as a result. This leads to an accumulation of adaptive genetic information over time, as quantified by the mutual information $I^t(G;E)$. This quantity increases any time populations in different environments change in ways that make them more distinct. Given the environment-dependent nature of selection, this process leads to increasing adaptive information whenever different populations evolve in selective conditions that are non-identical with respect to the available genetic variation.

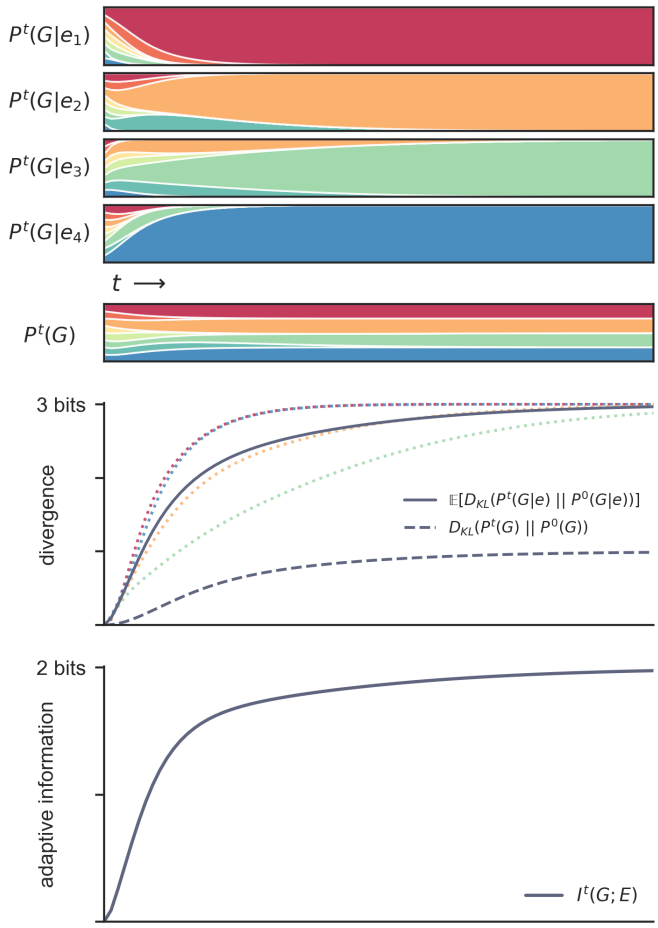


FIGURE 1.1. **The accumulation of adaptive information by natural selection.** Populations consisting of 8 alternative genotypes evolve by natural selection in 4 environments characterized by unique NK fitness landscapes (N=3, K=1). Natural selection proceeds according to discrete-time replicator dynamics with an infinite population size, no mutation, and no migration. (**Top**) Genotype frequency distributions for the populations in each environment are depicted over time by Muller plots (i.e., stacked area plots) labeled $P^t(G|e_j)$. The frequencies of the 8 genotypes are represented by the thickness of the 8 colored bands. The Muller plot labeled $P^t(G)$ gives the distribution of genotype frequencies when populations from all environments are pooled. (**Middle**) Colored dotted lines track the KL divergences of genotype frequency distributions in each environment, where the color of the line corresponds to the color of the optimal allele in the respective environment (i.e., red: e_1, gold: e_2, green: e_3, blue: e_4). The solid line tracks the KL divergence of genotype distributions averaged over all environments, and the dashed line tracks the KL divergence of genotype frequencies when populations from all environments are pooled. (**Bottom**) The amount of adaptive genetic information that has been accumulated by selection is plotted over time. This quantity is equal to the difference between the solid and dashed lines in the middle panel, as given by Equation 1.8.

Figure 1.1 illustrates an example of information acquisition by natural selection. Populations initially consisting of 8 alternative genotypes at uniform frequency are split into 4 environments with distinct fitness

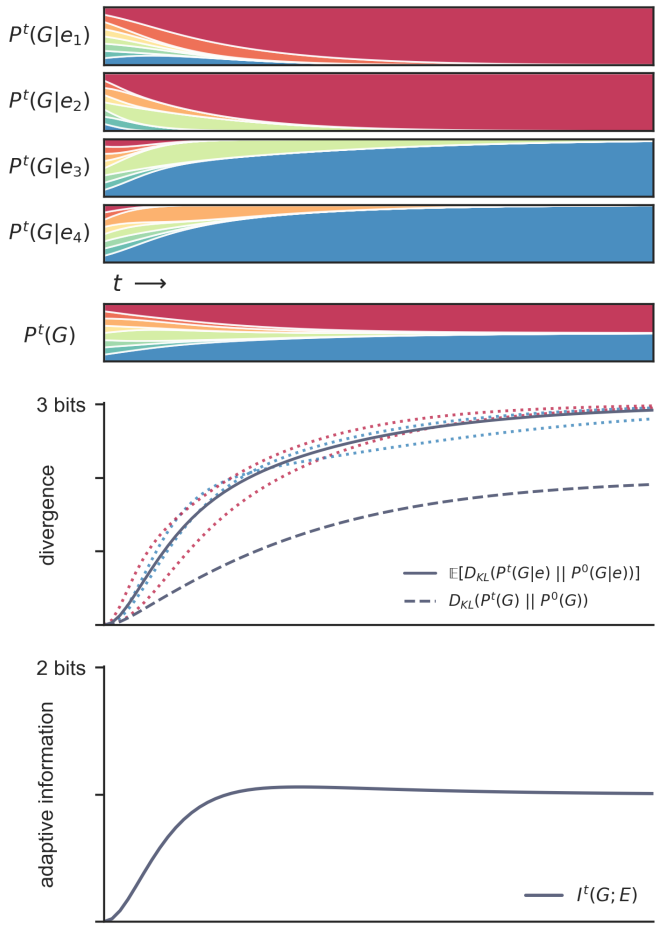


FIGURE 1.2. **Convergent evolution leads to less adaptive information.** Populations consisting of 8 alternative genotypes evolve by natural selection in 4 environments characterized by unique NK fitness landscapes ($\mathcal{N}$=3, $\mathcal{K}$=1). The fitness landscapes in environments e_1 and e_2 are correlated such that the same genotype is optimal in both; likewise for environments e_3 and e_4. Natural selection proceeds according to discrete-time replicator dynamics with an infinite population size, no mutation, and no migration. (**Top**) Genotype frequency distributions for the populations in each environment are depicted over time by Muller plots labeled $P^t(G|e_j)$. The Muller plot labeled $P^t(G)$ gives the distribution of genotype frequencies when populations from all environments are pooled. (**Middle**) Colored dotted lines track the KL divergences of genotype frequency distributions in each environment, where the color of the line corresponds to the color of the optimal allele in the respective environment. The solid line tracks the KL divergence of genotype distributions averaged over all environments, and the dashed line tracks the KL divergence of genotype frequencies when populations from all environments are pooled. (**Bottom**) The amount of adaptive genetic information that has been accumulated by selection is plotted over time.

landscapes. A different genotype confers the highest fitness in each environment, and selection fixes the optimal genotypes in their respective environments (Figure 1.1, Muller plots at top). In this case, selection establishes a one-to-one correspondence between genotypes and environments such that observing the presence

of a particular genotype determines the environment from which it was drawn with certainty. Equation 1.8 allows us to track the accumulation of adaptive information throughout the process. Selection causes substantial shifts in the genotype frequencies in each environment, as measured by the KL divergence of pre- and post-selection genotype frequencies in each environment (Figure 1.1, colored dotted lines in middle panel), and the average amount of change across environments approaches 3 bits worth of divergence (Figure 1.1, solid line in middle panel). However, not all of the selective change is specific to particular environments: 4 genotypes are lost in all of the environments, which does not contribute information that differentiates adaptations to these environments in particular. The amount of change that is common to all environments is quantified by the KL divergence of the genotype frequencies pooled from all environments, which amounts to 1 bit of total redundant divergence in this example (Figure 1.1, dashed line in middle panel). Subtracting the redundant change from the average total change, we find that selection ultimately establishes 2 bits of adaptive genetic information. These 2 bits of information reflect the ability of the evolved genetic representations to completely disambiguate 4 environments in this scenario.

The amount of adaptive information that selection can accumulate depends in part on the similarity of the fitness landscapes in different environments. When there is sufficient genetic variation that the optimal genotype in each environment is unique, selection can establish representations that resolve between each environment. However, if the fitness landscapes associated with different environments are correlated to some degree, then populations will converge in their evolution. This limits the total amount of adaptive genetic information that can be acquired, because a particular genotype may represent any one of the environments that favor that type. Figure 1.2 illustrates a scenario where the fitness landscapes in environments e_1 and e_2, although nonidentical, both favor the same genotype; likewise for environments e_3 and e_4. The amount of divergence in each environment is similar to the example with non-correlated fitness landscapes in Figure 1.1, but more of this change is redundant across environments in the correlated landscape case. Selection establishes 1 bit of adaptive genetic information, which provides information about convergent adaptation to environments e_1 and e_2 in contrast to e_3 and e_4, but does not differentiate e_1 from e_2. This reflects that populations have adapted to one set of environments or the other, but adaptation to individual environments has not occurred. Perhaps if more genetic variation is introduced selection may be able to evolve representations that further differentiate individual environments and thereby provide more information (see Mutation section below).

Genetic drift. In finite populations, random sampling of genotypes in the formation of each new generation can cause populations to differentiate. If stochastic differences between populations correlate with distinct environments, then genetic drift can lead to mutual information between genotypes and environments by chance.

However, information accumulated by drift is not adaptive. Consider two neutral alleles that fix in two different populations, respectively, as the result of drift alone. Given that these alleles are neutral (have the same fitness) in both populations, the environments they occupy are identical with respect to selection at this locus. While the observation of one allele or the other reduces uncertainty about the population from which it is drawn, this observation provides no information about distinguishing features of the environment and is not associated with a benefit to expected fitness.

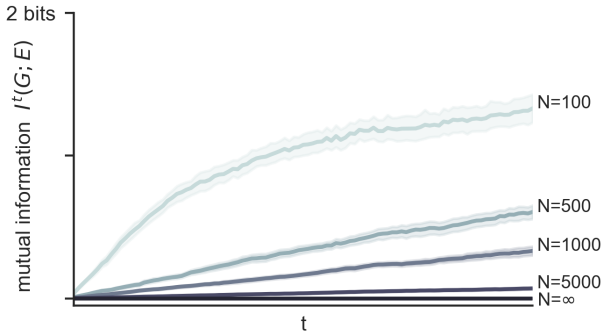


FIGURE 1.3. **Population size and information attributed to drift.** Populations consisting of 8 neutral genotypes experience genetic drift in 4 distinctly labeled but selectively neutral environments. The composition of each generation is determined by uniform random sampling of N genotypes from the parental generation. The accumulation of mutual information $I^t(G;E)$ between genotypes and environment labels due to drift is plotted over time for various population sizes (N). Lines give the mean accumulation over 30 replicate simulations for each population size; shaded bands give the 95% confidence interval.

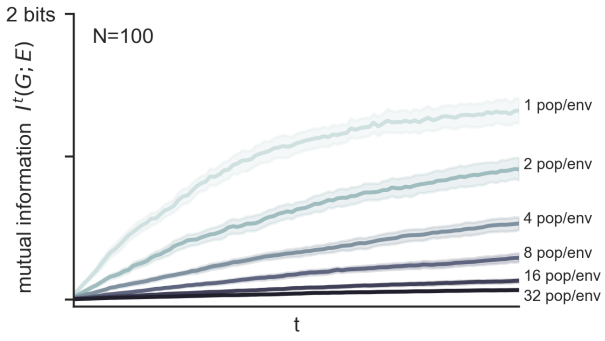


FIGURE 1.4. **Environmental replicates and information attributed to drift.** Populations consisting of 8 neutral genotypes experience genetic drift in 4 distinctly labeled but selectively neutral environments. The composition of each generation is determined by uniform random sampling of N=100 genotypes from the parental generation. In each simulation, multiple replicate populations evolve independently in each environment, and the frequencies of replicate populations in the same environment are pooled when calculating divergences with respect to that environment. The accumulation of mutual information $I^t(G;E)$ between genotypes and environment labels due to drift is plotted over time for various numbers of replicate populations per environment (*pop/env*). Lines give the mean accumulation over 30 replicate simulations for each treatment; shaded bands give the 95% confidence interval.

By definition, changes due to genetic drift do not depend on characteristic features of the environment. Therefore, genotype-environment correlations established by drift are not generally robust or repeatable. For example, genotype frequency fluctuations due to sampling error are less pronounced in larger populations, and the expected accumulation of mutual information $I(G;E)$ attributable to drift vanishes as population size increases (Figure 1.3). In addition, when multiple populations occupying the same kind of environment

are considered, drift will change replicate populations in idiosyncratic ways that will tend to cancel out when these populations are taken as a environmental group. Therefore, if multiple populations are used to determine the genotype frequencies associated with a given environment, less differentiation to specific environments will be observed. As the number of replicate populations per environment increases, the contribution of drift to the mutual information $I(G;E)$ becomes negligible, even when individual populations are small (Figure 1.4). As the population size or number of environmental replicates increase, more and more of the stochastic change observed in particular environments is seen to be redundant with change across environments, and these changes do not contribute adaptive information.

How do we reconcile the fact that mutual information is observed when there are few independent populations experiencing each environment, but not when there are many? Suppose a single representative population occupies each of two selectively neutral environments. Random sampling of genotypes in each generation causes the compositions of the two populations to diverge, and mutual information accumulates. This mutual information reflects the differential experiences of these populations, but in this case this case the information represents which genotypes happened to be in the right places at the right times to avoid falling victim to acts of God and leave descendants. Of course, this information is not about selective features that define the environment and is not associated with adaptive value. When multiple populations are measured in each class of environments, any observed mutual information reflects the degree to which random forces acted differently in each environment. As the number of populations per environment increases, the likelihood that the sum of random events in each environment are statistically distinguishable shrinks and the expected mutual information goes to zero.

Mutual information between genotypes and environments measures differentiation at the level at which environmental conditions are delineated. This quantity may refer to adaptive information, having been established by selection, or non-adaptive differentiation, having been established by drift. If environments are defined very narrowly (e.g., an individual field site versus another in an extreme example), acquisition of adaptive information about these specific conditions requires that sufficient genetic variation exists for distinct types to be present and favored by selection in one narrow set of conditions as opposed to another. If the available genetic variation is too limited or correlated in fitness effect in the defined environments, then observed differentiations are more likely to be attributable to drift. The more narrowly environmental conditions are delineated, the more replicate populations are needed to distinguish adaptive information about the history of selective conditions from information about the history of random events.

Whereas selection always increases the mutual information between genotypes and environments and thereby accumulates adaptive information, drift does not accumulate mutual information in the limit of large populations or multiple environmental replicates. Therefore, drift is not fundamentally a process that accumulates adaptive information.

Mutation. Spontaneous mutation is a random process that does not contribute to adaptive information by itself. In the absence of selection, mutation spreads the genetic composition of populations throughout genotype space in a diffusion-like process. This does not produce adaptive environment-specific differentiation, but rather makes populations more similar. When populations are small and the space of possible genotypes is large, different populations may sample mutations in different orders. In the long-term limit, recurrent

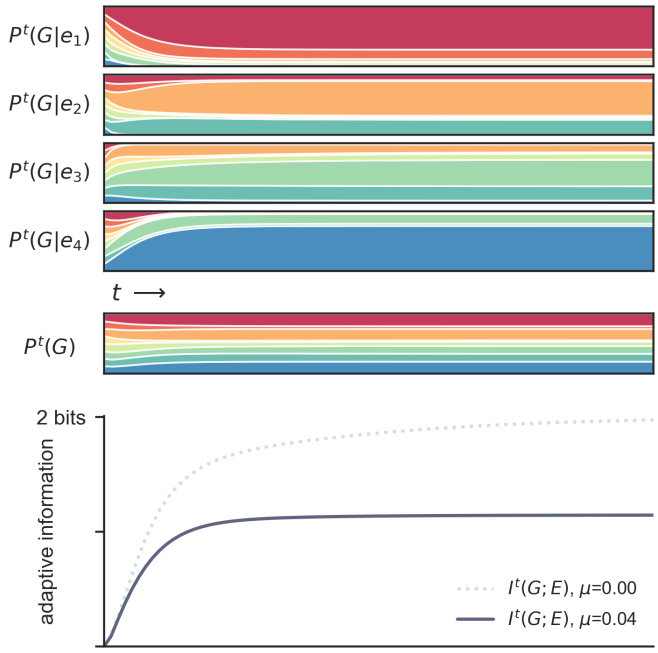


FIGURE 1.5. **Mutation counters information accumulation by selection.** Populations consisting of 8 alternative genotypes evolve by natural selection with mutation in 4 environments characterized by the same NK fitness landscapes as in Figure 1.1. Natural selection proceeds according to discrete-time replicator dynamics with infinite population size and no migration. Recurrent mutation occurs according to the genotype adjacencies of the NK landscape with a mutation rate μ=0.04. (**Top**) Genotype frequency distributions for the populations in each environment are depicted over time by Muller plots labeled $P^t(G|e_j)$. The Muller plot labeled $P^t(G)$ gives the distribution of genotype frequencies when populations from all environments are pooled. Populations evolve to mutation-selection balance compositions in each environment. (**Bottom**) The amount of adaptive genetic information accumulated by selection with mutation (solid line, μ=0.04) is compared to the accumulation of adaptive information in the absence of mutation (dotted line, μ=0, as was shown in Figure 1.1).

neutral mutation in each population will eventually make populations converge on increasingly similar genotype distributions, but populations may be differentiated by their mutational trajectories in the transient. Like drift, random mutational differentiation can contribute to the mutual information between genotypes and environments by chance. But also like drift, this information is not adaptive if not established by selection and vanishes in the limit of large population size or many replicate populations per environment.

The combination of recurrent mutation and natural selection moves a population toward a mutation-selection equilibrium. The genotype frequency distribution of a population at mutation-selection balance reflects the relative fitnesses of genotypes for the given environment, but recurrent mutations degrade some of the adaptive information that would otherwise be accumulated by preventing selection from fully differentiating populations according to their fitness landscapes (Figure 1.5).

On the other hand, mutation can also increase the maximum amount of adaptive genetic information available by introducing new genotypes that can differentially represent additional environments. If the genetic

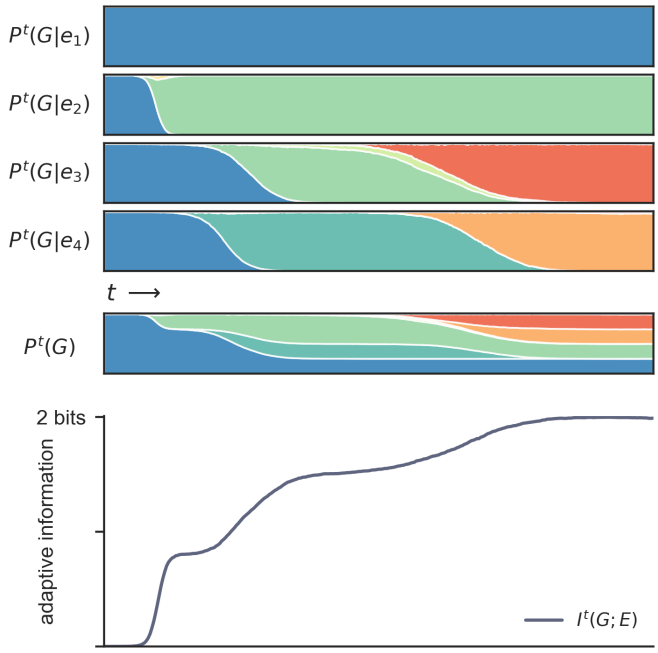


FIGURE 1.6. **Mutation increases the maximum amount of information available.** Populations that initially consist of a single common genotypes evolve by natural selection with mutation in 4 environments characterized by unique NK fitness landscapes ($\mathcal{N}$=3, $\mathcal{K}$=1). Natural selection proceeds according to discrete-time replicator dynamics with population size N=10,000 and no migration. Mutation occurs according to the genotype adjacencies of the NK landscape with a mutation rate μ=0.00002. (**Top**) Genotype frequency distributions for the populations in each environment are depicted over time by Muller plots labeled $P^t(G|e_j)$. The Muller plot labeled $P^t(G)$ gives the distribution of genotype frequencies when populations from all environments are pooled. Different mutants arise and sweep to fixation in environments e_2, e_3, and e_4. (**Bottom**) The amount of adaptive genetic information accumulated by selection with mutation is plotted over time. Adaptive information repeatedly increases and plateaus as unique adaptive mutations are discovered and selected for in each environment.

variation in a metapopulation is too limited to differentially represent all environments, the adaptive genetic information that can be encoded is likewise limited. Mutation avails additional genotypes for selection to act on, which may lead to the accumulation of additional information if selection comes to favor unique genotypes in environments that were previously convergent (Figure 1.6).

In many applications, it is problematic to calculate mutual information for a system with a changing number of bins (i.e., genotypes in the case of mutation). However, the mutual information quantifying adaptive genetic information is equivalent to the Jensen-Shannon (JS) divergence of the set of environment-specific genotype frequency distributions

$$I^t(G;E) = D_{JS}\left(P^t(G|e_1), P^t(G|e_2), \ldots, P^t(G|e_m)\right) \tag{1.9}$$

$$= H^t(G) - \mathbb{E}_E\left[H^t(G|e)\right] , \tag{1.10}$$

where $H^t(G)$ is the entropy of the pooled distribution of genotype frequencies and $\mathbb{E}_E\left[H^t(G|e)\right]$ gives the average entropy of the environment-specific distributions at time t . Environment-specific entropies are computed using the number of genotypes present in each environment at time t, and the entropy of the pooled distribution is computed using a superset of all extant genotypes. This method of calculation makes it feasible to track adaptive genetic information over time in systems with open-ended evolution.

Migration (gene flow). Migration tends to de-differentiate populations due to the homogenizing effect of gene flow. If populations in each environment receive and contribute migrants to each of the other environments at the same rate (as in Wright's island model), then migration moves the genetic composition of all populations toward the same distribution. Even if migration rates are asymmetric, any amount of gene flow from one environment to another will tend to make genotype frequencies in the respective environments more similar than they would be otherwise. Like mutation, migration counters environment-specific changes and reduces the amount of adaptive information that selection can establish. Thus restricted gene flow is necessary for local adaptation and the accumulation of adaptive genetic information.

If all migration occurs within the metapopulation of interest, then the mutual information between genotypes and environments accurately quantifies the amount of adaptive genetic information. However, genotype frequency shifts due to exogenous immigration can contribute to non-adaptive mutual information if different environments receive immigrants at different rates. If the exogenous immigration is not reliably observed in replicate populations, then this non-adaptive information will vanish in expectation as with drift and mutation. However, if different rates of exogenous immigration are consistently observed in each defined environment, then this source of genetic differentiation will confound adaptive differentiation due to selection. In such a case, not all of the mutual information between genotypes and environments represents adaptive information, but all of it is informative about characteristic features of the environment, where both selective properties and migration patterns are considered. Given that neither gene flow within the focal metapopulation nor genetic drift are expected to register non-zero mutual information, observing mutual information between environments and neutral genetic markers may be indicative of exogenous gene flow.

Environmental change. Up to this point, we have considered descriptions of environments that are consistent over the evolutionary period of interest. Environmental conditions need not be static, but some degree of characteristic autocorrelation in environmental states and corresponding fitness landscapes is implicitly assumed. After all, if the environment fluctuates randomly from one generation to the next, then genetic material inherited from the previous generation will be uninformative in any case.

If there is a fundamental shift in the conditions that define environments, pre-existing genotype-environment correlations are likely to become maladaptive. Following environmental change, selection will tend to initially decrease mutual information before acquiring adaptive genetic information about the new conditions.

Adaptive genetic information as a measure of adaptive differentiation

Assessing the relative influence of adaptive and non-adaptive forces on population differentiation and identifying regions of the genome that are under selection are important problems in evolutionary genetics. Because selection often reduces genetic variation faster than drift, mutation, or gene flow, many conventional

approaches for detecting adaptive evolution involve comparisons of allele frequency variance among populations. For example, Wright's fixation index F_{ST} and related statistics (e.g., Q_{ST}, G_{ST}, R_{ST}, Φ_{ST}) are commonly used to measure the proportion of overall genetic diversity that is attributable to allele frequency differences between populations (Wright 1951). Wright defined F_{ST} as the ratio of the variance of allele frequencies between subpopulations σ_S^2 to the expected total variance of allelic states in the metapopulation σ_T^2 assuming panmixis (Holsinger and Weir 2009, Wright 1951)

$$F_{ST} = \frac{\sigma_S^2}{\sigma_T^2} \,. \tag{1.11}$$

Divergent selection leads to higher values of F_{ST} than are expected as a result of genetic drift. Therefore, regions with high F_{ST} relative to the genomic background or to the expectations of null model of neutrality are candidates for adaptive differentiation.

Our definition of adaptive genetic information as quantified by the mutual information between genotypes and environments $I(G;E)$ provides an alternative method for measuring adaptive differentiation. Like F_{ST}, the value of the mutual information $I(G;E)$ reflects the amount of genetic variation that is associated with particular populations relative to the overall variation in the metapopulation. Both quantities increase as a result of differentiation due to selection or drift, and both are reduced by gene flow. Genetic sequences are expected to carry zero adaptive genetic information in the absence selection, so measuring $I(G;E)$ against this expectation or against background levels of mutual information may be used to identify candidate regions of adaptive evolution. Both quantities are measured using distributions of genotype frequencies from different populations. Similar to F_{ST}, estimates of $I(G;E)$ may be biased due to sampling error in the estimation of genotype frequencies, but various methods for estimating and correcting bias in estimates of mutual information exist in the information theory literature. Developing theory and estimators for the mutual information $I(G;E)$ as a measure of adaptive differentiation is an interesting direction for future research.

The direct connection between $I(G;E)$ and formal interpretations of adaptive genetic information confers certain benefits. Whereas F_{ST} takes values in (0, 1) in all contexts, the maximum value of $I(G;E)$ scales with the representational capacity of genetic variation (as determined by the number of distinct genotypes and environments), which allows for more meaningful comparisons across regions and over time. In addition, the delineation of conditional environments for adaptive genetic information provides insights regarding the conditions to which populations have adapted in their evolutionary divergence.

Summary

Genetic information has been described as "a metaphor in search of a theory" (Griffiths 2001). The relative lack of rigorous treatments of genetic information is due in part to misconceptions about the meaning of this information and its connection to evolution by natural selection. Here we have developed a formal definition of adaptive information that refers to contents of the genome that represent selective features of particular environments. This view integrates notions of correlational and causal genetic information with a transmission sense of information that captures the inheritance and evolution of genetic information. Adaptive genetic information is quantified in Shannon-theoretic terms as the mutual information between genotypes and environments, which is accumulated by divergent selection but not by non-adaptive processes such

as drift, mutation, or gene flow in expectation. This quantity can be used to measure the amount of adaptive genetic differentiation between populations, and we present multiple equivalent methods for computing adaptive genetic information from genotype frequency distributions.

Frank (2012) argues that "information is a primary quantity with intuitive meaning in the study of selection, and the history of evolutionary theory has it backwards, using statistical expressions of variances and covariances in place of the equivalent and more meaningful expressions of information." We see a rigorous and quantifiable definition of adaptive genetic information as contributing to a foundation for population-genetic theory centered on this fundamental evolutionary currency.

CHAPTER 2

The cost of natural selection as an information acquisition process.

Are God and Nature then at strife,
That Nature lends such evil dreams?
So careful of the type she seems,
So careless of the single life;

That I, considering everywhere
Her secret meaning in her deeds,
And finding that of fifty seeds
She often brings but one to bear.
...
Oh, yet we trust that somehow good
Will be the final goal of ill,
To pangs of nature, sins of will,
Defects of doubt, and taints of blood;

That nothing walks with aimless feet;
That not one life shall be destroy'd,
Or cast as rubbish to the void,
When God hath made the pile complete;

That not a worm is cloven in vain;
That not a moth with vain desire
Is shrivell'd in a fruitless fire,
but subserves another's gain.

Alfred, Lord Tennyson, *In Memoriam A.H.H.*

Living organisms are astonishingly complex and intricately adapted to their environments. To implement adaptations to particular environments, organisms require extensive information about the environmental conditions and how to function in response, much of which is carried by the genome. The core insight of the modern synthesis is that this adaptive information information is acquired by the process of natural selection acting on genetic variation.

Fundamentally, natural selection is a process that shifts the frequencies of types in a population according to their fitnesses. Types that are well-suited for the environment become more common, and poorly fit types are eliminated. The effects of selection over successive generations leave an imprint of the selective conditions on the population's composition, and the relative frequencies of types provide an increasing amount

of information about the population's environmental history. In this way, we might view selection as a kind of learning process by which the population gains information about which types are best suited for its environment. While population genetics has developed a large body of theory regarding how genetic variance changes in the process of evolution, we lack correspondingly rich theory describing how adaptive information changes as a consequence of these same dynamics.

Motoo Kimura was among the first to consider natural selection as an information acquisition process in a formal sense. Kimura (1961a) proposed a simple relationship between information gain by natural selection and the long-term growth of a population. When poorly adapted types are present, the population's mean fitness is suboptimal. As selection updates the composition of the population to better match the environment, information about the environment is acquired and the mean fitness of the population increases. But selection is gradual and there is a prolonged depression in mean fitness until an optimal composition is reached. Kimura noted that the amount of information gained in a simple allele substitution is proportional to this loss of potential fitness incurred by the incremental nature of selection, which he termed substitution load. This result led Kimura to suppose that there is a cost of selection that limits how much information can be gained by this process. Kimura's result points to a fundamental relationship between two essential quantities in evolution—fitness and information—but this insight has not been developed further.

Here we test, extend, and reinterpret this theory to illuminate a rigorous and meaningful relationship between information gain and fitness under very general conditions. First, we formalize selection as an information acquisition process and review Kimura's result relating information gain and substitution load. We then experimentally validate this basic result, generalizing this concept of load for heterogeneous and variable environments in the process. Next, we take the view of natural selection as a learning process seriously and reframe selection in learning theoretic terms. This allows us to clarify and formalize the cost of selection in terms of regret—a relative measure of load. From there we establish general bounds on this cost and show that there is indeed a fundamental and universal fitness cost associated with information acquisition by selection.

Natural selection as an information acquisition process

Learning is the process of iteratively updating a hypothesis in light of new evidence. In the context of natural selection, the learner is a population that consists of a number of replicators of varying types (e.g., genotypes, phenotypes). Each alternative type represents a strategy for how to function and survive, and the distribution of type frequencies $\boldsymbol{p}^t$ can be seen as the population's hypothesis about which type is most suited for the environment at time t. The fitness W_i of the ith type provides evidence about its suitability for the current state of the environment. Selection changes the frequency of each type i accordingly following the well-known replicator dynamics (Appendix 2.A.1.1)

$$p_i^{t+1} = \frac{p_i^t W_i}{\sum_k p_k^t W_k} \,. \tag{2.1}$$

The process by which selection updates a population's type distribution is formally analogous to Bayesian learning (Campbell 2016, Czégel et al. 2020, Harper 2009a, Shalizi 2009, Watson and Szathmáry 2016) (Appendix 2.B.1). Selection increases the frequency of types with high relative fitness in the same way that Bayes' rule increases the weight of alternatives that give high relative likelihood to the observed evidence. The population's hypothesis regarding the fit of types to the environment is refined over time as new information about

which types have obtained high fitness is encoded into the frequency distribution. We can measure how much a population learns by measuring how much its hypothesis changes during this process. The Kullback-Leibler divergence between the population's initial type distribution $\boldsymbol{p}^0$ and its updated distribution $\boldsymbol{p}^T$ quantifies the amount of **information gain** I^T that selection provides over a learning period of duration T (Figure 2.1a,b,d)

While other processes such as drift, mutation, and migration may change type frequencies, we focus here on changes due to selection. Selection is unique among these processes in that the type frequency changes it enacts are, by definition, explicitly dependent on the environment. Selection establishes an adaptive matching between types and environments such that observing a particular type in a given population reduces uncertainty about the kind of environment it occupies to some extent (Bergstrom and Rosvall 2011, Shea 2007a). In this sense, frequency changes attributable to selection directly contribute to an increase in adaptive information about the environment, whereas any frequency changes attributable to drift or other processes do not (see Chapter 1).

As a population evolves with respect to a particular environment, selection moves the population toward an **evolutionarily stable state** (ESS) $\overline{\boldsymbol{p}}$ that maximizes expected fitness for the current conditions (Maynard Smith 1982). The divergence $D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)$ of the ESS from the population's current type distribution defines the **potential information** of the system: the information gain that is still available from the population's current state (Figure 2.1a). As selection proceeds, information is gained (Figure 2.1d) and the potential information continually decreases (Harper 2009b, Harper and Fryer 2015). In fact, the route that selection takes follows an information gradient, where each update shifts the type distribution in the direction that maximizes information gain relative to the population's current state (Harper 2009b, Harper and Fryer 2015). When the population reaches an equilibrium composition, it has gained all of the information that can be gained about the current environment. Therefore natural selection can be understood at its most fundamental as an information acquisition process, both in its outcomes and in its underlying dynamics.

Fitness loss associated with information gain

Populations do not gain information for free. Just as we learn from our mistakes, populations learn from the shortcomings of poorly adapted types. Differential fitnesses provide the evidence that drives the population's learning, but selection incorporates this information by gradually adjusting frequencies, rather than by instantly fixing highly-fit types. Until less-fit types are eliminated, a population that is still gaining information about the environment has lower mean fitness than a population that already has this information. A population can either grow optimally or acquire information, but not both.

Haldane (1937, 1957) and Kimura (1960) introduced the concept of substitution load to describe the relative cost of using selection to fix beneficial alleles in a population. Haldane and Kimura considered the simple case of an allele substitution in an infinite haploid population with multiple segregating types. The optimal type with the highest fitness increases in frequency toward fixation, but until the substitution is complete the presence of less-fit types holds the mean fitness of the population below the maximum. The **substitution load**

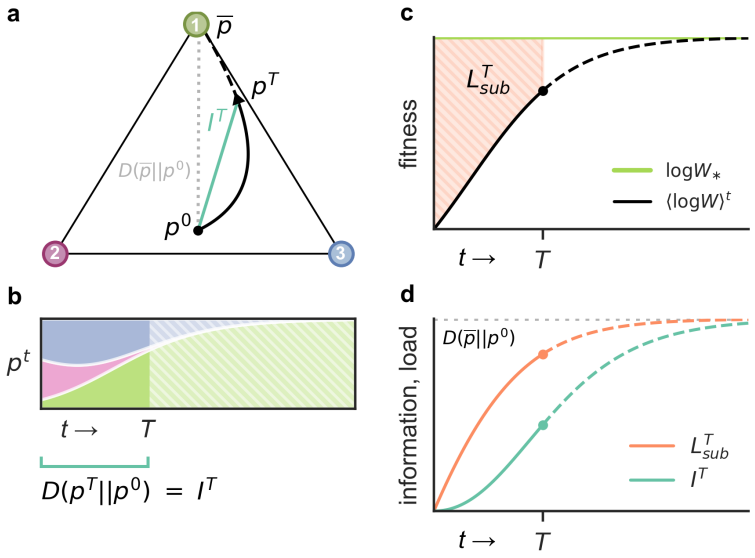


FIGURE 2.1. **Selection accumulates information gain and substitution load.** Natural selection changes the frequencies of types in a population according to their relative fitnesses. **(a,b)** The black trajectory through the simplex in (a) tracks the composition of the population over time as selection moves the population from its initial type frequency distribution $\boldsymbol{p}^0$ toward fixation of the optimal Type 1 (green vertex), which is an evolutionarily stable state (ESS) $\overline{\boldsymbol{p}}$. Change in the population's type frequency distribution over time is also depicted by a Muller plot in (b), where the height of each colored band represents the frequency of the corresponding type at a given time (type colors are indicated in (a)). The population's information gain I^T at time T is measured by the KL divergence of the population's state $\boldsymbol{p}^T$ from its initial state $\boldsymbol{p}^0$. The teal line segment connecting the population's initial state with its trajectory through the simplex provides graphical intuition about how information gain changes over time (note that KL divergence is not a true distance metric). The gray dotted line represents the initial potential information, which is defined as the KL divergence of the ESS from the population's initial state. **(c)** As selection proceeds this fixed environment, the mean fitness of the population increases as suboptimal types decrease in frequency (black line). As the population approaches fixation of Type 1, the mean fitness converges on the optimal fitness of this type (green line). The substitution load at time T measures the cumulative depression in population mean fitness below the optimal fitness level up to that time, which corresponds to the area of the orange shaded region. **(d)** The accumulation of information gain (teal line) and substitution load (orange line) are plotted over the course of selection. Substitution load and information gain converge on the same value as the population approaches fixation (Proposition 1), and substitution load always exceeds information gain (Proposition 2).

L^T_{sub} measures the cumulative difference between the optimal growth rate and the population's average growth rate over T generations of selection (Figure 2.1c,d, Appendix 2.C.1.1)

$$L^T_{\text{sub}} = \sum_{t=1}^{T} \sum_i p_i^t \left(\log W_* - \log W_i\right) \tag{2.2}$$

$$= \sum_{t=1}^{T} \left(\log W_* - \langle \log W \rangle^t\right) , \tag{2.3}$$

where W_* and W_i denote the fitnesses of the optimal and ith types, respectively. The log fitness $\log W_i$ corresponds to the growth rate of the ith type, and $\langle \log W \rangle^t = \sum_i p_i^t \log W_i$ denotes the population's average growth rate (Appendix 2.A.1.1). This quantity gives the total fold loss of potential growth that a population suffers by evolving its composition incrementally relative to a population that grows optimally all along.

Kimura (1961a) noticed an interesting connection between load and information: the total information gain accumulated in a complete allele substitution is equal to the total substitution load incurred during that process. Irrespective of the strength of selection, the information gain and substitution load ultimately converge on the same value, which depends only on the initial frequency of the optimal type p_*^0 (Figure 2.1d, Appendix 2.C.1.1).

Proposition 1. *(Kimura 1961a)*

$$\lim_{T\to\infty} I^T = \lim_{T\to\infty} L_{sub}^T = -\log p_*^0 \,. \tag{2.4}$$

In addition, we prove here that substitution load exceeds information gain throughout the course of selection, with equality in the limit of fixation (Figure 2.1d)

Proposition 2. *(Appendix 2.C.4.3)*

$$I^T \leq L_{sub}^T \;\forall\, T \,. \tag{2.5}$$

This implies that in order to gain a given amount of information, there is a minimum load that must be paid to achieve the type substitutions necessary to encode that information. In other words, the process of selection acquiring one bit of information, which provides a two-fold reduction in uncertainty about the environment, requires at least a two-fold depression in mean fitness along the way.

Measuring load in realistic environments

This theory describes a relationship between two fundamental quantities in evolutionary biology—fitness and information—but its predictions (Proposition 1, Proposition 2) have yet to be measured directly. Here we do so using an experimental evolution system that closely adheres to the assumptions made in Kimura's theory and that enables high-resolution measurement of information gain and substitution load. We used four strains of *Escherichia coli* with distinct mutations in the *rpoB* gene, which encodes the β subunit of RNA polymerase. These *rpoB* alleles confer differing growth rates in a standard growth medium. Each strain was transformed with a plasmid carrying a constitutive fluorescent protein marker for strain identification and cell enumeration. To track the growth of the optimal allele in particular, the strain with the highest fitness was transformed with a GFP marker, while all other strains were transformed with RFP markers.

To test the predictions of Kimura's theory, we ran competition experiments with various combinations of strains starting from different initial frequencies. Each population was incubated in well-mixed rich media for 36 hours, which was sufficient time for all populations to approach fixation. We developed a transfer dilution protocol to ensure that populations remained in exponential growth phase for the duration of the experiment (Materials & Methods). Populations were sampled every four hours, and allele frequencies and population sizes were measured using flow cytometry. Information gain was calculated at each time point as the KL-divergence between the population's current and initial allele frequency distributions. Load was

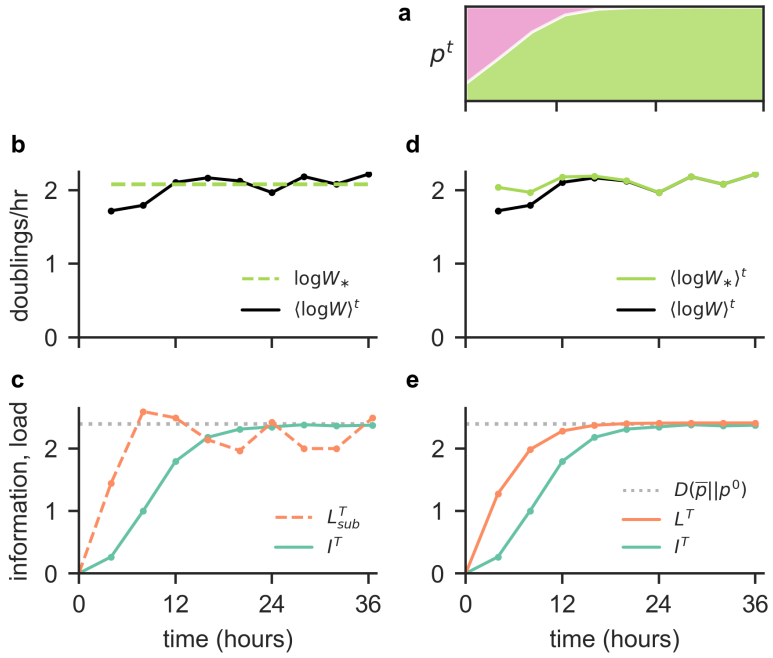


FIGURE 2.2. **Information gain and load measured in a selection experiment.** Results from a single representative selection experiment are presented in detail. **(a)** Changes in the frequencies of genotypes over time are depicted by a Muller plot, where the height of the green and pink bands give the frequency of the GFP-labeled WT strain and the RFP-labeled M1 strain, respectively. The WT strain has the optimal growth rate and selection moves the population toward fixation of this type. **(b)** The mean growth rate of the population (doublings per hour) is plotted over the course of the experiment (black line). Each point represents the estimated average population growth rate over the preceding 4 hour interval. The growth rate of the optimal WT strain was also measured at 4 hour intervals, and the average these measurements is used as a fixed estimate of this type's characteristic growth rate (green dashed line). **(c)** Information gain (teal line) and *substitution* load (orange dashed line) measurements are plotted over time. Here substitution load is computed as the difference between the population growth rate (black line in (b)) and the characteristic growth rate of the optimal type (green dashed line in (b)), as per the classical definition. Under this definition, the observed substitution load does not conform to the predictions of Proposition 1 or Proposition 2. **(d)** In addition to the the estimated mean population growth rates over time (black line, same as in (b)), we also plot the estimated growth rates of the optimal WT strain at each interval (green line). While growth rates fluctuate, the growth rate of the optimal allele always exceeds the mean growth rate, and the population growth rate converges on the optimal rate as it nears fixation of the WT strain. **(e)** Information gain (teal line) and *mismatch* load (orange dashed line) measurements are plotted over time. Using the extended definition of mismatch load, which accommodates time-varying fitnesses, the observed load and information adhere to the the predictions of Proposition 1 or Proposition 2.

quantified as the cumulative difference between the population's average growth rate and the growth rate of the optimal allele in each interval.

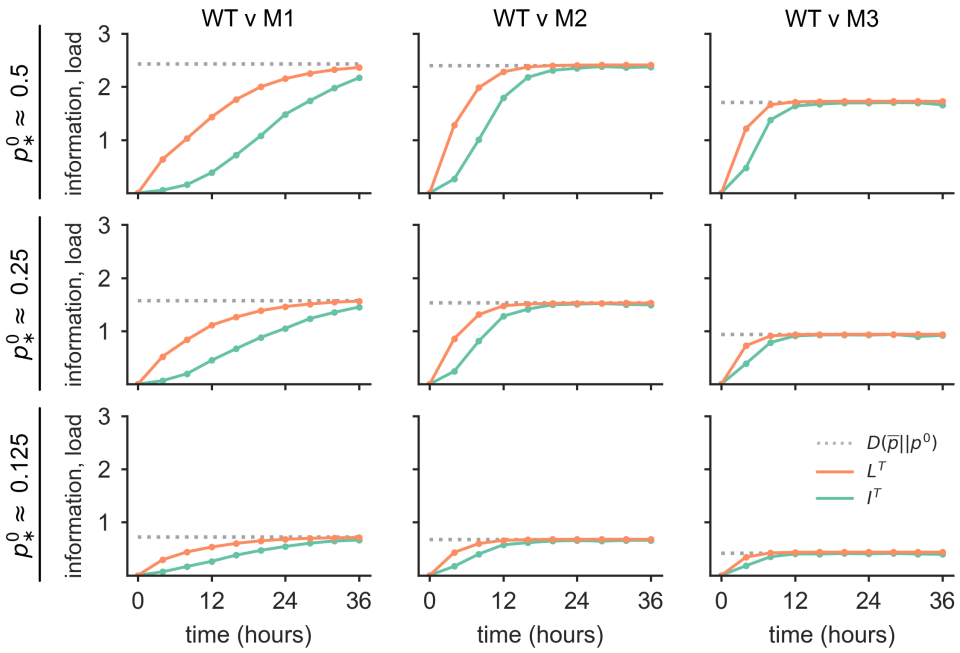


FIGURE 2.3. **The relationship between load and information predicted by theory is reliably observed in selection experiments.** Results from 9 selection competition experiments that involved different combinations of types (column headings) and that initialized batch culture populations with with optimal type at different frequencies p_*^0 (row headings) are shown. The WT strain had the optimal growth rate and approached fixation in all competitions. Theory predicts that information gain and load converge on the same value, equal to the negative log of the initial frequency of the optimal type ($-\log p_*^0$, Proposition 1). The convergence value depends on the initial frequencies of each population, but this convergence is observed in all contexts. Notably, neither the total load nor total information gain depend on the strength of selection—which varies from one group of types to another (Appendix 2.D)—in accordance with theory. Mismatch load is expected to exceed information gain at all times in this setting (Proposition 2), and this relationship is observed in all competitions.

The results of these selection experiments are shown in Figure 2.2, Figure 2.3, and Figure 2.D.8. As expected, the optimal allele increases in frequency and approaches fixation in all competitions. However, the observed relationship between information and load deviates from the theoretical predictions: the substitution load does not always exceed information gain or converge on information gain at fixation (Figure 2.2c). Why not? While Kimura's definition of substitution load assumes a static environment with constant fitnesses, growth rates fluctuated stochastically in our experiments despite maintaining cultures in exponential growth phase. Given that this assumption cannot be satisfied even in a highly controlled laboratory population, applying this fixed definition of substitution load to natural systems is problematic.

However, we can extend this concept of load to realistic contexts where conditions and fitnesses change over time. Suppose there is a set of distinct environmental conditions, which may be characterized by abiotic properties, type frequencies, population densities, or other factors (Appendix 2.A.2). Let the probability that a given organism experiences the jth condition be denoted by x_j, and let the fitness of an organism of the ith type in the jth condition be given by W_{ij}. Then scenarios with fluctuating conditions can be represented with

a probability distribution $\boldsymbol{x}^t$ that may change arbitrarily over time. For this more general context, we define the **mismatch load** (Appendix 2.C.1.2)

$$L^T = \sum_{t=0}^{T} \sum_{i,j} p_i^t x_j^t \left(\log W_{*j} - \log W_{ij}\right) \tag{2.6}$$

$$= \sum_{t=0}^{T} \left(\langle \log W_* \rangle^t - \langle \log W \rangle^t \right) , \tag{2.7}$$

where W_{*j} is the fitness of whichever type is optimal in condition j (i.e., $W_{*j} = \max_i W_{ij}$). Mismatch load measures the the cumulative loss of potential fitness due to the mismatch of types and environmental conditions. Substitution load is a special case of mismatch load for an allele substitution in an environment with only a single, unchanging state.

In our experiments, the environmental conditions (e.g., inoculum size, nutrient concentration, temperature) change slightly from one interval to the next, which causes growth rates to fluctuate, but all cells in each well-mixed batch culture experience the same conditions, and the identity of the optimal type does not change. This still constitutes an allele substitution scenario, albeit a more noisy one, and the theoretical relationship between information gain and load (Proposition 1, Proposition 2) is predicted to hold in such a case. Using the more appropriate definition of mismatch load, our empirical results conform remarkably well to the predictions of theory: information gain is always exceeded by load, and these quantities approach equality as the populations approach fixation (Figure 2.2e). The theory tells us that encoding new information by means of natural selection requires a load of at least the information gain. Because the information gain converges to the mismatch load at fixation, our experimental populations are paying the requisite minimum load necessary for that information, but no more.

The cost of selection as a learning process

Substitution load was originally conceived of and studied as a "cost of selection" (Crow 1970, 1958, Haldane 1937, 1957, Kimura 1968, 1961a,b). Many have since argued that it is counterintuitive to view selection acting on beneficial variation as costly, since populations that evolve are surely better off than those that do not (Brues 1964, Maynard Smith 1968, Moran 1970, O'Donald 1969, Sved 1968, Van Valen 1963). Part of this disagreement stems from a fundamental difference between substitution load and other genetic loads that is obfuscated by their nomenclature. Like other genetic loads, substitution load tracks the fitness penalty that accrues due to the presence of suboptimal types in a population. But whereas mutation load and migration load refer to the deleterious effects of processes that introduce suboptimal types, substitution load refers to a process that removes them. The "cost of selection" that Haldane (1957) and Kimura (1961a) refer to is not the disadvantage of a population that undergoes selection relative to one that does *not* evolve, but rather the disadvantage of a population needing to evolve relative to one that is *already* optimal. As Felsenstein (1971) points out, "it is appropriate to speak of a cost of selection, since the cost comes from the fact that natural selection is less efficient than divine intervention."

That said, it is true that the generalization of substitution load—mismatch load—does not quantify precisely this notion of cost in all cases. Fundamentally, mismatch load measures the loss of potential fitness due to ongoing mismatch between types and environmental conditions. When there is only one environmental

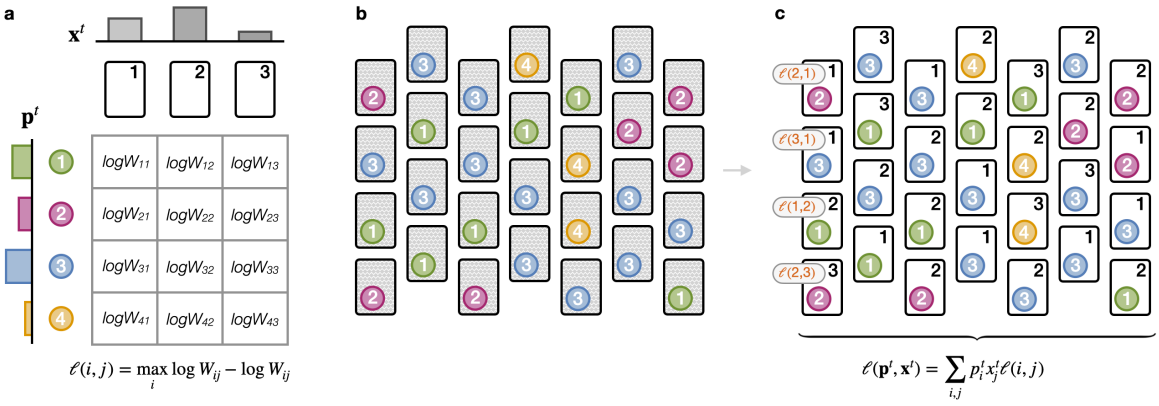


FIGURE 2.4. **The Population versus Environment game.** The learning problem faced by an evolving population can be modeled as a repeated game against the environment, which is illustrated here using the metaphor of a tabletop game. **(a)** The game is defined by the set of types available to the population player (tokens, rows), the set of conditions available to the environment player (cards, columns), and a game matrix $\boldsymbol{G}$ that specifies the payoff that each type receives when played against each environmental condition. Here the payoffs received by individual tokens are defined as the log fitness (growth rate) $\log W_{ij}$ of each type i in each condition j. **(a)** In each round t, the environment player puts out an array of environmental conditions (cards) with condition frequencies given by the environmental strategy $\boldsymbol{x}^t$. The population player selects a collection of types to play against the environment according to their type frequency strategy $\boldsymbol{p}^t$. Each individual type is assigned to an independent environmental condition (one token placed on each condition card), but the pairing of types and conditions is random and out of both players' control (condition cards are shuffled and remain face down until the type tokens are arbitrarily placed). **(c)** The environmental conditions that individuals have been paired with are then revealed. An individual with type i that encountered condition j receives a *loss* $\ell(i,j) = \max_i G_{ij} - G_{ij} = \max_i \log W_{ij} - \log W_{ij}$. The population player's score in round t is the expected (mean) loss $\ell(\boldsymbol{p}^t, \boldsymbol{x}^t)$ of all of the individuals (tokens) they played across the distribution of conditions (cards) played by the environment. The population player's goal is minimize their cumulative expected loss over many rounds of the game.

state (as in Haldane and Kimura's substitution models) or when the same type is optimal in all conditions (as in our experiments), an optimal population is fixed for the most fit type and experiences no mismatch or load. In these cases, all mismatch is attributable to the process of selection being incomplete, and mismatch load quantifies the cost of selection's gradual nature as desired. However, when the environment includes multiple conditions that favor different types, a type distribution that maximizes expected fitness over all conditions may still experience some load due to mismatch resulting from stochastic associations of types and conditions. In this case, not all of the cumulative load is attributable to the gradual nature of selection, and mismatch load does not accurately reflect the cost of the process itself. Instead, it would be more appropriate to quantify the cost of selection as the amount of *excess* load that accrues during evolution relative to the baseline load that is expected of an already adapted population.

A learning theoretic model of evolution. Analyzing the cost and relative performance of learning processes is the purview of computational learning theory. Here we cast evolution by natural selection in a learning

theoretic framework in order to clarify the meaning of mismatch load and develop a rigorous measure for the cost of selection as a learning process.

The general learning problem faced by an evolving population can be modeled as repeated play of a game between two players: the population and the environment (Figure 2.4, Appendix 2.B.2). A play of this **population versus environment** (PvE) game in round t (i.e., generation t) consists of the population 'choosing' a distribution of types and the environment choosing a distribution of environmental conditions to be experienced by the population. The distribution of types $\boldsymbol{p}^t$ and the distribution of conditions $\boldsymbol{x}^t$ represent the strategies of the population player and the environment player, respectively, in round t of the game. A particular learning problem instance can be defined by specifying an n-by-m game matrix $\mathbf{G}$ that gives the log fitness (i.e., growth rate) for each of the n types in each of the m environmental conditions (i.e., $G_{ij} = \log W_{ij}$) and by specifying the process by which the environment updates its strategy $\boldsymbol{x}^t$ over time. A wide range of biological contexts can be modeled with appropriate choices of $\mathbf{G}$ and $\boldsymbol{x}^t$.

Each individual in the population experiences a micro-environment characterized by an independent condition drawn from the environment's distribution of conditions $\boldsymbol{x}^t$. The association of particular micro-environments (conditions) to specific individuals is assumed to be random and out of both players' control. The **loss** of potential long-term fitness that an individual of type i incurs in condition j is defined as the difference in log fitness (i.e., growth rate) between the optimal type in the jth condition and the ith type in the same condition

$$\ell(i,j) = \log W_{*j} - \log W_{ij} \,. \tag{2.8}$$

After each round t, the population player 'observes' the expected loss of each ith type across the distribution of all conditions

$$\ell(i,\boldsymbol{x}^t) = \sum_j x_j^t \left(\log W_{*j} - \log W_{ij}\right) \,, \tag{2.9}$$

which provides evidence about the relative suitability of each type for the current environment. The population player can then use this type loss information to update the type frequencies that define its strategy $\boldsymbol{p}^t$ according to some learning process.

The population incurs the expected fitness loss of its strategy in each round t

$$\ell(\boldsymbol{p}^t,\boldsymbol{x}^t) = \sum_{i,j} p_i^t x_j^t \left(\log W_{*j} - \log W_{ij}\right) \,. \tag{2.10}$$

The **cumulative loss** of potential fitness accrued after T rounds of the population's learning process is equivalent to the mismatch load L^T (Equation 2.7, Appendix 2.C.1.2)

$$\sum_{t=1}^{T} \ell(\boldsymbol{p}^t,\boldsymbol{x}^t) = \sum_{t=1}^{T} \sum_{i,j} p_i^t x_j^t \left(\log W_{*j} - \log W_{ij}\right) = L^T \,. \tag{2.11}$$

This connection reinforces the concept of mismatch load (and substitution load) as the cumulative loss of potential fitness due to mismatch between types and environmental conditions.

Regret measures the cost of learning. The goal of the game for the population is to minimize its regret about unfulfilled fitness when looking back after many generations. In learning theory, **regret** is defined as the difference between the cumulative loss that the population experiences as it updates its strategy $\boldsymbol{p}^t$ over

time and the loss it could have achieved had it played an optimal fixed strategy from the beginning. Regret is the true cost of a learning process: the excess loss one experiences in having to learn an effective strategy as opposed to knowing the optimal solution all along.

In general, the optimal strategy for the population depends on the particular sequence of environments it experiences and is determined in hindsight with knowledge of this sequence. There are multiple ways to define an optimal strategy with respect to a given environmental sequence, and we will examine two such definitions here (Appendix 2.C.2.2).

We are interested in the performance of natural selection as a learning process, and we first consider a notion of optimality that follows from these dynamics. An evolutionarily stable state (ESS) $\overline{\boldsymbol{p}}_{\boldsymbol{x}^t}$ is a type distribution that locally maximizes expected fitness for a particular environment $\boldsymbol{x}^t$. In each update, selection always moves the population in the direction of an ESS for the environment at that time. If the ESS accessible to the population remains constant over a sequence of environments, then selection will continually approach and settle on this **stationary evolutionarily stable state** $\overline{\boldsymbol{p}}$ (i.e., $\overline{\boldsymbol{p}} = \overline{\boldsymbol{p}}_{\boldsymbol{x}^t}$ for all $t \in [0, T]$; Appendix 2.B.3). A stationary ESS is the optimal strategy accessible to the population in the sense that it locally maximizes cumulative expected fitness over the corresponding sequence of environments. The ESS that the population approaches may not be globally optimal, but it is the best strategy that can be reached using replicator dynamics from the population's initial state.

The population will accrue mismatch load as it gradually converges on the optimum. The population's regret $\bar{R}^T$ with respect to the stationary ESS $\overline{\boldsymbol{p}}$ is the amount by which its load L^T exceeds the load $\bar{L}^T$ that it would have experienced had it used the optimal strategy $\overline{\boldsymbol{p}}$ all along over the same sequence of environments.

$$\bar{R}^T = L^T - \bar{L}^T \tag{2.12}$$

$$= \sum_{t=1}^{T} \sum_{i,j} \left(p_i^t - \overline{p}_i\right) x_j^t \left(\log W_{*j} - \log W_{ij}\right) . \tag{2.13}$$

Note that substitution load was defined in a setting where one type always has the highest fitness, and thus the optimal strategy that fixes that type incurs no load (i.e., $\bar{L}^T = 0$ for all T). Therefore, substitution load is equivalent to regret when there is a single universally optimal type.

For a biological population, regret is equal to the cumulative difference in expected growth between a hypothetical population using the optimal strategy $\overline{\boldsymbol{p}}$ and the evolving population using $\boldsymbol{p}^t$

$$\bar{R}^T = \sum_{t=1}^{T} \sum_{ij} \overline{p}_i x_j^t \log W_{ij} - p_i^t x_j^t \log W_{ij} . \tag{2.14}$$

If we consider a lineage to be the collection of individuals derived from an initial population, this quantity can be equivalently expressed as the negative log ratio of the population's total lineage size Γ^T to the size of the optimal lineage $\overline{\Gamma}^T$

$$\bar{R}^T = -\log\left(\frac{\Gamma^T}{\overline{\Gamma}^T}\right) . \tag{2.15}$$

Just as relative fitness can refer to the short-term reproductive output of a type relative to the maximum output of other types, we can interpret the ratio $\Gamma^T/\overline{\Gamma}^T$ as the **relative lineage fitness**: the cumulative growth of the evolving lineage relative to the lineage that maximizes cumulative growth for the given sequence of

environments. Therefore, the population's goal in the game can be interpreted as seeking to learn a type distribution that approximates the optimal strategy in order to maximize its relative lineage fitness and thus minimize its regret.

Natural selection as a no-regret learning process.

Intuitively, it would be reasonable for the population player to adapt their strategy by increasing the weight of types that have done well (i.e., observed low fitness loss) in the past. The learning algorithm known as Follow the Leader (FTL) is an extreme implementation of this heuristic, which places all of the weight in each round on the single type that has received the lowest cumulative loss against the previous conditions of the environment (Appendix 2.B.2.3). The FTL algorithm can achieve very low loss when the environment is stable, but it breaks down when the population faces a dynamic environment, because there is no guarantee that a strategy that fared well in past observations will continue to do so against future environments. A more robust learning heuristic is to shift strategies in a gradual fashion that balances minimizing loss based on prior losses with maintaining diversity over types to hedge against the future.

Computer scientists and economists have shown that a simple, yet powerful, learning algorithm known as **Multiplicative Weights Updating** (MWU) is an effective solution for online learning in settings where the environment can change arbitrarily over time (Arora et al. 2012, Cesa-Bianchi and Lugosi 2006, Freund and Schapire 1999). Effectively, MWU balances concentrating weight on types that have performed well (i.e., incurred low loss) in the past against spreading weight over types that may perform well in the future (Appendix 2.B.2.3). A learning rate parameter modulates the emphasis given to reacting to loss versus maintaining spread in each update. When the learning rate can be tuned in response to the learning problem, MWU is guaranteed to learn a strategy that converges on the performance of the optimal strategy in hindsight, which ensures that the learner's per-round regret will approach zero in the long run (Cesa-Bianchi and Lugosi 2006, Freund and Schapire 1999).

Recent work has shown that replicator dynamics is equivalent to a fitness-based implementation of MWU with a constant learning rate that equally balances minimizing fitness loss (maximizing growth) and maintaining diversity (Chastain et al. 2014, Chastain 2017, Mehta et al. 2015, Meir and Parkes 2015)(Appendix 2.B.2.4). Because the learning rate implicit in replicator dynamics is not tuned, the no-regret guarantees from the analysis of MWU do not automatically carry over to selection. Here we apply a form of amortized analysis common in learning theory to establish upper bounds on load and regret for natural selection (Appendix 2.C.3).

When the sequence of environments is characterized by a stationary optimal strategy $\overline{\boldsymbol{p}}$, the total load for a population undergoing selection is bounded.

Proposition 3. *(Appendix 2.C.3.1)*

$$L^T \leq \bar{L}^T + D(\overline{\boldsymbol{p}} \big|\big| \boldsymbol{p}^0) \quad \forall\, T\,. \tag{2.16}$$

This bound guarantees that an evolving population's load will be no greater than the load of the optimal strategy plus the divergence between the population's initial type distribution $\boldsymbol{p}^0$ and the optimal composition $\overline{\boldsymbol{p}}$. The divergence term represents how much learning the population has to do at the outset, which translates

to excess load—regret—that accrues while the population is shifting its strategy toward the optimal. In fact, this divergence is an upper bound on regret with respect to a stationary ESS.

Theorem 1. *(Appendix 2.C.3.1) For any game matrix $\boldsymbol{G}$ and for any sequence of environmental conditions $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$ such that the population's initial type distribution $\boldsymbol{p}^0$ is in the basin of attraction of an evolutionary stable state $\overline{\boldsymbol{p}}$ that remains stationary for all $t \in [0, T]$, the total regret $\bar{R}^T$ with respect to $\overline{\boldsymbol{p}}$ of the trajectory of type distributions $\boldsymbol{p}^0, \ldots, \boldsymbol{p}^T$ generated by replicator dynamics is bounded from above at all times by*

$$\bar{R}^T \leq D(\overline{\boldsymbol{p}} || \boldsymbol{p}^0) \quad \forall\, T\,, \tag{2.17}$$

with equality as $T \to \infty$

$$\lim_{T \to \infty} \bar{R}^T = D(\overline{\boldsymbol{p}} || \boldsymbol{p}^0)\,. \tag{2.18}$$

In other words, the total cost of selection with respect to a particular ESS is equal to the amount of learning that the population must do to arrive there. The total regret is finite, which guarantees that the per-round regret of selection approaches zero in the long term (i.e., $\bar{R}^T/T \to 0$ as $T \to \infty$). Therefore, selection is a no-regret algorithm with respect to a fixed learning target, and selection is guaranteed to arrive at a type distribution that maximizes relative lineage fitness in this setting. This means that evolution by natural selection is an asymptotically optimal solution to the basic learning problem faced by evolving populations.

Learning problems faced by evolving populations.

The flexibility of this learning theoretic view of selection and its implications can be better understood by considering a few concrete cases.

Constant environments. First, let us revisit the simple case of a constant environment (i.e., $\boldsymbol{x}^t = \boldsymbol{x}^{t+1}$ for all t). When the makeup of the environment does not change, a single type will be the best option in every round. Selection will continually approach fixation of this optimal type, which is a stationary ESS. We know from Theorem 1 that a population that evolves by natural selection will achieve vanishing per-round regret in learning this optimal strategy.

In a static environment, achieving the goal of minimizing regret is equivalent to maximizing fitness in an absolute sense. Haldane and Kimura considered the simplest possible environment with only one unchanging condition (i.e., $m = 1$). In this case, selection increases the frequency of the type that has the highest fitness in this condition, which maximizes the mean fitness of the population as described by Fisher's fundamental theorem (Fisher 1930). Similarly, if the environment consists of multiple conditions (i.e., $m > 1$) but the frequency of these conditions does not change, then selection will consistently favor the single type with the highest expected fitness over the constant distribution of conditions (Figure 2.5, left), which maximizes the geometric mean fitness in keeping with Gillespie's geometric mean principle (Gillespie 1974).

When the same type always has the highest fitness, an FTL learner will fix the optimal type after a single round, and its only regret will be the excess loss observed for the initial type distribution. Replicator dynamics will not be as fast, but selection will still converge on the optimal type. In the special case of a constant environment, a population that learns using selection will have lower relative lineage fitness and accrue more load than one using FTL, but this "inefficiency" is bounded. Specifically, Theorem 1 assures that the population's regret is bounded by the divergence between its initial composition and the optimal strategy.

Frequency dependence. Constant environments are only a small special case of what the learning theoretic view of evolution can express. In fact, there are no restrictions on the how the environment's distribution of conditions can change over time. For example, the environment's strategy in each round may depend on the population's current or prior strategies, which allows for cases of frequency-dependent selection. In particular, if the frequencies of conditions that individuals experience are set by the frequencies of types in the population (i.e., $m = n$, $\boldsymbol{x}^t = \boldsymbol{p}^t$), then the make up of the "environment" is defined by the composition of the population itself, and fitnesses reflect interactions between types. This context is of interest to evolutionarily biology as it opens up new kinds of dynamics (e.g., social dilemmas) and is known to have a transformative effect on the adaptive capabilities of evolving populations (Kaznatcheev 2020, 2017).

The frequency-dependent setting reveals why the goal of the game is to maximize relative lineage fitness (i.e., minimize regret) rather than to maximize fitness outright. To illustrate this, consider a specific learning problem corresponding to the Prisoner's Dilemma (note that this is a scenario where $\boldsymbol{x}^t = \boldsymbol{p}^t$). In the Prisoner's Dilemma the best outcome occurs if the population plays an all-cooperate strategy, which achieves the maximum possible population mean fitness. However, replicator dynamics leads toward a population of all defectors, which continually reduces the population's mean fitness and clearly does not maximize fitness outright.

Recall that the only information the population player receives is the average fitness loss of each type in each round of the game. At each step, the population observes that no matter the environmental distribution the defector type always has lower expected loss (higher average fitness) than the cooperator. Based on this evidence, the defector type is seen to be better off in every round, and replicator dynamics increases its frequency accordingly. Arriving at an all-cooperate strategy would require the population player to use information that it is not available to it (e.g., knowledge of $\mathbf{G}$ or knowledge of the environment's process for updating $\boldsymbol{x}^t$) or to adopt a learning algorithm that is not generally robust across learning problems (e.g., up-weighting types with high relative loss). Thus when facing some environments, the population may be simply unable to maximize fitness in absolute terms by using an online learning process.

In general, the attainable goal for the population is to maximize its cumulative fitness relative to the maximum possible fitness that could be achieved by a fixed strategy given the same sequence of environments. In the Prisoner's Dilemma setting, if the observed sequence of environments is fixed, then the population would have been best off playing all-defect from the beginning. The all-defect strategy is a stationary ESS, and the population does indeed learn this locally optimal strategy and minimize regret using replicator dynamics (Theorem 1). Selection does not maximize the absolute fitness of the population, but it does maximize the relative lineage fitness with respect to the cumulative growth of the all-defect strategy in the same sequence of environments.

Evaluating the performance of a learning process with respect to the fixed sequence of environments in hindsight may seem strange, particularly when that sequence of environments was dependent on the trajectory of the population while the learning was underway. Reframing regret minimization in terms of relative lineage fitness maximization lends intuition to the fixed retrospective nature of this quantity. For example, consider a large, well-mixed population of cells that evolves according to replicator dynamics in the Prisoner's Dilemma scenario. Suppose we identify two very small subpopulations: one that has the same initial type distribution as the overall population and one that is initially all defectors. In the first subpopulation, replicator dynamics increases the frequency of defectors and tracks the change in the overall population. The second subpopulation

remains fixed for defectors, which turns out to be optimal. Given that these subpopulations are small relative to the overall population, their trajectories do not affect the makeup of the overall population that constitutes the "environment." The two subpopulations therefore experience the same sequence of environments, and it is natural to ask how the cumulative growth of these two lineages compare given the conditions that play out. Relative lineage fitness quantifies this comparison. The more effective the learning process used by the first subpopulation, the greater its relative lineage fitness and the less regret it experiences. Just as the growth of learning subpopulation can be compared to a fixed subpopulation in the same physical environment, the cost of a learning process is measured with respect to the fixed optimal strategy in hindsight, whether or not such an optimal subpopulation actually existed.

Arbitrary environmental change. In this section, we extend the concept of regret to the fully general setting where no assumptions are made about the environment. Regret measures the cumulative fitness loss of an evolving population with respect to a fixed optimal strategy for the same sequence of environments. So far, we have considered the fixed optimal strategy to be an ESS that is stationary over the observed sequence of environments. However, when the distribution of environmental conditions $\boldsymbol{x}^t$ is free to fluctuate arbitrarily, shifts in the environment may cause the evolutionarily stable states to change. Whereas an environmental sequence with a stationary ESS represents a basic learning problem with a fixed solution, environments where the stable state switches from one interval to the next can be seen as a series of problems that the population must integrate its learning over.

If we make no restrictions on the sequence of environments, then we cannot determine a fixed optimal strategy based on the process that generates the environmental sequence or on the dynamical trajectory of the population, since these are free to change at any time. Instead, we consider the **empirically optimal strategy**

$$\widetilde{\boldsymbol{p}}^T = \arg\min_{\boldsymbol{q}} \sum_{t=1}^{T} \sum_{i,j} q_i x_j^t \left(\log W_{*j} - \log W_{ij}\right) , \tag{2.19}$$

which is the fixed type distribution that would have minimized cumulative loss (Equation 2.11) for the observed sequence of environments in hindsight after T rounds. In the learning problems that we have considered so far (i.e., constant environments and the Prisoner's Dilemma) the empirically optimal strategy is always a stationary ESS as well (i.e., $\exists\ \overline{\boldsymbol{p}}$ and $\overline{\boldsymbol{p}} = \widetilde{\boldsymbol{p}}^T\ \forall\ T$), but this is not always the case. For example, in a Hawk-Dove game replicator dynamics carries the population to a polymorphic ESS, while the empirically optimal strategy is to play all-defect, which globally minimizes loss (maximizes fitness) for the given sequence of environments. In scenarios like the Rock-Paper-Scissors and cyclical environment games shown in Figure 2.5, the strategy that is empirically optimal in hindsight may change in response to environmental shifts as the game progresses.

We refer to regret measured with respect to the empirically optimal strategy $\widetilde{\boldsymbol{p}}^T$ as **empirical regret**

$$\widetilde{R}^T = L^T - \widetilde{L}^T \tag{2.20}$$

$$= \sum_{t=1}^{T} \sum_{i,j} \left(p_i^t - \widetilde{p}_i^T\right) x_j^t \left(\log W_{*j} - \log W_{ij}\right) , \tag{2.21}$$

where $\widetilde{L}^T$ is the cumulative loss that would have been achieved using $\widetilde{\boldsymbol{p}}^T$ all along in the observed sequence of environments.

In general, the environment may fluctuate in such a way that the evolutionarily stable states change as time goes on. The accumulation of excess load while learning strategies that turn out to be suboptimal later on increases the cost of the overall learning process. Selection is a no-regret algorithm with respect to fixed learning targets, but the regret bound from Theorem 1 does not apply when the ESS accessible to the population is not stationary.

Nevertheless, results from analysis of the MWU algorithm provide an upper bound on the cumulative loss—mismatch load—for selection in the general setting.

Proposition 4. *(Freund and Schapire 1999)*

$$L^T \leq \frac{e}{e-1}\left(\widetilde{L}^T + D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0)\right) \; \forall \, T \, . \tag{2.22}$$

This bound is similar to the bound on load in the setting of a stationary ESS (Proposition 3), but is worse by a constant factor (where e denotes Euler's number) that represents the additional cost of the population being caught out and having to change direction due to changes in the learning target. Even still, this bound guarantees that the mismatch load for an evolving population in any environment will not be much more than the minimum possible load $\widetilde{L}^T$ in the long run.

This result leads to a corresponding upper bound on the population's empirical regret in any environment.

Theorem 2. *(Appendix 2.C.3.2) For any game matrix $\boldsymbol{G}$ and for any sequence of environmental conditions $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$, the total empirical regret $\widetilde{R}^T$ with respect to $\widetilde{\boldsymbol{p}}^T$ of the trajectory of type distributions $\boldsymbol{p}^0, \ldots, \boldsymbol{p}^T$ generated by replicator dynamics is bounded from above at all times by*

$$\widetilde{R}^T \leq \frac{1}{e-1}\left(\widetilde{L}^T + eD(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0)\right) \; \forall \, T \, . \tag{2.23}$$

In the fully general case, the population's regret is not bounded by a finite value and may continue to increase over time. The gradual nature of selection causes the population to accrue regret whenever it must move to a new ESS. If environmental fluctuations cause the ESS to change such that that the population cannot settle into an optimum for the long haul, then the population will continue to accumulate regret as it adapts to each new learning target. This general regret bound makes no assumptions whatsoever about the types, conditions, fitnesses, or sequence of environments that constitute the learning problem. Therefore, this bound provides an upper limit on the total empirical regret a population can possibly experience, even in a worst case scenario.

It is standard practice in computer science to evaluate the relative performance of an algorithm in terms of its worst-case performance. The worst case scenario for an evolving population is an environment that selects its distribution of conditions adversarially—that is, with knowledge of the population's strategy and an intent to inflict maximal losses. Theorem 2 provides a benchmark for comparing the worst-case performance of selection to other learning processes. While learning processes such as FTL that react more strongly than selection to each round of observed losses can achieve substantially lower regret than selection in simple environments, selection's worst-case regret is bounded lower than that of more aggressive learning processes Adversarial or not, environments that are highly variable or unpredictable can drive highly reactive learning processes like FTL to high regret. However, the balance that selection strikes in minimizing loss based on previous observations and maintaining diversity to hedge against the future make it more robust to arbitrary environmental change. Theorem 2 establishes the worst-case superiority of selection over more aggressive

processes, but delineating the classes of non-adversarial environments where selection typically outperforms other learning processes is an area for future research.

At first, an environment that responds adversarially to a population might seem like something that would only exist in a theorist's imagination. But on closer inspection, both physicians and the immune system play an adversarial role in responding to infection. Consider the context of evolutionary medicine concerned with mitigating a burden population of viruses, microbes, or cancer cells that are evolving within the patient. From our perspective, the burden population is the evolving player and the physician is the adversarial environment that administers treatment conditions that affect the burden's fitness. The physician is often seen as having an advantage due to being able to plan rationally while the burden population can only react to its past circumstances. But, as we have seen, there is a fundamental limit on how much regret can be inflicted on a population that adapts using natural selection. This can help explain why certain burdens such as cancers frequently escape treatment and should be a lesson for us to not underestimate the effectiveness of the burden's response to the physician.

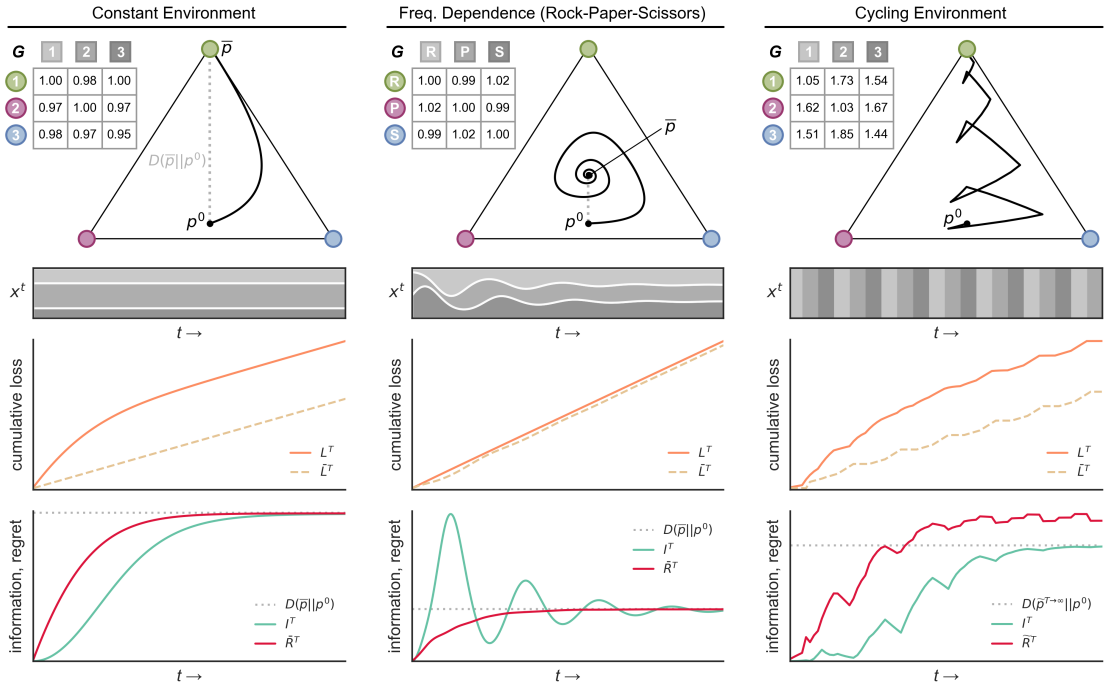


FIGURE 2.5. **Load, regret, and information gain in different environmental contexts.** Results from simulations of replicator dynamics in three different contexts (columns) are shown. In each column, the matrix $\boldsymbol{G}$ defines the log fitnesses (growth rates) for 3 types in each of 3 environmental conditions. Selection updates the type composition of the population along the black trajectory in the simplex. Gray dotted lines in the simplex depict the KL divergence $D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)$ of a stationary ESS $\overline{\boldsymbol{p}}$ from the population's initial state $\boldsymbol{p}^0$, where applicable. Muller plots show the distribution of environmental conditions $\boldsymbol{x}^t$ over time. The mismatch load of the evolving population (L^T, orange line) and of either the ESS composition (L^T, gold dashed line, left and center columns) or the empirically optimal composition ($\widetilde{L}^T$, gold dashed line, right column) are plotted over time. The bottom-most plot in each column gives the information gain (teal line) and either the ESS regret ($\bar{R}^T$, red line, left and center columns) or empirical regret ($\widetilde{R}^T$, red line, right column) over the course of selection. **(Constant environment, left)** The environment is heterogeneous but the distribution of conditions is constant. Type 1 (green) has the highest expected fitness and increases in frequency in every generation. However, Type 1 incurs non-zero loss in Condition 2, so the optimal composition that is fixed for Type 1 accrues load over time (gold dashed line). Fixation of Type 1 is a stationary ESS, so regret is bounded by the initial potential information $D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)$. Information gain and regret converge on this value in the long-run. **(Frequency-dependent selection, center)** Here the environmental conditions are defined by the types themselves, and the distribution of conditions is set to the type distribution in every generation (i.e., $\boldsymbol{x}^t = \boldsymbol{p}^t$). This frequency dependence and choice of $\boldsymbol{G}$ define a "Rock-Papers-Scissors" scenario, where each type ('R', 'P', or 'S', respectively) is advantaged over one type and disadvantaged to the other type. Selection leads to a dampened oscillation in type frequencies that converges on a mixed stationary ESS. The existence of a stationary ESS implies that regret converges on the initial potential information $D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)$. Information gain fluctuates before converging on the regret and initial potential information. **(Cycling environment, right)** The environment consists of a single condition that changes cyclically over time. The ESS changes with each environmental shift, and selection moves the population in different directions in response to each condition. Type 1 has the highest expected fitness over the entire cycle and tends to increase in frequency. The empirically optimal strategy initially alternates in response to the environmental cycle before settling on the Type 1 fixation composition for the long-run. In this example, selection causes the population to converge toward the empirically optimal strategy, so the empirical regret is finite. The empirical regret exceeds the information gain at all times.

The cost of information acquisition by natural selection.

Populations that evolve by natural selection accumulate both information and load as a result of this learning process. Kimura noted a proportional relationship between information and load in the special case of an allele substitution (Proposition 1), and we observe that the accumulation of load outpaces the acquisition of information for experimental populations in such cases (Proposition 2, Figure 2.2, Figure 2.3). This suggests that there is a fitness cost associated with acquiring information using natural selection. By embracing a learning theoretic view of evolution, we have formalized and generalized the cost of selection in terms of regret, and we can now investigate the the cost of information acquisition by natural selection more generally.

The relative fitnesses of types provide evidence about the relative suitability of types for the current environment. In the context of the population versus environment game, we say that the population player receives information about the average fitness loss of each type. Selection changes the frequencies of types in such a way that this relative suitability information is encoded into the population's type frequency distribution. The following result relates the information gained by selection in a single generation (as measured by the divergence $D(\boldsymbol{p}^{t+1}||\boldsymbol{p}^t)$) to the expected fitness loss due to type mismatch in the current and next generations.

Proposition 5. *(Appendix 2.C.4.1)*

$$D(\boldsymbol{p}^{t+1}||\boldsymbol{p}^t) \leq \ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{p}^{t+1}, \boldsymbol{x}^t) \tag{2.24}$$

$$= \sum_{i,j} \left(p_i^t - p_i^{t+1}\right) x_j^t \left(\log W_{*j} - \log W_{ij}\right) . \tag{2.25}$$

The fact that KL divergence is always non-negative reinforces that selection always updates the population such that the next generation has less expected loss in the environment experienced by the parent generation. The fold reduction in fitness loss conferred by a single selective update is always greater than or equal to the information gained in that step. We can view the loss difference in Proposition 5 as the single-step regret that the population has about not already having the composition enjoyed by the next generation. If the population were to have no regret relative to the next generation, then it would have nothing to gain—fitness or information—in changing its composition.

The population's cumulative information gain I^T after T generations is measured by the divergence $D(\boldsymbol{p}^T||\boldsymbol{p}^0)$ of its evolved type distribution from its initial composition (**??**). The total amount of information that can be gained for a particular ESS $\overline{\boldsymbol{p}}$ is given by the initial potential information $D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)$. So long as the ESS remains stationary, the population will eventually converge on the ESS and acquire all of the potential information (i.e., $I^T \to D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)$ as $T \to \infty$). We also recognize the initial potential information as the same divergence that bounds the population's regret with respect to a stationary ESS (Theorem 1). Therefore, we recover Kimura's finding that the cost of selection is proportional to the amount of information the population lacks out the outset (i.e., Proposition 1 is a special case of Theorem 3).

Theorem 3. *(Appendix 2.C.4.2) For any game matrix $\boldsymbol{G}$ and for any sequence of environmental conditions $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$ such that the population's initial type distribution $\boldsymbol{p}^0$ is in the basin of attraction of an evolutionary stable state $\overline{\boldsymbol{p}}$ that remains stationary for all $t \in [0, T]$, the total information gain I^T and the total regret $\bar{R}^T$ with respect to $\overline{\boldsymbol{p}}$ of the trajectory of type distributions $\boldsymbol{p}^0, \ldots, \boldsymbol{p}^T$ generated by replicator dynamics both converge on the value of the initial*

potential information

$$\lim_{T\to\infty} I^T = \lim_{T\to\infty} \bar{R}^T = D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)\,. \tag{2.26}$$

It follows that the cost of each bit of information is 1 fold reduction in relative lineage fitness relative to the optimal in the long run.

Corollary 3.1.

$$\lim_{T\to\infty} \frac{\bar{R}^T}{I^T} \to 1\,. \tag{2.27}$$

While a population that evolves by natural selection may experience more regret than it would using some other learning processes in the context of a constant learning target, all of the excess fitness loss information it receives is translated into information encoded in its distribution of types. The acquisition of this information maximizes the relative lineage fitness of the population to the point that it no longer experiences regret on a per-generation basis.

When the environment is variable, the population's information gain may rise and fall over time as it adapts to changing conditions. We can interpret these fluctuations as the population gaining information about the present environment and "unlearning" information it had acquired about spurious trends in the past. For example, consider the Rock-Paper-Scissors scenario shown in Figure 2.5 (center). In response to conditions that initially favor the S type, the population "overlearns" a S-dominant strategy, and the information gain actually exceeds the initial potential information as the population overshoots the ESS composition that is optimal for the long term. Ongoing shifts in the frequency-dependent environment lead to oscillations in information gain as the population cyclically learns and unlearns strategies dominated by each type in turn. Eventually the population learns a balanced optimal strategy as it settles into the stationary ESS, and the total information gain converges on the initial potential information.

Despite possible fluctuations in information gain, selection continually makes progress in learning about the current ESS. That is, the potential information $D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)$ that remains from the population's current state $\boldsymbol{p}^t$ decreases monotonically over intervals where the ESS $\overline{\boldsymbol{p}}$ remains constant.

In all cases, we find that the information gain is bounded by the empirical regret at all times

Theorem 4. *(Appendix 2.C.4.3) For any game matrix $\boldsymbol{G}$ and for any sequence of environmental conditions $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$, the total information gain I^T of the trajectory of type distributions $\boldsymbol{p}^0, \ldots, \boldsymbol{p}^T$ generated by replicator dynamics is bounded at all times by the total empirical regret $\widetilde{R}^T$ with respect to the empirically optimal strategy $\widetilde{\boldsymbol{p}}^T$*

$$I^T \leq \widetilde{R}^T \;\forall\, T\,. \tag{2.28}$$

For a biological population that evolves by natural selection, every bit of information about the environment requires at least a commensurate reduction in relative lineage fitness. This result holds for any evolutionary learning problem (i.e., any choice of $\mathbf{G}$ and sequence of $\boldsymbol{x}^t$) and confirms that there is a minimum fitness cost of information for natural selection in all contexts.

In general, when the environment changes arbitrarily the population pays a higher price for information. We have seen that the upper bound on regret is higher in environments for which the optimal strategy in hindsight changes over time due to additional load associated with the population learning evolutionarily stable states that are not ultimately optimal in the long run (Theorem 2). In environments that fluctuate

with a pattern that can be learned, selection will eventually settle on a strategy that is optimal for that central tendency and achieve vanishing per-round regret. In such a case, the cost per bit converges on a finite value in the long run. When the environment varies unpredictably or adversarially, regret and the cost per bit may not converge. In any case, the regret cost per bit can be bounded, and the maximum cost of information for selection is lower than that for learning processes that respond more aggressively to changes in the environment, such as FTL.

Discussion

These results highlight the intimate relationship between fitness and information. When a population has types with suboptimal fitnesses for the current conditions, natural incorporates this information about the environment by gradually adapting its composition. If the population does not suffer suboptimal fitnesses then there is nothing it needs to learn. Naturally, then, the accumulation of load and regret must precede and outpace the acquisition of information, and selection effectively converts observed fitness losses to information gain. As a population gains information it arrives at a more adaptive matching of types to environmental conditions and achieves higher relative lineage fitness than a population without this information. Potential information left to learn is unactualized fitness still to gain, and information not yet held is potential growth forsaken.

A learning theoretic view of evolution establishes regret, a measure of relative load, as the appropriate definition of the cost of selection. In environments where the learning target is fixed (i.e., where there is a stationary ESS), populations that evolve by selection reliably learn a strategy that maximizes relative lineage fitness and achieves vanishing per-generation regret. In general, selection balances load minimization and diversity maximization, which enables the process to fare well in complex environments.

These results generalize the concept of substitution load and breathe new intuition into the concept of "the cost of selection." In addition, we revive, formalize, and extend Kimura's initial insight that there is fundamental fitness cost for every bit of information that is acquired by selection. That said, the price of information is better for selection than for some alternative learning processes in general.

Kimura's interpretation of this cost as a limit on the rate by which adaptive genetic information can be accumulated was instrumental in the development of the neutral theory (Kimura 1968, 1961a, 1983). Others have since challenged the notion that there is a bound on the amount of information that can be acquired when multiple loci are considered (O'Donald 1969, Sved 1968, Van Valen 1963). While there is a minimum regret associated with every bit of information gain at a single locus, the total regret across loci does not scale additively due to epistasis and genetic hitchhiking. However, our results hold when types are defined such that type fitnesses integrate the epistatic fitness effects of multiple loci. Selection acting on multiple loci simultaneously is expected to induce regret that exceeds the total information gain with respect to multi-loci genotypes, as our experimental results show, but disentangling the contributions of individual loci to the total load and information gain is indeed more complicated.

Maynard Smith (1968) suggested that there is "no necessary limit on the number of genes which can be selected for simultaneously" under truncation selection. Maynard Smith "assumes random association of alleles at different loci, which requires a large amount of recombination," which is out of the scope of our model. Indeed, sexual reproduction and recombination may allow a population to accumulate information at

an arbitrarily high rate that exceeds the regret bounds discussed here. Determining the extent to which sexual reproduction impacts the rate of information acquisition and whether any such effects have advantages at the level of the individual is an interesting direction for future research.

Here we have focused on natural selection as a process that acquires adaptive information about standing variation. However, selection interacts with other forces in biological evolution. Notably, mutation introduces new types and expands the set of compositions that selection can learn about. Together mutation and selection move populations across adaptive landscapes, but finding optima in complex landscapes is a computationally hard search problem (Kaznatcheev 2019). Reconciling the effectiveness of selection as a learning process with the complexity of evolutionary search is important to understanding the performance of Darwinian evolution in an algorithmic sense.

We see continued integration of evolutionary theory, information theory, and learning theory leading to a richer understanding of adaptive evolution.

Methods

Bacterial strains. *Escherichia coli* B (REL606) were used in selection experiments and related assays. A "wild type" strain (WT) and three strains with unique mutations in the *rpoB* gene (M1, M2, M3) were obtained with permission from the $-80°$C strain archive from Lindsey et al. (2013). Mutations to the *rpoB* gene conferred each mutant strain with a distinct exponential growth rate that was reduced from that of the WT strain (Figure 2.D.6). A full description of all strains used in this study can be found in Appendix Appendix 2.D.

Plasmids. Each strain was transformed with an engineered marker plasmid. The pBR322 plasmid, which carries a *bla* gene conferring ampicillin resistance (Amp^R) and a *tetA* gene conferring tetracycline resistance (Tet^R), served as a vector backbone. Using Gibson assembly, the *bla* gene and corresponding promoter region was removed and replaced with an insert carrying a fluorescent protein gene under a strong constitutive *proC* promoter. As the strain with the optimal growth rate, the WT strain was transformed with a plasmid engineered to carry the green fluorescent protein (GFP) gene *mGFPmut2*. The *rpoB* mutant strains were each transformed with plasmids engineered to carry the red fluorescent protein (RFP) gene *mScarlet-I*. Additional details and plasmid schematics can be found in Appendix Appendix 2.D and Figure 2.D.7.

Growing and storing bacteria. Strains were cultured in standard Luria-Bertani (LB) broth with with 15μg/mL tetracycline for plasmid retention (hereafter referred to as "media"). A single large batch of media was prepared and used for all cultures across all stages of all experiments. Cultures were incubated in 25 mL of media in 125 mL capacity flasks at 37°C with shaking. Strains and culture samples were stored by mixing 1 ml of culture with 160 μl of 80% glycerol and freezing at $-80°$C. Strains were revived from freezer scrapings and incubated overnight prior to all assays and experiments.

Selection experiments. Selection competition experiments were conducted between three pairs of strains: WT vs. M1, WT vs. M2, and WT vs. M3. Each pair of strains participated in three competitions with different initial strain frequency compositions: WT at approximately 12.5%, 25%, and 50% of the initial population, respectively. All competitions were conducted simultaneously using the same incubator in order to mitigate potential day or block effects. Strains were revived and incubated overnight to saturation density

(~1×10^9 cfus/mL). Prior to the start of competitions, individual strain cultures were diluted 1000-fold into 25 mL of fresh media (~1×10^6 cfus/mL) and incubated for 1 hour, which was sufficient time for cells to begin resuming exponential growth. At the end of this "reanimation" period, the density of each strain culture was spot-checked using flow cytometry. Selection experiments were initiated by combining the strains to be competed in the designated ratios at a density of ~1×10^6 cfus/mL in 25 mL of fresh media. (Because the strains had entered exponential growth phase prior to initiating competitions, some growth occurred during the initiation process that caused minor imprecisions in the initial strain frequencies.) Every 4 hours, a 1 mL sample was taken from each competition culture to measure strain densities and frequencies using flow cytometry (see Flow cytometry section below for more information). At the same time, a sample of the culture was transferred to 25 mL of fresh media. An algorithmic process was used to determine the transfer volume for each competition culture at each transfer in order to ensure that the culture would remain in exponential growth while maintaining a measurable density for flow cytometry measurements in the next growth interval (see Appendix 2.D.1.1 for more information). Competitions proceeded in this fashion for 36 hours.

Flow cytometry. Flow cytometry was used to enumerate cells in culture samples and estimate culture densities. Fixed fluorescence gates were established to separately enumerate cells expressing GFP and RFP markers (see **??**). Culture samples were immediately centrifuged at 13,000 rpm for 5 minutes to remove growth media supernatant. Pelleted cells were then resuspended in 1 mL of flow buffer (1X PBST with EDTA added at a 500-fold dilution). Samples were then transferred to a round-bottom 96-well plate where a 10-fold dilution series was performed. The accuracy of cell enumeration was maximized when the sample density was in the range of ~$1 \times 10^4 - 1 \times 10^6$ cfus/mL, and samples were measured at 10-fold, 100-fold, and 1000-fold dilutions to ensure each sample had measurements in this range. Three replicate flow cytometer enumerations were performed for each sample at each dilution.

Appendix 2.A. *Modeling natural selection.*

2.A.1. The replicator dynamics model of natural selection. Replicator dynamics is a standard model of evolution that describes the action of natural selection in large populations of replicators without mutation, recombination, or other sources of new variation. Replicators are entities that generate copies of themselves and may represent alleles, genotypes, strategies, beliefs, etc. Populations may consist of multiple types of replicators, each with a characteristic fitness that gives their rate of reproduction. Replicator types with above-average fitness will increase their frequency in the population, and types with below-average fitness will decrease in frequency. Replicator dynamics are widely studied in population genetics, ecology, evolutionary game theory, learning theory, economics, and other contexts.

2.A.1.1. *Discrete-time replicator dynamics (discrete generations).*

Fitness and growth. First we consider a population of replicators with a set of m alternative types that reproduce in discrete generations. Let N_i^t be the number of individuals of the ith type in generation t. The expected number of offspring left by each individual of type i in a generation is given by W_i, which is referred to as the **Wrightian fitness** of the ith type (In general, the fitness of a type may depend on the current environmental conditions, which we will develop further in Appendix 2.A.2). The number of individuals of type i after a total of T generations is given by

$$N_i^T = N_i^0 \prod^T W_i \,. \tag{2.A29}$$

where N_i^0 is the initial number of individuals of type i. The expected number of descendants per initial individual after T generations is given by the ratio $N_i^T/N_i^0 = \prod^T W_i$.

Suppose that we observe N_i^T individuals of type i after T generations, and we wish to know the effective growth rate of this subpopulation over this interval. In other words, what average growth rate r_i corresponds to the reproductive output of individuals of this type? Solving the following expression for r_i, we have

$$e^{r_i T} = \frac{N_i^T}{N_i^0} \tag{2.A30}$$

$$r_i T = \log\left(\frac{N_i^T}{N_i^0}\right) \tag{2.A31}$$

$$r_i = \frac{1}{T}\log\left(\frac{N_i^T}{N_i^0}\right) \tag{2.A32}$$

$$r_i = \frac{1}{T}\log\left(\frac{N_i^0 \prod_{t=0}^T W_i}{N_i^0}\right) \qquad \text{using Equation 2.A29} \tag{2.A33}$$

$$r_i = \frac{1}{T}\log\left(\prod_{t=0}^T W_i\right) \tag{2.A34}$$

$$r_i = \frac{1}{T}\sum_{t=0}^T \log W_i \,. \tag{2.A35}$$

The effective growth rate of the ith type over many generations is equal to the expected logarithm of the type's fitness, and we refer to the **log fitness** $\log W_i$ as the effective growth rate of type i in a single generation (i.e., $r_i = \log W_i$ for $T = 1$).

The **relative fitness** of the ith type, w_i, is considered with respect to the highest fitness of any type in the population, W_*

$$w_i = \frac{W_i}{W_*} \,. \tag{2.A36}$$

By this definition, the relative fitness of the optimal type is equal to 1 (i.e., $w_* = W_*/W_* = 1$), and the relative fitnesses of all other types take values in $[0, 1]$. As such, we may also choose to represent the relative fitness of the ith type as

$$w_i = 1 - k_i \,, \tag{2.A37}$$

where $k_i \in [0, 1]$ is the **Wrightian selection coefficient** of the ith type, and $k_* = 0$ is the selection coefficient of the optimal type. This expression of relative fitness can be interpreted as saying that for every 1 offspring left by the optimal type there are $1 - k_i$ offspring left by the ith type.

Dynamics of type frequency change. The frequency of type i in the population in generation t is defined as

$$p_i^t = \frac{N_i^t}{\sum_z N_z^t} \,. \tag{2.A38}$$

Then the change in the frequency of type i from one generation to the next is given by

$$p_i^{t+1} = \frac{N_i^t W_i}{\sum_z N_z^t W_z} \tag{2.A39}$$

$$= \frac{p_i^t W_i}{\sum_z p_z^t W_z} \tag{2.A40}$$

$$= \frac{p_i^t w_i}{\sum_z p_z^t w_z} \,. \tag{2.A41}$$

Discrete-time replicator dynamics refer to this dynamical map

$$p_i^{t+1} = \frac{p_i^t W_i}{\sum_z p_z^t W_z} \,, \tag{2.A42}$$

where $\sum_z p_z^t W_z$ is the mean fitness of the population at time t. This dynamic re-weights the frequencies of types proportional to their fitnesses such that types with above average fitness increase in frequency and those with below average fitness decrease in frequency.

2.A.1.2. *Continuous-time replicator dynamics (overlapping generations).*

Fitness and growth. While not the focus of the main text, we can also consider the evolution of populations with overlapping generations. Rather than adding individuals to the population at discrete unit intervals, individuals are added to the population at multiple instances per unit time. To capture this, we can express the expected reproductive output per unit time as

$$W_i = \left(1 + \frac{r_i}{\epsilon}\right)^{\epsilon} \,, \tag{2.A43}$$

where r_i gives the expected number of offspring above replacement (or below, if negative) that each i individual contributes to the population per unit time, and ϵ gives the number of instances that offspring are added per unit time. When $\epsilon = 1$, then this definition reduces to the discrete generation case, where $W_i = 1 + r_i$ defines the absolute Wrightian fitness of the ith type. On the other hand, fully continuous growth is represented by

allowing new individuals to be added to the population at an unlimited number of points per unit time, which results in the familiar exponential growth term

$$\lim_{\epsilon\to\infty}\left(1+\frac{r_i}{\epsilon}\right)^{\epsilon}=e^{r_i}\,. \tag{2.A44}$$

Then the number of i individuals at time t in a continuously growing population is given by

$$N_i^t=N_i^0e^{r_it}\,, \tag{2.A45}$$

which is known as the Malthusian growth model. Here we see that the parameter r_i sets the exponential growth rate of the ith type, and we refer to this rate r_i as the **Malthusian fitness** of the ith type. The **Malthusian selection coefficient** s_i of the ith type is defined as the difference between the Malthusian fitness (i.e., growth rate) of the optimal type and that of the ith type

$$s_i=r_*-r_i\,. \tag{2.A46}$$

Dynamics of type frequency change. In this section, we consider the continuous analog of the discrete replicator dynamics presented above. The frequency of the ith type at time t, $p_i^t=N_i^t/\sum_z N_z^t$, is defined as it was for the discrete-time model. We find an expression for the change in this frequency by taking the derivative

$$\frac{dp_i^t}{dt}=\frac{d}{dt}\frac{N_i^t}{\sum_z N_z^t} \qquad p_i^t=\frac{N_i^t}{\sum_z N_z^t} \tag{2.A47}$$

$$=\frac{\left(\frac{d}{dt}N_i^t\right)\sum_z N_z^t-N_i^t\left(\frac{d}{dt}\sum_z N_z^t\right)}{\left(\sum_i N_i^t\right)^2} \qquad \text{quotient rule} \tag{2.A48}$$

$$=\frac{N_i^tr_i\sum_z N_z^t-N_i^t\sum_z N_z^tr_z}{\left(\sum_z N_z^t\right)^2} \qquad \frac{dN_i^t}{dt}=\frac{d}{dt}N_i^0e^{r_it}=N_i^0e^{r_it}r_i=N_i^tr_i \tag{2.A49}$$

$$=\frac{N_i^t\left(r_i\sum_z N_z^t-\sum_z N_z^tr_z\right)}{\left(\sum_z N_z^t\right)^2} \tag{2.A50}$$

$$=\frac{p_i^t\left(r_i\sum_z N_z^t-\sum_z N_z^tr_z\right)}{\sum_z N_z^t} \qquad p_i^t=\frac{N_i^t}{\sum_z N_z^t} \tag{2.A51}$$

$$=\frac{p_i^tr_i\sum_z N_z^t}{\sum_z N_z^t}-\frac{p_i^t\sum_z N_z^tr_z}{\sum_z N_z^t} \tag{2.A52}$$

$$=p_i^tr_i-p_i^t\sum_z\frac{N_z^t}{\sum_v N_v^t}r_z \tag{2.A53}$$

$$=p_i^t\left(r_i-\sum_z p_z^tr_z\right)\,. \tag{2.A54}$$

Thus, the **continuous-time replicator dynamics** are given by the system of differential equations

$$\frac{dp_i^t}{dt}=p_i^t\left(r_i-\sum_z p_z^tr_z\right)\,, \tag{2.A55}$$

where $\sum_z p_z^tr_z$ is the mean Malthusian fitness (i.e., growth rate) of the population at time t. Like the discrete-time dynamics, this dynamic re-weights the frequencies of types according to their fitnesses relative to the population mean fitness.

2.A.2. Modeling variable environments. In general, the fitness of a individual depends on both its type and its environment, and we are ultimately interested in modeling the process of natural selection in contexts where conditions and fitnesses change over time. In this section we extend the standard replicator dynamics models to support both heterogeneous and time-varying environments, broadly defined.

Suppose the population occupies an environment that is comprised of a set of m distinct environmental conditions. We assume that every individual in the population experiences an independent environmental condition, as if each individual occupies a separate part (i.e., micro-environment) of the physical environment. The association of specific individuals to particular conditions is assumed to be random, and the probability that a given individual experiences the jth condition at time t is given by x_j^t. The condition distribution $\boldsymbol{x}^t$ defines the make up of the environment at time t, and scenarios of environmental change can be modeled by specifying a particulars sequence of environments $\{\boldsymbol{x}^0, \boldsymbol{x}^1, \ldots, \boldsymbol{x}^T\}$.

We take a very expansive view of what constitutes an "environment." Essentially any set of contextual factors that have a conditional impact on the fitness of types can be considered an environmental condition. For example, conditions may represent the abiotic conditions experienced by individuals, the types of other replicators encountered by individuals, or even the identity of other alleles carried by individuals. Similarly, the sequence of environments $\{\boldsymbol{x}^0, \boldsymbol{x}^1, \ldots, \boldsymbol{x}^T\}$ may be generated according to any process, including as a function of time, as a function of type frequencies, as a stochastic or random process, as a "pre-programmed" sequence, etc. Throughout this work we consider various environmental contexts of interest as well as the fully general setting where no assumptions are made about the sequence of environmental condition distributions.

In the main text, we frame this model of discrete-time replicator dynamics in variable environments in terms of repeated play of a two-player, population versus environment game, which lends itself to learning theoretic analysis; see Appendix 2.B.2.2 for more information about this formalization.

Discrete-time replicator dynamics in variable environments: Let W_{ij} denote the Wrightian fitness of an individual of type i in environmental condition j. The average fitness of type i across the distribution of environmental conditions $\boldsymbol{x}^t$ is denoted $\langle W_i \rangle^t = \sum_j x_j^t W_{ij}$, and the average fitness the population across all types and conditions is denoted $\langle W \rangle^t = \sum_{ij} p_i^t x_j^t w_{ij}$. Then the discrete-time replicator dynamics in a heterogeneous environment becomes

$$p_i^{t+1} = \frac{p_i^t \langle W_i \rangle^t}{\langle W \rangle^t} \,. \tag{2.A56}$$

Continuous-time replicator dynamics in variable environments: Let r_{ij} denote the Malthusian fitness (i.e., growth rate) of an individual of type i in environmental condition j. The average growth rate of type i across the distribution of environmental conditions $\boldsymbol{x}^t$ is denoted $\langle r_i \rangle^t = \sum_j x_j^t r_{ij}$, and the average growth rate of the population across all types and conditions is denoted $\langle r \rangle^t = \sum_{ij} p_i^t x_j^t r_{ij}$. Then the continuous-time replicator dynamics in a heterogeneous environment becomes

$$\frac{dp_i^t}{dt} = p_i^t \left(\langle r_i \rangle^t - \langle r \rangle^t \right) \,. \tag{2.A57}$$

Appendix 2.B. *Natural selection as a learning process.*

2.B.1. Natural selection and Bayesian learning. Bayesian learning refers to the process of iteratively updating a hypothesis in light of new evidence using Bayes' rule. Suppose that a learner wishes to infer the true state of a system, which may take any one of m possible states. The learner holds a hypothesis about the state of the system, which is represented by a probability distribution that assigns a weight $\Pr(h_i)$ to each alternative. When new evidence E is obtained, the learner updates their prior hypothesis such that alternatives that give relatively high likelihood $\Pr(E|h_i)$ to the observed evidence are up-weighted in the posterior hypothesis. This Bayesian hypothesis update process is formalized by Bayes' rule:

$$\Pr(h_i|E) = \frac{\Pr(E|h_i)\Pr(h_i)}{P(E)} \tag{2.B58}$$

In the context of natural selection, each type in the population can be thought of as an alternative strategy for survival, and the distribution of type frequencies $\boldsymbol{p}$ can be seen as the population's hypothesis about which types are most suited for the current environment. Like Bayesian learning, natural selection updates the population's prior hypothesis in light of new evidence provided by the fitness landscape. As selection changes the frequencies of types according to their relative fitnesses, the distribution of types shifts to favor those that have generated organisms well-adapted to the environment. Indeed, there is a formal connection between Bayesian learning and replicator dynamics. These two update rules share the same form — selection increases the frequency of types with high relative fitness in the same way that Bayes' rule increases the weight of alternatives that give high relative likelihood to the observed evidence:

Bayesian Learning:	Replicator Dynamics:
$\Pr(h_i^{t+1}\|E^t) = \frac{\Pr(E^t\|h_i^t)\Pr(h_i^t)}{P(E^t)}$	$p_i^{t+1} = \frac{W_i p_i^t}{\langle W\rangle^t}$

Natural selection can thus be understood as a learning process that "infers" which types are best adapted to the environment, and a number of recent studies have formalized selection and other evolutionary processes in the framework of Bayesian computations (Campbell 2016, Czégel et al. 2020, Harper 2009a, Shalizi 2009, Watson and Szathmáry 2016). In fact, Bayesian updating is a special case of discrete replicator dynamics, where the "fitness" of each alternative is given by its likelihood, and many results for replicator dynamics can be applied to Bayesian analysis more generally (Shalizi 2009).

2.B.2. Natural selection as an online learning process.

2.B.2.1. ***The online learning problem.*** In the field of computational learning theory, *online learning* refers to a process of iteratively updating a strategy for responding to problems given information about the quality of past responses. That is, a learner is presented with a sequence of problems to which they must provide a response (e.g., an answer to a question, a prediction about the state of a system), and the learner holds a strategy that is used to generate their response to each problem in turn. After providing a response to each problem, the learner receives information about the quality of their response via a *loss* function that measures the discrepancy between the learner's response and the optimal response for that problem. The learner's goal is to minimize the cumulative loss that they suffer in the long run, which they may achieve by updating their strategy after each round such that their responses are more accurate in subsequent rounds. Online learning

algorithms are commonly analyzed in terms of *regret*, which measures the difference between the cumulative loss of a learner and that of a competing strategy, such as the optimal strategy in hindsight. The problem of adapting a strategy for sequential prediction has been studied in a number of fields, including machine learning, game theory, and information theory.

Here we formalize the online learning problem in the game-theoretic framework of playing repeated games against Nature. In the following section, we interpret this formalism for the context of an evolving population (Appendix 2.B.2.2). The repeated games framework is closely related to the prediction with expert advice framework that has been widely studied in computer science.

Consider repeated play of a two-player game in normal form. The game is defined by a matrix $\boldsymbol{G}$ with m rows and n columns representing the pure strategies available to the learner (row player) and the environment (column player), respectively. A *pure strategy* refers to the choice of a specific row or column, while a *mixed strategy* refers to a distribution over rows or columns. A *loss* function $\ell(i, j)$ defines the outcome associated with the learner's pure strategy i being paired with the environment player's pure strategy j. An optimal pure strategy i given the environment's choice of pure strategy j receives zero loss (i.e., $\ell(i, j) = 0$), whereas a pure strategy that fares more poorly against the environment's play receives a correspondingly higher loss. We may assume without loss of generality that all losses are scaled to the range $[0, 1]$, and the environment player's payoffs or losses are left unspecified.

A play of the game in round t consists of the population player choosing a mixed strategy $\boldsymbol{p}^t$ and the column player choosing a mixed strategy $\boldsymbol{x}^t$. The **expected loss** of the learner when mixed strategies $\boldsymbol{p}^t$ and $\boldsymbol{x}^t$ are used is given by

$$\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) = \mathbb{E}_{i\sim\boldsymbol{p}^t, j\sim\boldsymbol{x}^t}\left[\ell(i, j)\right] = \sum_{i,j} p_i^t x_j^t \ell(i, j)\,. \tag{2.B59}$$

Following Freund and Schapire Freund and Schapire (1999), the loss that the learner suffers in each round is equal to the expected loss of their mixed strategy. The loss received by the learner is revealed after the selection of mixed strategies in each round of the game, but the environment player's strategies and the game matrix as a whole are unknown to the learner. The goal of the learner is to minimize its **cumulative loss**

$$L^T = \sum_{t=0}^{T} \ell(\boldsymbol{p}^t, \boldsymbol{x}^t) = \sum_{t=0}^{T} \sum_{i,j} p_i^t x_j^t \ell(i, j)\,. \tag{2.B60}$$

The learner may achieve this goal by iteratively updating their mixed strategy so as to learn a mixed strategy that has minimal per-round expected loss.

Depending on the definition of the game $\boldsymbol{G}$ and the environment player's strategies $\boldsymbol{x}^t$, even the optimal strategy available to the learner may still incur non-zero expected loss in each round, which results in a cumulative loss than continues to increase over time. In addition, we will consider the general case where we make no assumptions about the sequence of mixed strategies chosen by the environment player. In general, the environment may select strategies deterministically, stochastically, arbitrarily, or even adversarially (i.e., the environment can select $\boldsymbol{x}^t$ with knowledge of $\boldsymbol{p}^t$). An adversarial environment can make the learner's cumulative loss arbitrarily large by choosing strategies $\boldsymbol{x}^t$ that give a worst-case loss for the learner's choice of $\boldsymbol{p}^t$. For these reasons, it difficult to evaluate the relative performance of online learning algorithms on the basis of the learner's cumulative loss alone.

Instead, online learning algorithms are commonly analyzed in terms of **regret**, which measures the difference between the cumulative loss of the learner and the cumulative loss of a fixed reference strategy $\boldsymbol{q}$.

$$R_{\boldsymbol{q}}^T = L^T - L_{\boldsymbol{q}}^T = \sum_{t=0}^{T} \ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \sum_{t=0}^{T} \ell(\boldsymbol{q}, \boldsymbol{x}^t)\,. \tag{2.B61}$$

In computational learning theory, regret is typically evaluated with respect to the optimal strategy in hindsight

$$\widetilde{R}^T = \sum_{t=0}^{T} \ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \min_{\tilde{\boldsymbol{p}}} \sum_{t=0}^{T} \ell(\tilde{\boldsymbol{p}}, \boldsymbol{x}^t)\,, \tag{2.B62}$$

where the optimal strategy $\tilde{\boldsymbol{p}}$ is defined as the fixed mixed strategy that minimizes the cumulative loss for the sequence of environment plays observed through round T. In this work, we also evaluate regret with respect to an optimal strategy defined in terms of an evolutionarily stable state, which has particular relevance for the learning problems faced by evolving populations; see Appendix 2.B.3 for more information.

Regret analysis captures the ability of a learning algorithm to arrive at a strategy that does as well as an optimal strategy would in a given scenario. Thus, the learner's goal can be restated as seeking to learn a mixed strategy that minimizes regret with respect to an optimal strategy $\boldsymbol{q}$. A learning process is said to be *no-regret* if it updates the learner's mixed strategy such that their per-round regret vanishes in the long run, achieving

$$\frac{1}{T} R_{\boldsymbol{q}}^T = \frac{1}{T}\left(L^T - L_{\boldsymbol{q}}^T\right) \xrightarrow{T\to\infty} 0\,. \tag{2.B63}$$

A no-regret learning process is optimal in the sense that it achieves the lowest possible per-round regret, which implies that the learner arrives at a strategy that has a per-round loss very nearly equal to that of the optimal strategy.

2.B.2.2. ***The Population versus Environment game.*** The repeated play formalism of the online learning problem (described in detail in Appendix 2.B.2.1) can be interpreted as the basis for a model of natural selection in variable environments[1]. We consider a game between two players: the population[2] (acting as the learner) and the environment. The set of pure strategies available to the population player (i.e., rows of the game matrix) represents the set of replicator types in the population, and the population's strategy $\boldsymbol{p}^t$ represents the type frequency distribution of the population at time t. The set of pure strategies available to the environment player (i.e., columns of the game matrix) represent a set of possible environmental conditions that make up the environment. We assume that every individual in the population experiences an independent environmental condition, as if each individual occupies a separate part (e.g., micro-environment) of the physical environment. Then the environment player's mixed strategy $\boldsymbol{x}^t$ represents the probability distribution of environmental conditions that are experienced by individuals in the population. A play of the game in round t can be interpreted as the population assigning a type to each individual with frequencies $\boldsymbol{p}^t$ and the environment assigning a condition to each micro-environemnt with probabilities $\boldsymbol{x}^t$.

The outcome of each play of the game is scored according to a a game matrix $\boldsymbol{G}$. Each element of the game matrix G_{ij} gives the payoff for a type i that experiences an environmental condition j. Each individual of type i that experiences environmental condition j receives a loss $\ell(i,j)$. Typically the loss $\ell(i,j)$ is given

[1]The Population versus Environment Game model is wholly consistent with the models of replicator dynamics in variable environments outlined in Appendix 2.A.2.

[2]Note that this differs from standard evolutionary game theory where each individual within the population is a separate player

by a function that measures the difference between the payoff for the ith type in the jth condition and the optimal payoff achievable in that condition. For example,

$$\ell(i,j) = \max_i G_{ij} - G_{ij} \,. \tag{2.B64}$$

As we will see in Appendix 2.B.2.4, payoffs and losses are closely related to fitness when considering a population player that updates their strategy according to natural selection in this setting. As in the general online learning setting described above, the loss suffered by the population player in round t is defined as the expected loss $\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) = \sum_{ij} p_i^t x_j^t \ell(i,j)$, which is equivalent to the mean loss of all individuals in the population. The goal of the population player is to adapt its mixed strategy (type frequency distribution) $\boldsymbol{p}^t$ in order to minimize its cumulative loss and achieve low regret. As we will see in Appendix 2.B.2.4, the replicator dynamics model of natural selection is equivalent to an online learning algorithm for precisely this setting.

We take a very expansive view of what constitutes an "environment." Environmental conditions may represent the abiotic conditions experienced by individuals, the types of other replicators encountered by individuals, or even the identity of other alleles carried by individuals. Similarly, the environment player may select their sequence of strategies $\{\boldsymbol{x}^0, \boldsymbol{x}^1, \ldots, \boldsymbol{x}^T\}$ according to any process, including as a function of time, as a function of the population player's strategy (i.e., as a function of type frequencies), as a stochastic or random process, as a "pre-programmed" sequence, etc. While the environment may select the frequency distribution of conditions $\boldsymbol{x}^t$ with knowledge of the population's type frequency distribution $\boldsymbol{p}^t$, the association of particular environmental conditions (micro-environments) to specific individuals is assumed to be random and out of the control of best players. Throughout this work we consider various environmental contexts of interest as well as the fully general setting where no assumptions are made about the sequence of environmental condition distributions.

2.B.2.3. ***The Multiplicative Weights Updating (MWU) algorithm.*** Online learning problems of this form are common, and a simple yet powerful learning algorithm known as the **Multiplicative Weights Updating (MWU)** has been shown to be an effective solution to such problems in computer science, machine learning, economics, and other fields. Here we highlight the conceptual intuitions underlying this algorithm by going through a loss-optimization derivation of MWU before stating the MWU update rule.

Derivation of MWU. [*This section derives the MWU update rule using the method of Lagrange multipliers, the details of which are not critical for understanding this work. The logic that sets up the optimization may provide some insights, even if Lagrange optimization is unfamiliar. Readers simply interested in the definition of MWU can skip to the next section.*]

A learner facing an online learning problem seeks a method for adaptively selecting strategies such that their long-term cumulative loss is minimized. Intuitively, it might seem reasonable for the learner to choose the mixed strategy that would perform the best against previously observed plays by the environment—this is known as the **Follow the Leader (FTL)** algorithm

$$\boldsymbol{p}^{T+1} = \min_{\boldsymbol{p}'} \sum_{t=0}^{T} \ell(\boldsymbol{p}', \boldsymbol{x}^t) = \min_{\boldsymbol{p}'} \sum_{t=0}^{T} \sum_i p_i' \ell(i, \boldsymbol{x}^t) = \min_{\boldsymbol{p}'} \sum_i p_i' L_i^T \,. \tag{2.B65}$$

This update rule effectively places all of the weight on the single pure strategy that has had the lowest empirical loss so far. This Follow the Leader algorithm is inherently unstable as all of the weight can switch from one

pure strategy to another from round to round. Furthermore, there is no guarantee that a strategy that performs well on past observations will continue to do so against future plays of the environment, so a learner using this algorithm can easily be tricked into adopting strategies with poor long-term loss.

In order to add stability to the mixed strategy updates, we can add a regularization term $\mathcal{R}(\boldsymbol{p}')$ to the optimization that generates the updated mixed strategy, where the regularizer $\mathcal{R}$ is some convex function. An appropriately chosen regularizer function $\mathcal{R}$ can offset the influence of the loss minimization term by promoting more uniform weight distributions that do not change too quickly. Such regularization makes the algorithm less vulnerable to short term plays of the environment that might otherwise fool the learner into adopting strategies with high long-term loss. The resulting update rule is known as the **Follow the Regularized Leader (FTRL)** algorithm:

$$\boldsymbol{p}^{T+1} = \arg\min_{\boldsymbol{p}'} \left(\sum_i p'_i L_i^T + \frac{1}{\eta} \mathcal{R}(\boldsymbol{p}') \right) , \tag{2.B66}$$

where $\eta \in (0, \infty)$ is a learning rate parameter that modulates the relative emphasis that is given to minimizing loss in hindsight versus constraining change in each update via the regularization term. When the learning rate η is large, the weights can change greatly in each update (approximating the un-regularized FTL as $\eta \to \infty$). When the learning rate is small, the FTRL update favors more uniform distributions and prevents weight consolidating on a small number of pure strategies in the short-term. In general, the FTRL algorithm balances minimizing long-term loss while maintaining diversity in the mixed strategy distribution during the transient of the learning process.

Intuitively, the negative entropy $-H(\boldsymbol{p}')$ is a reasonable choice for the regularization function. First, the negative entropy is a strongly convex function, which is a requirement of our choice of regularizer function. More importantly, entropy is a measure of the spread of a distribution that is maximized for the uniform distribution, so the *negative* entropy is *minimized* for the uniform distribution. Therefore, using the negative entropy of the mixed strategy as the regularizer will lead the minimization in the FTRL update rule to favor more uniform weight distributions:

$$\boldsymbol{p}^{T+1} = \arg\min_{\boldsymbol{p}'} \left(\sum_i p'_i L_i^T + \frac{1}{\eta} \mathcal{R}(\boldsymbol{p}') \right) \tag{2.B67}$$

$$= \arg\min_{\boldsymbol{p}'} \left(\sum_i p'_i L_i^T - \frac{1}{\eta} H(\boldsymbol{p}') \right) \tag{2.B68}$$

$$= \arg\min_{\boldsymbol{p}'} \left(\sum_i p'_i L_i^T + \frac{1}{\eta} \sum_i p'_i \log p'_i \right) \tag{2.B69}$$

We can set out to solve for the distribution $\boldsymbol{p}'$ that achieves this FTRL update by finding the minimum of this expression. This is a straightforward optimization problem, because the expression to be minimized is strictly concave and has a unique minimum. We solve this optimization using the method of Lagrange multipliers, where we include the constraint that $\sum_z p'_z = 1$. We begin by defining the Lagrange function $\mathcal{L}$ (not to be confused with L, which denotes cumulative loss).

$$\mathcal{L}(\boldsymbol{p}', \lambda) = \left(\sum_i p' L'_i + \frac{1}{\eta} \sum_i p'_i \log p'_i \right) + \lambda \left(1 - \sum_z p'_z \right) \tag{2.B70}$$

Set the derivative of the Lagrange function equal to 0 and evaluate the system of equations that includes the contraint

$$\frac{\partial \mathcal{L}}{\partial p'_i} = L_i^T + \frac{1}{\eta}(1 + \log p'_i) - \lambda = 0 \tag{2.B71}$$

$$\sum_z p'_z = 1 \tag{2.B72}$$

Solving for p'_i in terms of λ

$$L_i^T + \frac{1}{\eta}(1 + \log p'_i) = \lambda \tag{2.B73}$$

$$\frac{1}{\eta}(1 + \log p'_i) = -L_i^T + \lambda \tag{2.B74}$$

$$1 + \log p'_i = \eta(-L_i^T + \lambda) \tag{2.B75}$$

$$\log p'_i = -\eta L_i^T + \eta\lambda - 1 \tag{2.B76}$$

$$p'_i = \exp\left(-\eta L_i^T + \eta\lambda - 1\right) \tag{2.B77}$$

Plugging this expression for p'_i into the constraint to solve for λ

$$\sum_z p'_z = 1 \tag{2.B78}$$

$$\sum_z \exp\left(-\eta L_z^T + \eta\lambda - 1\right) = 1 \tag{2.B79}$$

$$\sum_z \exp\left(-\eta L_z^T\right) \exp\left(\eta\lambda - 1\right) = 1 \tag{2.B80}$$

$$\exp\left(\eta\lambda - 1\right) \sum_z \exp\left(-\eta L_z^T\right) = 1 \tag{2.B81}$$

$$\sum_z \exp\left(-\eta L_z^T\right) = \frac{1}{\exp\left(\eta\lambda - 1\right)} \tag{2.B82}$$

$$\sum_z \exp\left(-\eta L_z^T\right) = \exp\left(-\eta\lambda + 1\right) \tag{2.B83}$$

$$\log\left(\sum_z \exp\left(-\eta L_z^T\right)\right) = 1 - \eta\lambda \tag{2.B84}$$

$$\log\left(\sum_z \exp\left(-\eta L_z^T\right)\right) - 1 = -\eta\lambda \tag{2.B85}$$

$$-\frac{1}{\eta}\left(\log\left(\sum_z \exp\left(-\eta L_z^T\right)\right) - 1\right) = \lambda \tag{2.B86}$$

Plugging this expression for λ into the expression we previously found for p_i' in terms of λ (Equation 2.B77) we can ultimately solve for p_i':

$$p_i' = \exp\left(-\eta L_i^T + \eta\lambda - 1\right) \tag{2.B87}$$

$$p_i' = \exp\left(-\eta L_i^T + \eta\left[-\frac{1}{\eta}\left(\log\left(\sum_z \exp\left(-\eta L_z^T\right)\right) - 1\right)\right] - 1\right) \tag{2.B88}$$

$$p_i' = \exp\left(-\eta L_i^T + \left[-\left(\log\left(\sum_z \exp\left(-\eta L_z^T\right)\right) - 1\right)\right] - 1\right) \tag{2.B89}$$

$$p_i' = \exp\left(-\eta L_i^T + \left[-\log\left(\sum_z \exp\left(-\eta L_z^T\right)\right) + 1\right] - 1\right) \tag{2.B90}$$

$$p_i' = \exp\left(-\eta L_i^T - \log\left(\sum_z \exp\left(-\eta L_z^T\right)\right) + 1 - 1\right) \tag{2.B91}$$

$$p_i' = \exp\left(-\eta L_i^T - \log\left(\sum_z \exp\left(-\eta L_z^T\right)\right)\right) \tag{2.B92}$$

$$p_i' = \exp\left(-\eta L_i^T\right)\exp\left(-\log\left(\sum_z \exp\left(-\eta L_z^T\right)\right)\right) \tag{2.B93}$$

$$p_i' = \frac{\exp\left(-\eta L_i^T\right)}{\exp\left(\log\left(\sum_z \exp\left(-\eta L_z^T\right)\right)\right)} \tag{2.B94}$$

$$p_i' = \frac{\exp\left(-\eta L_i^T\right)}{\sum_z \exp\left(-\eta L_z^T\right)} \tag{2.B95}$$

$$p_i' = \frac{e^{-\eta L_i^T}}{\sum_z e^{-\eta L_z^T}} \tag{2.B96}$$

We have found an expression for the mixed strategy distribution $\boldsymbol{p}'$ that solves the FTRL optimization problem (Equation 2.B69). Therefore, we find that applying the update $p_i^{T+1} = e^{-\eta L_i^T} / \sum_z e^{-\eta L_z^T}$ for each pure strategy i updates the mixed strategy to a distribution that minimizes a combination of expected cumulative loss and negative entropy.

This is a useful update rule in and of itself, but we can go a step further to put this update rule in a form that depends only on the player's loss in the most recent round, which conveniently does not require tracking cumulative loss. Using the recursion $L_i^T = L_i^{T-1} + \ell(i, \boldsymbol{x}^t)$, we can express this update rule in terms of the

most recent single-round losses $\ell(\cdot, \boldsymbol{x}^t)$ alone

$$p_i^{T+1} = \frac{e^{-\eta L_i^T}}{\sum_z e^{-\eta L_z^T}} \tag{2.B97}$$

$$p_i^{T+1} = \frac{e^{-\eta(L_i^{T-1} + \ell(i, \boldsymbol{x}^t))}}{\sum_z e^{-\eta(L_z^{T-1} + \ell(z, \boldsymbol{x}^t))}} \tag{2.B98}$$

$$p_i^{T+1} = \frac{e^{-\eta L_i^{T-1} - \eta \ell(i, \boldsymbol{x}^t)}}{\sum_z e^{-\eta L_z^{T-1} - \eta \ell(z, \boldsymbol{x}^t)}} \tag{2.B99}$$

$$p_i^{T+1} = \frac{e^{-\eta L_i^{T-1}} e^{-\eta \ell(i, \boldsymbol{x}^t)}}{\sum_z e^{-\eta L_z^{T-1}} e^{-\eta \ell(z, \boldsymbol{x}^t)}} \tag{2.B100}$$

$$p_i^{T+1} = \frac{p_i^t \left(\sum_k e^{-\eta L_k^T}\right) e^{-\eta \ell(i, \boldsymbol{x}^t)}}{\sum_z p_z^t \left(\sum_k e^{-\eta L_k^T}\right) e^{-\eta \ell(z, \boldsymbol{x}^t)}} \qquad \text{using } p_i^{T+1} = \frac{e^{-\eta L_i^T}}{\sum_k e^{-\eta L_k^T}} \implies p_i^{T+1} \left(\sum_k e^{-\eta L_k^T}\right) = e^{-\eta L_i^T} \tag{2.B101}$$

$$p_i^{T+1} = \frac{p_i^t \left(\sum_k e^{-\eta L_k^T}\right) e^{-\eta \ell(i, \boldsymbol{x}^t)}}{\left(\sum_k e^{-\eta L_k^T}\right) \sum_z p_z^t e^{-\eta \ell(z, \boldsymbol{x}^t)}} \tag{2.B102}$$

$$p_i^{T+1} = \frac{p_i^t e^{-\eta \ell(i, \boldsymbol{x}^t)}}{\sum_z p_z^t e^{-\eta \ell(z, \boldsymbol{x}^t)}} \tag{2.B103}$$

Finally, we have an update rule (Equation 2.B103) that implements FTRL with negative entropy regularization.

The MWU Update Rule. In the previous section, we derived an update rule that implements the *Follow the Regularized Leader* algorithm with negative entropy regularization, and in so doing updates the learner's mixed strategy distribution so as to optimize a potential function of expected total loss and strategy entropy. This update rule has been repeatedly derived in many contexts and goes by several names, including the **Multiplicative Weights Updating (MWU)**:

$$p_i^{t+1} = \frac{p_i^t e^{-\eta \ell(i, \boldsymbol{x}^t)}}{\sum_z p_z^t e^{-\eta \ell(z, \boldsymbol{x}^t)}}. \tag{2.B104}$$

This algorithm is no-regret in many contexts and achieves reasonable regret bounds in general. This means that a learner who uses MWU to update their strategy learns an approximation of an optimal strategy that achieves a cumulative loss that is nearly as low as that of the optimal strategy. In some settings, such as where the learning rate can be tuned or where the loss function has certain properties, it can be shown MWU in the sense that no other learning algorithm can possibly achieve a tighter bound on regret than MWU. In this work, we are particularly interested in the regime where MWU is equivalent to natural selection (see Appendix 2.B.2.4). Regret bounds for natural selection are given in Appendix 2.C.3.

2.B.2.4. ***Natural selection as a fitness-based instance of MWU.*** In Appendix 2.B.2.2 we establish the online learning problem as the basis for a model of evolution, and in Appendix 2.B.2.3 we derive the MWU update rule, which is known to be an effective learning process for problems of this kind. Recent work has shown that discrete-time replicator dynamics is equivalent to an instantiation of the MWU learning process (Chastain et al. 2014, Chastain 2017, Mehta et al. 2015, Meir and Parkes 2015). In this section, we complete the

connection between online learning and natural selection by describing the equivalence of replicator dynamics and MWU and interpreting the loss function and learning rate implicit in this equivalence.

It is straightforward to recognize that MWU and replicator dynamics[3] have the same form — MWU increases the weight of alternatives that have relatively low loss, and selection increases the frequency of types with high relative fitness:

$$\text{Multiplicative Weights Updating:} \qquad \text{Replicator Dynamics:} \tag{2.B105}$$

$$p_i^{t+1} = \frac{p_i^t e^{-\eta \ell(i,\boldsymbol{x}^t)}}{\sum_z p_z^t e^{-\eta \ell(z,\boldsymbol{x}^t)}} \qquad p_i^{t+1} = \frac{p_i^t \langle w_i \rangle^t}{\sum_z p_z^t \langle w_z \rangle^t} \tag{2.B106}$$

These two update rules are equivalent where the following identity holds

$$e^{-\eta \ell(i,\boldsymbol{x}^t)} = \langle w_i \rangle^t \ . \tag{2.B107}$$

Evaluating this equation further we have

$$e^{-\eta \ell(i,\boldsymbol{x}^t)} = \langle w_i \rangle^t \tag{2.B108}$$

$$\log\left(e^{-\eta \ell(i,\boldsymbol{x}^t)}\right) = \log \langle w_i \rangle^t \tag{2.B109}$$

$$-\eta \ell(i,\boldsymbol{x}^t) = \log \langle w_i \rangle^t \tag{2.B110}$$

$$\eta \ell(i,\boldsymbol{x}^t) = -\log \langle w_i \rangle^t \, . \tag{2.B111}$$

At this point, if we let $\eta = 1$ then the expected loss of the ith type ought to be equal to the negative expected relative fitness of the ith type.

$$\ell(i,\boldsymbol{x}^t) = -\log \langle w_i \rangle^t \, , \text{ where } \eta = 1 \, . \tag{2.B112}$$

[3]Replicator dynamics expressed here in terms of the expected relative fitness $\langle w_i \rangle^t$ of each type i over all environmental conditions.

From here we would like to find the loss function function $\ell(i,j)$ for which this equivalence holds. Let us propose that the choice of loss function $\ell(i,j) = -\log w_{ij}$ satisfies this equation, which we will now prove.

$$\ell(i,\boldsymbol{x}^t) = -\log \langle w_i \rangle^t \text{, where } \eta = 1 \tag{2.B113}$$

$$\sum_j x_j^t \ell(i,j) = -\log \langle w_i \rangle^t \qquad \text{defintion of expected type loss } \ell(i,\boldsymbol{x}^t) \tag{2.B114}$$

$$\sum_j x_j^t \ell(i,j) = -\log \left(\sum_j x_j^t w_{ij} \right) \qquad \text{defintion of expected type relative fitness } \langle w_i \rangle^t \tag{2.B115}$$

$$\sum_j x_j^t \left(-\log w_{ij}\right) = -\log \left(\sum_j x_j^t w_{ij} \right) \qquad \text{choice of loss function } \ell(i,j) \tag{2.B116}$$

$$\sum_j x_j^t \left(\log w_{ij}\right) = \log \left(\sum_j x_j^t w_{ij} \right) \tag{2.B117}$$

$$\sum_j x_j^t \left(\log \left(1-k_{ij}\right)\right) = \log \left(\sum_j x_j^t (1-k_{ij}) \right) \qquad \text{definition of Wrightian selection coefficients } w_{ij} = 1 - k_{ij} \tag{2.B118}$$

$$\sum_j x_j^t \left(\log \left(1-k_{ij}\right)\right) = \log \left(1 - \langle k_i \rangle^t\right) \tag{2.B119}$$

$$\sum_j x_j^t (-k_{ij}) \approx -\langle k_i \rangle^t \text{ for small } k_{ij} \qquad \log(1-u) \approx -u \text{ for } 0 < u \lll 1 \tag{2.B120}$$

$$-\langle k_i \rangle^t \approx -\langle k_i \rangle^t \text{ as was to be shown.} \tag{2.B121}$$

Therefore, Multiplicative Weights Updating is equivalent to discrete-time replicator dynamics for haploid asexual populations under the following conditions:

(1) Loss function $\ell(i,j) = -\log w_{ij}$.
(2) Learning rate $\eta = 1$.
(3) In the limit of weak selection, where $k_{ij} \lll 1$ for all i, j (or in the limit of continuously overlapping generations).

Using the definition of relative fitness w_{ij}, the loss function $\ell(i,j) = -\log w_{ij}$ can be expressed and interpreted in another form.

$$\begin{aligned}
\ell(i,j) &= -\log w_{ij} && \text{(2.B122)}\\
&= -\log \frac{W_{ij}}{W_{*j}} && \text{(2.B123)}\\
&= \log \frac{W_{*j}}{W_{ij}} && \text{(2.B124)}\\
&= \log W_{*j} - \log W_{ij} && \text{(2.B125)}\\
&= \max_i (\log W_{ij}) - \log W_{ij}\,. && \text{(2.B126)}
\end{aligned}$$

Thus we see that this loss function measures the difference between the log fitness (growth rate) for the ith type in the jth condition and the optimal payoff achievable in that condition. This is a sensible way to define the validity of the individual "response" i to the "context" j in learning-theoretic terms. Moreover, this difference gives the loss of potential fitness that an individual experiences by having a suboptimal type for the condition they experience, which is meaningful in the context of a population evolving by natural selection. Identifying log fitness as the payoff compared by the loss function (Equation 2.B64) tells us that the n-by-m matrix $\boldsymbol{G}$ defining the Population versus Environment game gives the log fitnesses (growth rates) of all n types in all m conditions (i.e, $G_{ij} = \log W_{ij}$).

The learning rate η sets the relative emphasis that Multiplicative Weight Updating gives to minimizing loss based on previous options and maintaining spread over types as a form of regularization (Equation 2.B66). Learning rates $\eta > 1$ react strongly to each observed round of losses, while learning rates $\eta < 1$ put more emphasis on maintaining entropy in the learner's strategy. The fact that a learning rate $\eta = 1$ is implicit in the equivalence between natural selection and MWU tells us that selection balances maximizing expected cumulative fitness (minimizing expected fitness loss) and maintaining diversity exactly equally (Chastain et al. 2014).

2.B.3. Two classes of learning problems for evolving populations. For a population that learns by natural selection, it is useful to distinguish two classes of learning problems that lead to different conceptions of strategy optimality and corresponding definitions of regret.

The nature of the learning problem faced by a population depends on the sequence of environments it experiences. A given distribution of environmental conditions $\boldsymbol{x}^t$ determines a *fitness landscape*: the set of expected fitnesses for each type in the population (i.e., $\boldsymbol{G}\boldsymbol{x}^t$ gives the fitness landscape at time t). The core problem for the evolving population can be seen to be arriving at an equilibrium state that maximizes fitness locally with respect to the current fitness landscape. The dynamics of selection continually move the population toward an accessible stable state, which decreases potential information and increases fitness with respect to the present environment. In other words, what selection learns is an evolutionarily stable type composition for the current conditions. When the makeup of the environment changes the fitness landscape changes as well, but the evolutionarily stable state that represents the population's learning target may or may not change depending on the particular sequence of changes to the fitness landscape. A sequence of environments for which the evolutionarily stable state accessible to the population remains stationary represents a **fixed learning problem**. On the other hand, many sequences of environments do cause the evolutionarily stable states to

change over time, which we refer to as **variable learning problems**. These classes of learning problems are defined formally below.

Definition. *A type distribution $\overline{\boldsymbol{p}}_{\boldsymbol{x}^t}$ is an* **evolutionarily stable state** *of continuous-time replicator dynamics with respect to the distribution of environmental conditions $\boldsymbol{x}^t$ if for all distributions $\boldsymbol{q} \neq \overline{\boldsymbol{p}}_{\boldsymbol{x}^t}$ in a neighborhood of $\overline{\boldsymbol{p}}_{\boldsymbol{x}^t}$*

$$\boldsymbol{q}^\mathsf{T}\boldsymbol{G}\boldsymbol{x}^t < \overline{\boldsymbol{p}}_{\boldsymbol{x}^t}^\mathsf{T}\boldsymbol{G}\boldsymbol{x}^t \implies \begin{cases} \sum_{ij} q_i x_j^t \log W_{ij} < \sum_{ij} \overline{p}_{\boldsymbol{x}^t,i} x_j^t \log W_{ij} & \text{for discrete-time replicator dynamics} \\ \sum_{ij} q_i x_j^t r_{ij} < \sum_{ij} \overline{p}_{\boldsymbol{x}^t,i} x_j^t r_{ij} & \text{for continuous-time replicator dynamics} \end{cases} \tag{2.B127}$$

Definition. *The* **basin of attraction** *of an evolutionarily stable state $\overline{\boldsymbol{p}}_{\boldsymbol{x}^t}$ is the set of all initial conditions from which trajectories of replicator dynamics approach $\overline{\boldsymbol{p}}_{\boldsymbol{x}^t}$. Formally, the basin of attraction $\mathcal{B}(\overline{\boldsymbol{p}}_{\boldsymbol{x}^t})$ is a positively invariant set (i.e., all trajectories starting in $\mathcal{B}(\overline{\boldsymbol{p}}_{\boldsymbol{x}^t})$ remain in $\mathcal{B}(\overline{\boldsymbol{p}}_{\boldsymbol{x}^t})$) that contains $\overline{\boldsymbol{p}}_{\boldsymbol{x}^t}$ and for which there exists a Lyapunov function $V(\boldsymbol{p})$ that satisfies*

$$V(\boldsymbol{p}) \geq 0 \quad \forall \boldsymbol{p} \in \mathcal{B}(\overline{\boldsymbol{p}}_{\boldsymbol{x}^t}) \tag{2.B128}$$

$$V(\boldsymbol{p}) = 0 \quad \textit{iff } \boldsymbol{p} = \overline{\boldsymbol{p}}_{\boldsymbol{x}^t} \tag{2.B129}$$

$$\frac{d}{dt}V(\boldsymbol{p}) < 0 \quad \forall \boldsymbol{p} \in \mathcal{B}(\overline{\boldsymbol{p}}_{\boldsymbol{x}^t}) \setminus \boldsymbol{p} \tag{2.B130}$$

Then, by La Salle's Invariance Principle, every trajectory starting at $\boldsymbol{p}^0 \in \mathcal{B}(\overline{\boldsymbol{p}}_{\boldsymbol{x}^t})$ tends to $\overline{\boldsymbol{p}}_{\boldsymbol{x}^t}$ as $t \to \infty$.

Definition. *An evolutionarily stable state $\overline{\boldsymbol{p}}$ is said to be* **stationary** *over a sequence of environments $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$ if the equilibrium point remains constant and the population's trajectory $\boldsymbol{p}^t$ remains within its basin of attraction throughout the sequence. That is,*

$$\overline{\boldsymbol{p}} = \overline{\boldsymbol{p}}_{\boldsymbol{x}^t} \qquad \textit{and} \qquad \boldsymbol{p}^t \in \mathcal{B}(\overline{\boldsymbol{p}}_{\boldsymbol{x}^t}) \;\; \forall\, \boldsymbol{x}^t \in \{\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T\}\,. \tag{2.B131}$$

Definition. *Any sequence of environments $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$ for which the population's initial type distribution $\boldsymbol{p}^0$ is in the basin of attraction of a stationary evolutionarily stable state $\overline{\boldsymbol{p}}$ is said to represent a* **fixed learning problem** *for the population.*

Definition. *Any sequence of environments $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$ that does not meet the criteria of a fixed learning problem — that is, where the evolutionarily stable state accessible to the population changes throughout the sequence of environments — is said to represent a* **variable learning problem**.

Appendix 2.C. *The cost of natural selection.*

2.C.1. Mismatch load quantifies the cumulative fitness loss of selection. The *genetic load* of a population refers to "the extent to which the population is impaired by the fact that not all individuals in the population are of the optimum type" (Crow 1958). A population that includes suboptimal types for the current environment will have a lower average population fitness than a population that is fixed for the optimal type (i.e., the type with the maximum fitness), and genetic load measures the proportion by which the population fitness is decreased due to the presence of the inferior types. A number of factors can contribute to the presence of suboptimal types in a population and thus to the overall genetic load. For example, *mutation load* refers to the depression in population fitness due to the occurrence of deleterious mutations, and *segregation load* refers to a depression in population fitness due to the production of inferior homozygotes by allele segregation in the context of heterozygote advantage (Crow 1958, Kimura 1960). The component of load that is the focus of this paper is the *substitution load* and its generalization, *mismatch load*: the depression in population fitness that results from mismatch of individual types and environmental conditions.

2.C.1.1. ***Substitution load.*** In the following sections, we walk through Haldane's and Kimura's definitions of substitution load in terms of Wrightian fitnesses (i.e., discrete generations) and Malthusian fitness (i.e., continuous growth), respectively.

Substitution load for discrete-time replicator dynamics. The basic idea of genetic load was introduced by J.B.S. Haldane (1937). Haldane (1957) was also the first to consider a "cost of natural selection," in which he quantified the number of selective deaths that must occur in order to replace one allele with a more fit one.

Haldane (1957) considers a haploid population with discrete generations that undergoes selection in response to a change in the environment. The result of this change is that one or more types that were previously rare become beneficial. Overall, the population is less adapted to the new environment, and its reproductive capacity is lowered due to the predominance of types that have poor viability, fertility, etc. in the new environment. Natural selection gradually improves the composition of the population, but "meanwhile a number of deaths, or their equivalents in fertility, have occurred" (Haldane 1957). Haldane sets out to quantify the total number of "selective deaths" that occur during the process of selection substituting the newly optimal type.

Let the absolute Wrightian fitness of each type i in the new environment be given by W_i.[4] The fitness of the ith type relative to the fitness of the optimal type is given by

$$w_i = \frac{W_i}{W_*} = 1 - k_i \tag{2.C132}$$

where $k_i \in [0, 1]$ is the Wrightian selection coefficient of the ith type, and $k_* = 0$ is the selection coefficient of the optimal type. In other words, $1 - k_i$ individuals of type i survive and reproduce for every one of the optimal type. Therefore the selection coefficient k_i can be interpreted as the per capita number of "selective deaths" of type i. Then the expected number of selective deaths in the population in single a generation is given by $\sum_i p_i^t k_i$. Haldane (1957) goes on the evaluate the total number of expected selective deaths accumulated

[4]Haldane (1957) considered a 2 allele model, but it is straightforward to extend this to m types.

over T generations of selection

$$\sum_{t=0}^{T} \sum_{i} p_i^t k_i \,. \tag{2.C133}$$

Motoo Kimura further developed this concept into the more general definition of substitution load (Kimura 1961a, 1960). Kimura defined substitution load as the cumulative difference between a population's average growth rate and the growth rate of the optimal type over the course of an allele substitution. Kimura's definition clarifies that selection and the resulting substitution load need not involve actual deaths.

While Kimura defined this quantity in terms of continuous growth rates (see the next section), we consider an analogous definition for discrete generations in this paper. The log fitness $\log W_i$ gives the expected instantaneous growth rate of the subpopulation of type i individuals in a single generation (Appendix 2.A.1.1). The difference between the growth rate of the optimal type and the population's average growth rate in a single generation is then given by

$$\log W_* - \langle \log W \rangle^t = \sum_i p_i^t \left(\log W_* - \log W_i \right) \,. \tag{2.C134}$$

This difference gives the expected depression in single-generation growth that the population experiences due to its mixed composition relative to a population that consists of only the optimal type. Natural selection will gradually increase the the frequency of the optimal type in the evolving population, and this difference will decrease over time. Substitution load L_{sub}^T is defined as the total sum difference between the evolving population's growth rate and the optimal growth rate over T generations of a gene substitution process in a constant environment

$$L_{\text{sub}}^T = \sum_{t=0}^{T} \log W_* - \langle \log W \rangle^t \tag{2.C135}$$

$$= \sum_i p_i^t \left(\log W_* - \log W_i \right) \,. \tag{2.C136}$$

This quantity gives the expected fold difference in long-term growth of the evolving population relative to a population with an optimal composition, where the 'fold' refers to the base of the logarithms. When logarithms to base 2 are used, a substitution load $L_{\text{sub}}^T = 1$ refers to one 2-fold reduction in the population's long term growth relative to the optimum.

Kimura's substitution load is equivalent to Haldane's expression for the number of selective deaths in the limit of weak selection or continuously overlapping generations. To see this, we can work out an expression

for the discrete-time definition of substitution load in terms of Wrightian selection coefficients.

$$L_{\text{sub}}^{T} = \sum_{t=0}^{T} \log W_* - \langle \log W \rangle^t \tag{2.C137}$$

$$= \sum_i p_i^t \left(\log W_* - \log W_i \right) \tag{2.C138}$$

$$= \sum_i p_i^t \left(\log \frac{W_*}{W_i} \right) \tag{2.C139}$$

$$= -\sum_i p_i^t \left(\log \frac{W_i}{W_*} \right) \tag{2.C140}$$

$$= -\sum_i p_i^t \left(\log w_i \right) \qquad \text{definition of relative fitness } w_i \tag{2.C141}$$

$$= -\sum_i p_i^t \left(\log \left(1 - k_i \right) \right) \tag{2.C142}$$

$$\approx -\sum_i p_i^t \left(-k_i \right) \text{ for small } k_i \qquad \log\left(1-u\right) \approx -u \text{ for } 0 < u \lll 1 \tag{2.C143}$$

$$= \sum_i p_i^t k_i \, . \tag{2.C144}$$

The final line is equal to the number of selective deaths in Haldane's model (Equation 2.C133), and thus Haldane's quantity is equivalent to discrete-time substitution load in the limit of weak selection (small k_i for all i). The connection between the discrete generations and continuous growth definitions of substitution load can be seen in the following section.

Substitution load for continuous-time replicator dynamics. Now we turn to Kimura's definition of substitution load for haploid populations with overlapping generations in a constant, homogeneous environment Kimura (1961a, 1960). Consider a population that consists of multiple types[5], where p_i^t and r_i give the frequency and Malthusian fitness (exponential growth rate) of the i-th type at time t, respectively. Suppose the optimal type, which confers the maximum growth rate $r_* = \max_i r_i$, has an initial frequency $0 < p_*^0 < 1$. The difference between the growth rate of the optimal type and that of the ith type is given by the Malthusian selection coefficient

$$s_i = r_* - r_i \, . \tag{2.C145}$$

This fitness difference can be thought of as the growth penalty that the population receives for each individual that is of the ith type rather than the optimal type for the environment. Then the population's expected fitness reduction due to the presence of inferior types at a given time t is the frequency-weighted average of these fitness differences

$$\sum_i p_i^t \left(r_* - r_i \right) = \sum_i p_i^t s_i \, . \tag{2.C146}$$

[5] Likewise, we extend the 2 allele model from Kimura (1961a, 1960) to m types.

This expected fitness reduction integrated over a period of selection of duration T defines the **substitution load**

$$L_{\text{sub}}^{T} = \int_0^T \sum_i p_i^t \left(r_* - r_i\right) dt \,. \tag{2.C147}$$

Substitution load can be equivalently expressed as the cumulative difference between the optimal fitness and the population's mean fitness $\langle r \rangle^t = \sum_i p_i^t r_i$

$$L_{\text{sub}}^{T} = \int_0^T r_* - \langle r \rangle^t \; dt \,, \tag{2.C148}$$

or as the integral of the population's mean selection coefficient $\langle s \rangle^t = \sum_i p_i^t s_i$

$$L_{\text{sub}}^{T} = \int_0^T \langle s \rangle^t \; dt \,. \tag{2.C149}$$

We can evaluate this integral to obtain a closed-form expression for substitution load. Evaluating this integral with respect to time is not straight-forward, but we can make use of the definition of continuous-time replicator dynamics to change the integration variable to one that is simpler to work with. Recall the continuous-time replicator dynamics that governs how type frequencies change over time

$$\frac{dp_i}{dt} = p_i^t (r_i - \langle r \rangle^t) \,. \qquad \text{Equation 2.A55} \tag{2.C150}$$

We can rearrange this differential equation to find an expression for dt in terms of the frequency and fitness variables.

$$dt = \frac{dp_i}{p_i^t (r_i - \langle r \rangle^t)} \,. \tag{2.C151}$$

If we consider the replicator dynamics for the optimal allele in particular, the expression we obtain for dt includes in the denominator a term that appears in the substitution load integral, which will prove useful for simplifying the integral

$$dt = \frac{dp_*}{p_*^t (r_* - \langle r \rangle^t)} = \frac{dp_*}{p_*^t \langle s \rangle^t} \,. \tag{2.C152}$$

Now we can use this expression for dt to change the integration variable and evaluate the simplified integral

$$L_{\text{sub}}^{T} = \int_{0}^{T} \langle s \rangle^{t} \; dt \tag{2.C153}$$

$$= \int_{p_*^0}^{p_*^T} \langle s \rangle^{t} \; \frac{dp_*}{p_* \langle s \rangle^{t}} \qquad \text{using Equation 2.C152} \tag{2.C154}$$

$$= \int_{p_*^0}^{p_*^T} \frac{\langle s \rangle^{t}}{\langle s \rangle^{t}} \frac{dp_*}{p_*} \tag{2.C155}$$

$$= \int_{p_*^0}^{p_*^T} \frac{dp_*}{p_*} \tag{2.C156}$$

$$= \log p_* \Big|_{p_*^0}^{p_*^T} \tag{2.C157}$$

$$= \log p_*^T - \log p_*^0 \tag{2.C158}$$

$$= \log \left(\frac{p_*^T}{p_*^0} \right) \tag{2.C159}$$

Here we see that the substitution load after a period of selection of duration T is a function of the pre- and post-selection frequencies of the optimal allele. In the long run, as the population approaches fixation for the optimal allele, the total substitution load converges to

$$\lim_{T \to \infty} L_{\text{sub}}^{T} = \int_{0}^{\infty} \langle s \rangle^{t} \; dt = \ln \left(\frac{1}{p_*^0} \right) = -\ln p_*^0 \,, \tag{2.C160}$$

a value that depends only on the initial frequency of the optimal allele. Interestingly, the total substitution load for a complete allele substitution does not depend on the strength of selection; the relative fitness advantage of the optimal allele over other alleles in the population will modulate the time it takes for selection to fix this allele, but the total load will be the same in any case.

2.C.1.2. ***Mismatch load.*** The classical definitions of substitution load introduced by Haldane (1957) and Kimura (1961a, 1960) consider a model of evolution in a homogeneous and static environment. In such a case, there is a strictly optimal type with maximal fitness for the given environment, and the result of natural selection is a sweep where the optimal type substitutes all others. Substitution load measures the cumulative loss of fitness associated with selection completing this substitution gradually. This is a useful thing to measure, but this classical setting is restrictive.

We would like to extend this notion of load to more general settings with heterogeneous and time-varying environments. In Appendix 2.A.2 and Appendix 2.B.2.2 we present a model of evolution in variable environments. The environment is considered to be constituted by a set of distinct conditions. Each individual in the population experiences an independent micro-environment characterized by particular condition, and the fitness W_{ij} (or r_{ij} for continuous dynamics) of each individual depends both on its type i and on the environmental condition j that it experiences. The core concept that substitution load measures is the accumulation of fitness losses due to the presence of types that are poorly suited for the conditions they experience. In a heterogeneous environment, different environmental conditions may favor different types. In principle, overall fitness is maximized if every individual possesses the type that is optimal for the specific environmental condition that they experience. However, we assume that types are distributed over micro-environments and

associated conditions randomly, so some type-condition mismatch is inevitable. We introduce the concept of **mismatch load**, which quantifies the cumulative loss of potential fitness due to the mismatch of types and environmental conditions.

Consider an environmental condition j. An individual of type i that experiences this condition has fitness W_{ij}, while the fitness of the type that is optimal for this condition is denoted by $W_{*j} = \max_i W_{ij}$. The amount of potential fitness that an individual misses out on by having a suboptimal type i in condition j is given by the difference

$$\log W_{*j} - \log W_{ij} \,. \tag{2.C161}$$

The expected amount of fitness loss incurred on the population due to type-condition mismatch over the distribution of types $\boldsymbol{p}^t$ and distribution of conditions $\boldsymbol{x}^t$ in generation t is given by

$$\sum_{ij} p_i^t x_j^t \left(\log W_{*j} - \log W_{ij}\right) \,. \tag{2.C162}$$

Mismatch load L^T measures the cumulative loss of fitness due to type-condition mismatch over T generations

$$L^T = \sum_{t=0}^{T} \sum_{i,j} p_i^t x_j^t \left(\log W_{*j} - \log W_{ij}\right) \tag{2.C163}$$

$$= \sum_{t=0}^{T} \left(\langle \log W_* \rangle^t - \langle \log W \rangle^t \right) \,. \tag{2.C164}$$

Mismatch load is equivalent to substitution load (Equation 2.C136) when there is only one environmental condition (i.e., $m = 1$, $\boldsymbol{x}^t$ is a point distribution). Like substitution load, mismatch load quantifies the fold reduction in total growth that an evolving population experiences due to having a composition that is not optimally matched to the environmental conditions. In the case of a gene substitution in a constant environment, substitution load converges on a finite value as selection approaches fixation of the optimal type. However, when the environment is heterogeneous there may not be a single type that is universally optimal in all conditions. In such a case, there is no type composition that can be expected to achieve zero mismatch given random associations of types with conditions. Therefore, mismatch load does not necessarily converge on a finite value and may accrue even for the type composition that gives the optimal expected fitness over all types and conditions.

In Appendix 2.B.2.4 we show that replicator dynamics are equivalent to an instantiation of the MWU online learning algorithm characterized by the learning rate $\eta = 1$ and fitness-based loss function

$$\ell(i, j) = -\log w_{ij} = \log W_{*j} - \log W_{ij} \,. \tag{2.C165}$$

This function defines the *loss* that an individual of type i incurs in environmental condition j. In biological terms, this loss function gives the loss of potential fitness that an individual of type i experiences due to having a suboptimal type in condition j. The expected loss of all type i individuals taken over the distribution of environmental conditions $\boldsymbol{x}^t$ at time t is given by

$$\ell(i, \boldsymbol{x}^t) = \sum_j x_j^t \left(\log W_{*j} - \log W_{ij}\right) \,. \tag{2.C166}$$

The expected loss of fitness experienced by the overall population in generation t (i.e., the expected loss taken by the population player in round t of the Population versus Environment game) is given by

$$\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) = \sum_j p_i^t x_j^t \left(\log W_{*j} - \log W_{ij}\right) . \tag{2.C167}$$

Then the cumulative expected loss of the population summed over T generations (i.e., T rounds of the PvE game) is given by

$$\sum_{t=0}^{T} \ell(\boldsymbol{p}^t, \boldsymbol{x}^t) = \sum_{t=0}^{T} \sum_j p_i^t x_j^t \left(\log W_{*j} - \log W_{ij}\right) . \tag{2.C168}$$

We recognize this quantity as being identical to the definition of mismatch load (Equation 2.7, Equation 2.C164, Appendix 2.C.1.2)

$$\sum_{t=0}^{T} \ell(\boldsymbol{p}^t, \boldsymbol{x}^t) = L^T . \tag{2.C169}$$

This connection reinforces the meaning of mismatch load and makes intuitive sense. Natural selection is equivalent to the MWU online learning algorithm when the learning-theoretic loss function is defined as the fitness loss each individual incurs given the environmental conditions they experience. The more type-condition mismatch occurs, the more loss the population is expected to incur and accrue, both in terms of learning-theoretic loss and long-term fitness loss.

2.C.2. Regret quantifies the cost of selection. In Appendix 2.B.2 we introduced the concept of regret in the context of online learning. Regret is used in learning theory to evaluate the performance of a learning process by comparing the responses of a learner to the responses produced by a fixed reference strategy. Responses are typically scored using a loss function that quantifies the validity of a response relative to the best response in a given situation. When the reference strategy is taken to be optimal, regret represents the cost of the learning process: the excess loss a learner suffers by having to learn an appropriate response as opposed to knowing the solution all along.

In this section, we show the connection between regret and load and establish two definitions of regret in the context of an evolving population.

2.C.2.1. ***Regret quantifies relative mismatch load.*** In this section, we evaluate the definition of regret that applies to natural selection as a learning process. In general, regret is defined as the difference between the cumulative expected loss of the learner and the cumulative expected loss of a fixed reference strategy $\boldsymbol{q}$ (Appendix 2.B.2.1).

$$R_{\boldsymbol{q}}^T = \sum_{t=0}^{T} \ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \sum_{t=0}^{T} \ell(\boldsymbol{q}, \boldsymbol{x}^t) \tag{2.C170}$$

$$= \sum_{t=0}^{T} \sum_{ij} p_i^t x_j^t \ell(i,j) - \sum_{t=0}^{T} \sum_{ij} q_i^t x_j^t \ell(i,j) . \tag{2.C171}$$

If we plug in the loss function $\ell(i,j) = \log W_{*j} - \log W_{ij}$ that is implicated in using natural selection as the learning process we have

$$R_{\boldsymbol{q}}^T = \sum_{t=0}^{T} \sum_{ij} p_i^t x_j^t \ell(i,j) - \sum_{t=0}^{T} \sum_{ij} q_i^t x_j^t \ell(i,j) \tag{2.C172}$$

$$= \underbrace{\sum_{t=0}^{T} \sum_{ij} p_i^t x_j^t \left(\log W_{*j} - \log W_{ij}\right)}_{\text{mismatch load } L^T} - \sum_{t=0}^{T} \sum_{ij} q_i^t x_j^t \left(\log W_{*j} - \log W_{ij}\right) . \tag{2.C173}$$

In Appendix 2.C.1.2 we show that the cumulative loss of natural selection is equivalent to the definition of mismatch load (Equation 2.7, eqn:app-mismatch-load, Appendix 2.C.1.2). The second term has the same form, but is calculated with respect to the fixed reference distribution $\boldsymbol{q}$. Therefore we can interpret the regret of selection as the difference between the evolving population's mismatch load and the mismatch load experienced by the reference type composition $\boldsymbol{q}$

$$R_{\boldsymbol{q}}^T = L^T - L_{\boldsymbol{q}}^T . \tag{2.C174}$$

Thus regret represents a measure of relative mismatch load. If the reference distribution $\boldsymbol{q}$ is taken to be an optimal type composition in some sense, then regret measures the amount by which the mismatch load experienced by the evolving population exceeds the mismatch load it could have achieved if it adopted the optimal fixed composition from the beginning. This is a rigorous measure of the cost of the selection process that quantifies how much loss of long-term fitness the population incurs by gradually evolving as opposed to enjoying an optimally adapted composition from the beginning. In this work we consider the regret of selection with respect to two alternative definitions of fixed optimal compositions; see Appendix 2.C.2.2 for more details.

2.C.2.2. ***ESS regret and empirical regret.*** Regret measures the performance of a learning process in terms of the cumulative loss of the learner's strategies relative to that of a fixed reference strategy. Typically the reference strategy is taken to be a strategy that is optimal in some sense. For a population that learns by natural selection, it is useful to distinguish two classes of learning problems that lead to different definitions of strategy optimality and thus regret.

ESS regret measures the cost of selection with respect to fixed learning problems. First we consider a definition of regret that pertains to a population that evolves by natural selection in response to a fixed learning problem. In this context, selection moves the population toward a stationary evolutionarily stable state. By definition, the evolutionarily stable state is an local optima where the population has higher expected fitness than all other neighboring type frequency distributions for the given environment (Maynard Smith 1982). Natural selection is a process by which the population learns this optimal composition. To evaluate the cost of natural selection as a learning process in this setting, it is natural to measure how the cumulative fitness loss of the evolving population compares to the cumulative fitness loss that would have have been achieved if the population used the optimal, evolutionarily stable composition all along in the same sequence of environments. This measures defines what we refer to as the *ESS regret*.

Definition. *For a sequence of environments $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$ that constitutes a fixed learning problem with a stationary evolutionarily stable state $\overline{\mathbf{p}}$, the regret of a learning process that generates a sequence of type frequency distributions*

$\boldsymbol{p}^0, \ldots, \boldsymbol{p}^T$ *measured with respect to* $\overline{\boldsymbol{p}}$ *is referred to as the* **ESS regret**

$$\bar{R}^T = \sum_{t=0}^{T} \ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\overline{\boldsymbol{p}}, \boldsymbol{x}^t)\,, \tag{2.C175}$$

where the notation $\bar{R}^T$ *is used as a shorthand for* $R^T_{\overline{\boldsymbol{p}}}$ *for typesetting clarity.*

Empirical regret measures the cost of selection in generalized settings. ESS regret is a natural measure that captures the cost of selection for the core problem of interest for evolving populations—finding evolutionarily stable states—however it is only applicable to learning problems characterized by stationary evolutionarily stable states. In general, a population may face a sequence of environments for which the accessible evolutionarily stable state shifts over time. In addition, it may not always be reasonable to assume that the evolutionarily stable state is known and usable as a reference distribution for the measurement of regret. For example, an experimentalist may not have sufficient knowledge of the empirical Population versus Environment game they are studying to be able to identify the stable state(s) that characterize the learning problem faced by an empirical population.

In computational learning theory, regret is typically measured with respect to the fixed strategy that would have had the minimum load possible against the observed sequence of environments in hindsight.

Definition. *For any sequence of environments* $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$*, the regret of a learning process that generates a sequence of type frequency distributions* $\boldsymbol{p}^0, \ldots, \boldsymbol{p}^T$ *measured with respect to the fixed strategy* $\widetilde{\boldsymbol{p}}^T$ *that gives the minimum loss over the sequence of environments observed in hindsight is referred to as the* **empirical regret**

$$\widetilde{R}^T = \sum_{t=0}^{T} \ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\widetilde{\boldsymbol{p}}^T, \boldsymbol{x}^t)\,, \tag{2.C176}$$

where the ***empirically optimal strategy*** $\widetilde{\boldsymbol{p}}^T$ *is defined as*

$$\widetilde{\boldsymbol{p}}^T = \arg\min_{\boldsymbol{q}} \sum_{t=0}^{T} \ell(\boldsymbol{q}, \boldsymbol{x}^t)\,, \tag{2.C177}$$

and where the notation $\widetilde{R}^T$ *is used as a shorthand for* $R^T_{\widetilde{\boldsymbol{p}}^T}$ *for typesetting clarity.*

This version of regret is defined solely with respect to the empirical sequence of environments that the observed by the learner and does not rely on any defining characteristics of the sequence of environments or the learning process itself. Therefore this definition of regret is fully general and can be used to measure the cost of any learning process in any online learning setting. Given the generality of this definition, bounds that are established for the empirical regret hold for all games and for all possible environmental sequences and therefore represent worst-case regret guarantees for a learning process.

2.C.3. Regret bounds for natural selection. Now that we have established regret as a meaningful measure of the cost of a learning process, we would like to establish bounds on the cost of selection. Following in the footsteps of foundational works in online learning theory (Cesa-Bianchi and Lugosi 2006, Freund and Schapire 1999, Kivinen and Warmuth 1995), our basic method for deriving bounds will be a kind of amortized analysis that tracks the progression of the learning process using a potential function. We will consider some

potential function of frequency vectors $\phi(\boldsymbol{p})$ that is assumed to be non-negative and bounded from above

$$\phi(\boldsymbol{p}) \geq 0 \qquad \forall \boldsymbol{p} \tag{2.C178}$$

$$\phi(\boldsymbol{p}) \leq \phi^0 \qquad \forall \boldsymbol{p}\,. \tag{2.C179}$$

The difference in potential $\phi(\boldsymbol{p}^t) - \phi(\boldsymbol{p}^{t+1})$ (or $-\frac{d}{dt}\phi(\boldsymbol{p}^t)$ for a continuous process) is one way of describing the progress made by the learning process at time t. If we can prove that

$$\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{q}, \boldsymbol{x}^t) \leq \phi(\boldsymbol{p}^t) - \phi(\boldsymbol{p}^{t+1}) \tag{2.C180}$$

for all possible updates made by the learning process, then by summing the progression over time we will obtain

$$R_{\boldsymbol{q}}^T \leq \phi(\boldsymbol{p}^0) - \phi(\boldsymbol{p}^{T+1}) \leq \phi^0\,. \tag{2.C181}$$

In other words, if the instantaneous regret is bounded by the progress in that step, and the total amount of progress possible is bounded, then we obtain a bound for the total regret.

The particular potential function that we will use in this analysis is the aptly named *potential information*

$$\phi(\boldsymbol{p}^t) = D(\boldsymbol{q}||\boldsymbol{p}^t)\,. \tag{2.C182}$$

The potential information is defined as the KL divergence of a reference distribution $\boldsymbol{q}$ from the learner's current distribution $\boldsymbol{p}^t$, which can be interpreted as the amount of information the learner stands to gain by updating to the distribution $\boldsymbol{q}$. This choice of potential function is not arbitrary but rather follows from the information geometry that underlies the dynamics of selection. The potential information is the dynamical potential for which selection is a gradient flow—the surface upon which selection moves the population along the steepest path (Harper 2009b, Harper and Fryer 2015). The tight connection between potential information and regret that will be developed in the following sections reinforces the fundamental role of information in understanding the dynamics and performance of natural selection.

The following sections work through the proof method outlined above to establish regret bounds for replicator dynamics.

2.C.3.1. ***Regret bound for fixed learning problems.*** First we consider a population evolving in a sequence of environments characterized by a stationary evolutionarily stable state (i.e., "fixed learning problems"). In this context, we can establish tight, finite, and converging upper and lower bounds on regret, as anticipated by Kimura. We do so first for continuous replicator dynamics, where dynamical and geometrical intuitions can be developed most naturally, before proving the analogous bounds for discrete replicator dynamics that are the focus of the main paper.

Continuous-time replicator dynamics. The main result of this section is the following theorem, which places an upper bound on the total regret of continuous-time replicator dynamics with respect to a stationary evolutionarily stable state.

Theorem 5. *For any game matrix $\boldsymbol{G}$ and for any sequence of distribution of environmental conditionss $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$ such that the population's initial type distribution $\boldsymbol{p}^0$ is in the basin of attraction of an evolutionary stable state $\overline{\boldsymbol{p}}$ that remains stationary for all $t \in [0, T]$, the total regret $\bar{R}^T$ with respect to $\overline{\boldsymbol{p}}$ of the trajectory of type distributions $\boldsymbol{p}^0, \ldots, \boldsymbol{p}^T$ generated by continuous-time replicator dynamics is bounded from above by the initial potential*

information

$$\bar{R}^T \leq D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) \quad \forall\, T\,, \tag{2.C183}$$

with equality as $T \to \infty$

$$\lim_{T\to\infty} \bar{R}^T = D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)\,. \tag{2.C184}$$

The proof of this theorem follows the logic outlined above (Equation 2.C178 - Equation 2.C181). First we relate the instantaneous regret to the instantaneous change in potential information. Then we establish the potential information $D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)$ as a non-negative and bounded potential function. The proof then follows from these propositions.

We begin by deriving a fitness-based expression for the change in potential information under replicator dynamics. This result will be handy for later steps.

Lemma 0. *Let* $\boldsymbol{p}^t$ *be the type frequency distribution of a population that evolves according to continuous-time replicator dynamics across a distribution of environmental conditions* $\boldsymbol{x}^t$ *at time* t*. Then the derivative of the potential information between* $\boldsymbol{p}^t$ *and an arbitrary type frequency distribution* $\boldsymbol{q}$ *can be expressed*

$$\frac{d}{dt} D(\boldsymbol{q}||\boldsymbol{p}^t) = \sum_{ij} \left(p_i^t - q_i\right) x_j^t r_{ij}\,, \tag{2.C185}$$

where r_{ij} *is the growth rate of type* i *in environmental condition* j*.*

PROOF.

$$\frac{d}{dt}D(\boldsymbol{q}||\boldsymbol{p}^t) = \frac{d}{dt}\sum_i q_i \log \frac{q_i}{p_i^t} \qquad \text{definition of KL divergence} \tag{2.C186}$$

$$= \frac{d}{dt}\sum_i \left(q_i \log q_i - q_i \log p_i^t\right) \tag{2.C187}$$

$$= -\sum_i q_i \frac{1}{p_i^t}\frac{dp_i^t}{dt} \qquad q_i \log q_i \text{ does not depend on } t \tag{2.C188}$$

$$= -\sum_i q_i \frac{p_i^t\left(\langle r_i\rangle^t - \langle r\rangle^t\right)}{p_i^t} \qquad \frac{dp_i^t}{dt} = p_i^t\left(\langle r_i\rangle^t - \langle r\rangle^t\right) \text{ is the continuous replicator} \tag{2.C189}$$

$$= -\sum_i q_i \left(\langle r_i\rangle^t - \langle r\rangle^t\right) \tag{2.C190}$$

$$= -\sum_i \left(q_i \langle r_i\rangle^t - q_i \langle r\rangle^t\right) \tag{2.C191}$$

$$= \sum_i q_i \langle r\rangle^t - \sum_i q_i \langle r_i\rangle^t \tag{2.C192}$$

$$= \langle r\rangle^t - \sum_i q_i \langle r_i\rangle^t \tag{2.C193}$$

$$= \sum_i p_i^t \langle r_i\rangle^t - \sum_i q_i \langle r_i\rangle^t \qquad \text{definition of } \langle r\rangle^t \tag{2.C194}$$

$$= \sum_i \left(p_i^t - q_i\right)\langle r_i\rangle^t \tag{2.C195}$$

$$= \sum_{ij} \left(p_i^t - q_i\right) x_j^t r_{ij} \qquad \text{definition of } \langle r_i\rangle^t \tag{2.C196}$$

□

Proposition 5.1. *Let $\boldsymbol{p}^t$ be the type frequency distribution of a population that evolves according to continuous-time replicator dynamics across a distribution of environmental conditions $\boldsymbol{x}^t$ at time t. Then the derivative of the potential information between $\boldsymbol{p}^t$ and an arbitrary type frequency distribution $\boldsymbol{q}$ can be expressed*

$$\frac{d}{dt}D(\boldsymbol{q}||\boldsymbol{p}^t) = -\left(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{q}, \boldsymbol{x}^t)\right), \tag{2.C197}$$

where $(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{q}, \boldsymbol{x}^t))$ is the instantaneous regret with respect to $\boldsymbol{q}$.

PROOF. We derive the following expression for the instantaneous regret in terms of fitness.

$$\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{q}, \boldsymbol{x}^t) = \sum_{ij} p_i^t x_j^t \ell(i,j) - \sum_{ij} q_i x_j^t \ell(i,j) \qquad \text{definition of expected loss} \tag{2.C198}$$

$$= \sum_{ij} p_i^t x_j^t \left(r_{*j} - r_{ij}\right) - \sum_{ij} q_i x_j^t \left(r_{*j} - r_{ij}\right) \qquad \text{definition of loss function} \tag{2.C199}$$

$$= \sum_{ij} \left(p_i^t x_j^t r_{*j} - p_i^t x_j^t r_{ij}\right) - \sum_{ij} \left(q_i x_j^t r_{*j} - q_i x_j^t r_{ij}\right) \tag{2.C200}$$

$$= \sum_{ij} p_i^t x_j^t r_{*j} - \sum_{ij} p_i^t x_j^t r_{ij} - \sum_{ij} q_i x_j^t r_{*j} + \sum_{ij} q_i x_j^t r_{ij} \tag{2.C201}$$

$$= \sum_{j} x_j^t r_{*j} - \sum_{ij} p_i^t x_j^t r_{ij} - \sum_{j} x_j^t r_{*j} + \sum_{ij} q_i x_j^t r_{ij} \qquad r_{*j} \text{ does not depend on } i \tag{2.C202}$$

$$= -\sum_{ij} p_i^t x_j^t r_{ij} + \sum_{ij} q_i x_j^t r_{ij} \tag{2.C203}$$

$$= \sum_{ij} \left(q_i - p_i^t\right) x_j^t r_{ij} \tag{2.C204}$$

$$\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{q}, \boldsymbol{x}^t) = -\sum_{ij} \left(p_i^t - q_i\right) x_j^t r_{ij} \tag{2.C205}$$

This expression for instantaneous regret is equal to the negation of the expression for the derivative of potential information that was derived as Lemma 0. Therefore we have that

$$\frac{d}{dt} D(\boldsymbol{q}||\boldsymbol{p}^t) = \sum_{ij} \left(p_i^t - q_i\right) x_j^t r_{ij} = -\left(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{q}, \boldsymbol{x}^t)\right) \qquad \text{using Lemma 0} \tag{2.C206}$$

as was to be shown. □

Proposition 5.2. *If $\overline{\boldsymbol{p}}$ is a stationary evolutionarily stable state, then the potential information $D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)$ is a Lyapunov function that satisfies for all $\boldsymbol{p}^t$ generated by continuous-time replicator dynamics in a neighborhood $\mathcal{Q}$ of $\overline{\boldsymbol{p}}$*

$$a) \qquad D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t) > 0 \quad \forall\, \boldsymbol{p}^t \in \mathcal{Q} \setminus \overline{\boldsymbol{p}} \tag{2.C207}$$

$$b) \qquad D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t) = 0 \quad \textit{iff } \boldsymbol{p}^t = \overline{\boldsymbol{p}} \tag{2.C208}$$

$$c) \qquad \frac{d}{dt} D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t) \leq 0 \quad \forall\, \boldsymbol{p}^t \in \mathcal{Q}\,, \tag{2.C209}$$

and $\overline{\boldsymbol{p}}$ is an asymptotically stable equilibrium point of the system at time t.

PROOF. Conditions (a) and (b) follow, respectively, from the facts that KL divergence is always non-negative and is equal to zero if and only if the two distributions are equal (Cover and Thomas 2006). To prove condition (c) we observe that $\boldsymbol{p}^t$ is in the neighborhood of an evolutionarily stable state $\overline{\boldsymbol{p}}$ by the premise of

the proposition, and thus we have

$$\sum_{ij} p_i^t x_j^t r_{ij} \leq \sum_{ij} \overline{p}_i x_j^t r_{ij} \qquad \text{definition of evolutionarily stable state} \tag{2.C210}$$

$$\sum_{ij} p_i^t x_j^t r_{ij} - \sum_{ij} \overline{p}_i x_j^t r_{ij} \leq 0 \tag{2.C211}$$

$$\sum_{ij} \left(p_i^t - \overline{p}_i\right) x_j^t r_{ij} \leq 0 \tag{2.C212}$$

$$\frac{d}{dt} D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t) \leq 0 \qquad \text{using Lemma 0} \tag{2.C213}$$

as was to be shown. Therefore the potential information $D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)$ satisfies the conditions of a Lyapunov function for the continuous-time replicator dynamics, the existence of which implies that $\overline{\boldsymbol{p}}$ is an *asymptotically stable equilibrium point* by the Lyapunov Asymptotic Stability Theorem. □

Corollary 5.2.1. *If $\overline{\boldsymbol{p}}$ is a stationary evolutionarily stable state, then the potential information $D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)$ is bounded from above by its initial value*

$$D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t) \leq D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) \qquad \textit{using Proposition 5.2c} \tag{2.C214}$$

Corollary 5.2.2. *If $\overline{\boldsymbol{p}}$ is a stationary evolutionarily stable state, then its basin of attraction $\mathcal{B}(\overline{\boldsymbol{p}})$ is the largest positively invariant set for which $\frac{d}{dt} D(\overline{\boldsymbol{p}}||\boldsymbol{p}) \leq 0$ under continuous-time replicator dynamics for all $\boldsymbol{p} \in \mathcal{B}(\overline{\boldsymbol{p}})$.*

Proof of Theorem 5. Proposition 5.1 relates the instantaneous change in potential information to the instantaneous regret. Considering the regret with respect to a stationary evolutionarily stable state $\overline{\boldsymbol{p}}$, we have that

$$-\left(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\overline{\boldsymbol{p}}, \boldsymbol{x}^t)\right) = \frac{d}{dt} D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)\,. \qquad \text{using Proposition 5.1} \tag{2.C215}$$

Integrating over a period of selection of duration T we obtain

$$-\int_0^T \ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\overline{\boldsymbol{p}}, \boldsymbol{x}^t) = \int_0^T \frac{d}{dt} D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t) \tag{2.C216}$$

$$-\left(L^T - L^T_{\overline{\boldsymbol{p}}}\right)\Big|_0^T = D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)\Big|_0^T \tag{2.C217}$$

$$-\bar{R}^T = D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)\Big|_0^T \qquad \text{definition of } \overline{R}^T \tag{2.C218}$$

$$-\bar{R}^T = D(\overline{\boldsymbol{p}}||\boldsymbol{p}^T) - D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) \tag{2.C219}$$

$$\bar{R}^T = D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) - D(\overline{\boldsymbol{p}}||\boldsymbol{p}^T) \tag{2.C220}$$

$$\bar{R}^T \leq D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)\,, \qquad \text{KL divergence is always non-negative, so } D(\overline{\boldsymbol{p}}||\boldsymbol{p}^T) \geq 0 \tag{2.C221}$$

which proves that the initial potential information is an upper bound on the total regret. The stationary evolutionarily stable state $\overline{\boldsymbol{p}}$ is an asymptotically stable equilibrium point, and the population's initial type distribution is within its basin of attraction as stated in the premise of the theorem. Therefore, $\boldsymbol{p}^T \to \overline{\boldsymbol{p}}$ as

$T \to \infty$, and we have that

$$\lim_{T\to\infty} \bar{R}^T = \lim_{T\to\infty} \left(D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) - D(\overline{\boldsymbol{p}}||\boldsymbol{p}^T) \right) \tag{2.C222}$$

$$\lim_{T\to\infty} \bar{R}^T = D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) - D(\overline{\boldsymbol{p}}||\overline{\boldsymbol{p}}) \tag{2.C223}$$

$$\lim_{T\to\infty} \bar{R}^T = D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)\,, \qquad \text{KL divergence is zero for equal distributions} \tag{2.C224}$$

which proves that the upper bound is tight, as the total regret converges on the initial potential information as $T \to \infty$. □

Discrete-time replicator dynamics. The main result of this section is the following theorem, which places an upper bound on the total regret of discrete-time replicator dynamics with respect to a stationary evolutionarily stable state.

Theorem 6. *For any game matrix $\boldsymbol{G}$ and for any sequence of distribution of environmental conditionss $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$ such that the population's initial type distribution $\boldsymbol{p}^0$ is in the basin of attraction of an evolutionary stable state $\overline{\boldsymbol{p}}$ that remains stationary for all $t \in [0, T]$, the total regret $\bar{R}^T$ with respect to $\overline{\boldsymbol{p}}$ of the trajectory of type distributions $\boldsymbol{p}^0, \ldots, \boldsymbol{p}^T$ generated by discrete-time replicator dynamics is bounded from above by the initial potential information*

$$\bar{R}^T \leq D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) + O(\left\langle k^2 \right\rangle^t) \quad \forall\, T\,, \tag{2.C225}$$

with equality as $T \to \infty$

$$\lim_{T\to\infty} \bar{R}^T = D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) + O(\left\langle k^2 \right\rangle^t)\,. \tag{2.C226}$$

where k_{ij} is the Wrightian selection coefficient of type i in state j (Appendix 2.A.1.1) and $\left\langle k^2 \right\rangle^t = \sum_{i,j} p_i^t x_j^t k_{ij}^2$ is the expected squared selection coefficient across types and conditions at time t.

This result is nearly identical to the corresponding bound for continuous-time replicator dynamics given in Theorem 5, save for an "error" term related to the strength of selection. This "error" can be interpreted as a small amount of additional regret that may be incurred due to slight imprecisions in discrete replicator dynamics' approximation of the continuous path of convergence to the evolutionarily stable state. When the strength is selection is high, discrete updates may carry the population further in tangential directions, and the expected imprecision of the discrete-time trajectory is greater. In the limit of weak selection, where most individuals have similar relative fitnesses and thus the selection coefficients k_{ij} are small, this imprecision is negligible and the discrete-time bound is equal to that of continuous-time replicator dynamics.

The proof of this theorem proceeds similarly to the proof of the bound for continuous replicator dynamics in the previous section. First we relate the instantaneous regret to the instantaneous change in potential information. Then we establish the potential information $D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)$ as a non-negative and bounded potential function. The proof then follows from these propositions.

Discrete-time replicator dynamics are a special case of the Multiplicative Weight Updating (MWU) algorithm where the learning rate $\eta = 1$ and a particular fitness-based loss function is used (Appendix 2.B.2.4). We begin by working through intermediate results for the general class of MWU algorithms where the learning rate η and the loss function $\ell(i, j)$ are left unspecified. This provides some insights about how the learning rate, which modulates the learning algorithm's balance of loss minimization and diversity maintenance, affects

the relationship between potential information and regret. Ultimately, Theorem 6 and the associated propositions are stated in terms of the replicator dynamics case where $\eta = 1$ and $\ell(i,j) = -\log w_{ij}$.

Proposition 6.1. *Let $\boldsymbol{p}^t$ be the type frequency distribution of a population that evolves according to discrete-time replicator dynamics over a distribution of environmental conditions $\boldsymbol{x}^t$ at time t. Then the derivative of the potential information between $\boldsymbol{p}^t$ and an arbitrary type frequency distribution $\boldsymbol{q}$ can be expressed*

$$D(\boldsymbol{q}||\boldsymbol{p}^{t+1}) - D(\boldsymbol{q}||\boldsymbol{p}^t) = -\left(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{q}, \boldsymbol{x}^t)\right) + O(\left\langle k^2 \right\rangle^t), \tag{2.C227}$$

where k_{ij} is the Wrightian selection coefficient of type i in condition j (Appendix 2.A.1.1) and $\left\langle k^2 \right\rangle^t = \sum_{i,j} p_i^t x_j^t k_{ij}^2$ is the expected squared selection coefficient across types and conditions at time t. We refer to the quantity $(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{q}, \boldsymbol{x}^t))$ as the instantaneous regret with respect to $\boldsymbol{q}$.

PROOF. We set out to evaluate an expression for the difference in the potential information across one update of the learning process.

$$D(\boldsymbol{q}||\boldsymbol{p}^{t+1}) - D(\boldsymbol{q}||\boldsymbol{p}^{t}) = \sum_i q_i \log \frac{q_i}{p_i^{t+1}} - \sum_i q_i \log \frac{q_i}{p_i^{t}} \qquad \text{definition of KL divergence} \tag{2.C228}$$

$$= \sum_i q_i \left(\log \frac{q_i}{p_i^{t+1}} - \log \frac{q_i}{p_i^{t}} \right) \tag{2.C229}$$

$$= \sum_i q_i \left(\log q_i - \log p_i^{t+1} - \log q_i + \log p_i^{t} \right) \tag{2.C230}$$

$$= \sum_i q_i \left(\log p_i^{t} - \log p_i^{t+1} \right) \tag{2.C231}$$

$$= \sum_i q_i \log \frac{p_i^{t}}{p_i^{t+1}} \tag{2.C232}$$

$$= \sum_i q_i \log \frac{p_i^{t}}{\frac{p_i^t e^{-\eta \ell(i,\boldsymbol{x}^t)}}{\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)}}} \qquad p_i^{t+1} = \frac{p_i^t e^{-\eta \ell(i,\boldsymbol{x}^t)}}{\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)}} \text{ is MWU (Equation 2.B103)} \tag{2.C233}$$

$$= \sum_i q_i \log \frac{\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)}}{e^{-\eta \ell(i,\boldsymbol{x}^t)}} \tag{2.C234}$$

$$= \sum_i q_i \left(\log \left(\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)} \right) - \log \left(e^{-\eta \ell(i,\boldsymbol{x}^t)} \right) \right) \tag{2.C235}$$

$$= \sum_i q_i \left(\log \left(\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)} \right) - \left(-\eta \ell(i,\boldsymbol{x}^t) \right) \right) \tag{2.C236}$$

$$= \sum_i q_i \left(\log \left(\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)} \right) + \eta \ell(i,\boldsymbol{x}^t) \right) \tag{2.C237}$$

$$= \sum_i \left(q_i \log \left(\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)} \right) + q_i \eta \ell(i,\boldsymbol{x}^t) \right) \tag{2.C238}$$

$$= \sum_i q_i \log \left(\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)} \right) + \eta \sum_i q_i \ell(i,\boldsymbol{x}^t) \tag{2.C239}$$

$$= \log \left(\sum_i p_i^t e^{-\eta \ell(i,\boldsymbol{x}^t)} \right) + \eta \ell(\boldsymbol{q}^t, \boldsymbol{x}^t) \qquad \text{def. of expected loss } \ell(\boldsymbol{q}^t, \boldsymbol{x}^t) \text{ (Equation 2.B59)} \tag{2.C240}$$

In order to simplify the first term further, we can invoke Jensen's inequality. Jensen's inequality states that for a *concave* function $g(u)$

$$g\left(\frac{\sum_i a_i u_i}{\sum_i a_i} \right) \geq \frac{\sum_i a_i g(u_i)}{\sum_i a_i} \,. \tag{2.C241}$$

This can be equivalently restated as

$$g\left(\frac{\sum_i a_i u_i}{\sum_i a_i}\right) = \frac{\sum_i a_i g(u_i)}{\sum_i a_i} + \xi\,, \tag{2.C242}$$

where ξ is the value of the *Jensen gap*—the difference between the left- and right-hand sides of Jensen's inequality

$$\xi = g\left(\frac{\sum_i a_i u_i}{\sum_i a_i}\right) - \frac{\sum_i a_i g(u_i)}{\sum_i a_i}\,. \tag{2.C243}$$

Using Equation 2.C242, we can express the first term in Equation 2.C240 as

$$\log\left(\sum_i p_i^t e^{-\eta\ell(i,\boldsymbol{x}^t)}\right) = \sum_i p_i^t \log\left(e^{-\eta\ell(i,\boldsymbol{x}^t)}\right) + \xi^t\,. \tag{2.C244}$$

Picking up where we left off in the derivation of an expression for the change in potential information, we have

$$\begin{aligned}
D(\boldsymbol{q}||\boldsymbol{p}^{t+1}) - D(\boldsymbol{q}||\boldsymbol{p}^t) &= \log\left(\sum_i p_i^t e^{-\eta\ell(i,\boldsymbol{x}^t)}\right) + \eta\ell(\boldsymbol{q}^t,\boldsymbol{x}^t) && \text{(2.C245)}\\
&= \left[\sum_i p_i^t \log\left(e^{-\eta\ell(i,\boldsymbol{x}^t)}\right) + \xi^t\right] + \eta\ell(\boldsymbol{q}^t,\boldsymbol{x}^t) && \text{(2.C246)}\\
&= \left[-\eta\sum_i p_i^t \ell(i,\boldsymbol{x}^t) + \xi^t\right] + \eta\ell(\boldsymbol{q}^t,\boldsymbol{x}^t) && \text{(2.C247)}\\
&= \left[-\eta\ell(\boldsymbol{p}^t,\boldsymbol{x}^t) + \xi^t\right] + \eta\ell(\boldsymbol{q}^t,\boldsymbol{x}^t) && \text{(2.C248)}\\
&= -\eta\ell(\boldsymbol{p}^t,\boldsymbol{x}^t) + \eta\ell(\boldsymbol{q}^t,\boldsymbol{x}^t) + \xi^t && \text{(2.C249)}\\
&= -\eta\left(\ell(\boldsymbol{p}^t,\boldsymbol{x}^t) - \ell(\boldsymbol{q}^t,\boldsymbol{x}^t)\right) + \xi^t && \text{(2.C250)}
\end{aligned}$$

At this point we see that the change in potential information across one update of the discrete-time learning process is proportional to (equal to for $\eta = 1$) the inverse of the instantaneous regret experienced in the most recent time step, plus a value equal to the Jensen gap ξ^t that arose in our derivation. From here, we would like to evaluate the size of this Jensen gap term. Solving for ξ^t using Equation 2.C243 and Equation 2.C244, we have

$$\begin{aligned}
\xi^t &= \log\left(\sum_i p_i^t e^{-\eta\ell(i,\boldsymbol{x}^t)}\right) - \sum_i p_i^t \log\left(e^{-\eta\ell(i,\boldsymbol{x}^t)}\right) && \text{(2.C251)}\\
&= \log\left(\sum_i p_i^t e^{-\eta\ell(i,\boldsymbol{x}^t)}\right) - \sum_i p_i^t \left(-\eta\ell(i,\boldsymbol{x}^t)\right) && \text{(2.C252)}\\
&= \log\left(\sum_i p_i^t e^{-\eta\ell(i,\boldsymbol{x}^t)}\right) + \eta\sum_i p_i^t \ell(i,\boldsymbol{x}^t)\,. && \text{(2.C253)}
\end{aligned}$$

In general, the size of this Jensen gap depends on the choice of loss function $\ell(i,j)$, but we can evaluate the size of this gap further by considering the particular loss function $\ell(i,j) = -\log w_{ij}$ and learning rate $\eta = 1$

that are implicit in the replicator dynamics instance of MWU.

$$\xi^t = \log\left(\sum_i p_i^t e^{-\eta \ell(i,\boldsymbol{x}^t)}\right) + \eta \sum_i p_i^t \ell(i,\boldsymbol{x}^t) \tag{2.C254}$$

$$= \log\left(\sum_i p_i^t e^{-\sum_j x_j^t \ell(i,j)}\right) + \sum_i p_i^t \sum_j x_j^t \ell(i,j) \qquad \text{def. of expected loss } \ell(i,\boldsymbol{x}^t) \text{ for type } i \tag{2.C255}$$

$$= \log\left(\sum_i p_i^t e^{-\sum_j x_j^t (-\log w_{ij})}\right) + \sum_i p_i^t \sum_j x_j^t \left(-\log w_{ij}\right) \qquad \ell(i,j) = -\log w_{ij} \tag{2.C256}$$

$$= \log\left(\sum_i p_i^t e^{\sum_j x_j^t \log w_{ij}}\right) - \sum_{ij} p_i^t x_j^t \log w_{ij} \tag{2.C257}$$

We can simplify the first term in this expression by making use of Jensen's inequality (Equation 2.C241) again. In particular, Jensen's inequality tells us that $\sum_j x_j^t \log w_{ij} \leq \log \sum_j x_j^t w_{ij}$, which implies that $\exp\left(\sum_j x_j^t \log w_{ij}\right) \leq \exp\left(\log \sum_j x_j^t w_{ij}\right)$.

$$\xi^t = \log\left(\sum_i p_i^t e^{\sum_j x_j^t \log w_{ij}}\right) - \sum_{ij} p_i^t x_j^t \log w_{ij} \tag{2.C258}$$

$$\xi^t \leq \log\left(\sum_i p_i^t e^{\log \sum_j x_j^t w_{ij}}\right) - \sum_{ij} p_i^t x_j^t \log w_{ij} \qquad \text{applying Jensen's inequality to exponential term} \tag{2.C259}$$

$$\xi^t \leq \log\left(\sum_i p_i^t \left(\sum_j x_j^t w_{ij}\right)\right) - \sum_{ij} p_i^t x_j^t \log w_{ij} \tag{2.C260}$$

$$\xi^t \leq \log\left(\sum_{ij} p_i^t x_j^t w_{ij}\right) - \sum_{ij} p_i^t x_j^t \log w_{ij} \tag{2.C261}$$

$$\tag{2.C262}$$

We have arrived at an inequality that bounds the Jensen gap in terms of relative fitnesses, but the scale of this bound is not particularly intuitive. We can derive a simpler expression for the magnitude of the Jensen gap by

expressing relative fitnesses in terms of Wrightian selection coefficients: $w_{ij} = 1 - k_{ij}$ (Appendix 2.A.1.1).

$$\xi^t \leq \log\left(\sum_{ij} p_i^t x_j^t w_{ij}\right) - \sum_{ij} p_i^t x_j^t \log w_{ij} \tag{2.C263}$$

$$\xi^t \leq \log\left(\sum_{ij} p_i^t x_j^t (1 - k_{ij})\right) - \sum_{ij} p_i^t x_j^t \log(1 - k_{ij}) \qquad w_{ij} = 1 - k_{ij} \tag{2.C264}$$

$$\xi^t \leq \log\left(1 - \sum_{ij} p_i^t x_j^t k_{ij}\right) - \sum_{ij} p_i^t x_j^t \log(1 - k_{ij}) \tag{2.C265}$$

$$\xi^t \leq \log\left(1 - \langle k\rangle^t\right) - \sum_{ij} p_i^t x_j^t \log(1 - k_{ij}) \tag{2.C266}$$

Applying the Taylor expansion $\log(1-u) = -\sum_{a=1}^{\infty} \frac{u^a}{a}$ to the logarithms in both terms and simplifying

$$\xi^t \leq \left(-\langle k\rangle^t - \frac{\left(\langle k\rangle^t\right)^2}{2} - \frac{\left(\langle k\rangle^t\right)^3}{3} - \dots\right) - \sum_{ij} p_i^t x_j^t \left(-k_{ij} - \frac{k_{ij}^2}{2} - \frac{k_{ij}^3}{3} - \dots\right) \tag{2.C267}$$

$$\xi^t \leq \left(-\langle k\rangle^t - \frac{\left(\langle k\rangle^t\right)^2}{2} - \frac{\left(\langle k\rangle^t\right)^3}{3} - \dots\right) - \left(-\sum_{ij} p_i^t x_j^t k_{ij} - \sum_{ij} p_i^t x_j^t \frac{k_{ij}^2}{2} - \sum_{ij} p_i^t x_j^t \frac{k_{ij}^3}{3} - \dots\right) \tag{2.C268}$$

$$\xi^t \leq \left(-\langle k\rangle^t - \frac{\left(\langle k\rangle^t\right)^2}{2} - \frac{\left(\langle k\rangle^t\right)^3}{3} - \dots\right) - \left(-\langle k\rangle^t - \frac{\langle k^2\rangle^t}{2} - \frac{\langle k^3\rangle^t}{3} - \dots\right) \tag{2.C269}$$

$$\xi^t \leq -\langle k\rangle^t - \frac{\left(\langle k\rangle^t\right)^2}{2} - \frac{\left(\langle k\rangle^t\right)^3}{3} - \dots + \langle k\rangle^t + \frac{\langle k^2\rangle^t}{2} + \frac{\langle k^3\rangle^t}{3} + \dots \tag{2.C270}$$

$$\xi^t \leq \frac{\langle k^2\rangle^t}{2} - \frac{\left(\langle k\rangle^t\right)^2}{2} + \frac{\langle k^3\rangle^t}{3} - \frac{\left(\langle k\rangle^t\right)^3}{3} + \dots \tag{2.C271}$$

$$\xi^t = O(\langle k^2\rangle^t) \tag{2.C272}$$

We find that the Jensen gap is bounded by an infinite series involving the expected values of the population's selection coefficients across types and conditions. Turning to Jensen's inequality once again gives $\langle k^2\rangle^t/2 \geq \left(\langle k\rangle^t\right)^2/2$, which tells us that this series is dominated by the first term $\langle k^2\rangle^t/2$. Thus we establish that the Jensen gap term ξ^t is on the order $O(\langle k^2\rangle^t)$ or less.

Therefore, in the case of discrete-time replicator dynamics where $\eta = 1$ and $\ell(i,j) = -\log w_{ij}$, the change in potential information across one update is given by

$$D(\boldsymbol{q}||\boldsymbol{p}^{t+1}) - D(\boldsymbol{q}||\boldsymbol{p}^t) = -\eta\left(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{q}^t, \boldsymbol{x}^t)\right) + \xi^t \qquad \text{Equation 2.C250} \tag{2.C273}$$

$$= -\left(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{q}^t, \boldsymbol{x}^t)\right) + O(\langle k^2\rangle^t) \tag{2.C274}$$

as was to be shown. □

Proposition 6.2. *If $\overline{\boldsymbol{p}}$ is a stationary evolutionarily stable state, then the potential information $D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)$ is a Lyapunov function that satisfies for all $\boldsymbol{p}^t$ generated by discrete-time replicator dynamics in a neighborhood $\mathcal{Q}$ of $\overline{\boldsymbol{p}}$*

$$a) \qquad D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t) > 0 \quad \forall\, \boldsymbol{p}^t \in \mathcal{Q} \setminus \overline{\boldsymbol{p}} \tag{2.C275}$$

$$b) \qquad D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t) = 0 \quad \textit{iff}\ \boldsymbol{p}^t = \overline{\boldsymbol{p}} \tag{2.C276}$$

$$c) \quad D(\overline{\boldsymbol{p}}||\boldsymbol{p}^{t+1}) - D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t) \leq 0 \quad \forall\, \boldsymbol{p}^t \in \mathcal{Q}\,, \tag{2.C277}$$

and $\overline{\boldsymbol{p}}$ is an asymptotically stable equilibrium point of the system at time t.

Corollary 6.2.1. *If $\overline{\boldsymbol{p}}$ is a stationary evolutionarily stable state, then the potential information $D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)$ is bounded from above by its initial value*

$$D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t) \leq D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) \qquad \textit{using Proposition 6.2c} \tag{2.C278}$$

Corollary 6.2.2. *If $\overline{\boldsymbol{p}}$ is a stationary evolutionarily stable state, then its basin of attraction $\mathcal{B}(\overline{\boldsymbol{p}})$ is the largest positively invariant set for which $D(\overline{\boldsymbol{p}}||\boldsymbol{p}^{t+1}) - D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t) \leq 0$ under discrete-time replicator dynamics for all $\boldsymbol{p} \in \mathcal{B}(\overline{\boldsymbol{p}})$.*

Proof of Theorem 6. Proposition 6.1 relates the instantaneous change in potential information to the instantaneous regret for discrete-time replicator dynamics. Considering the regret with respect to a stationary evolutionarily stable state $\overline{\boldsymbol{p}}$, we have that

$$-\left(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\overline{\boldsymbol{p}}, \boldsymbol{x}^t)\right) + O(\langle k^2 \rangle^t) = D(\overline{\boldsymbol{p}}||\boldsymbol{p}^{t+1}) - D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t) \qquad \text{Proposition 6.1} \tag{2.C279}$$

Summing over a period of selection of duration T we obtain

$$\sum_{t=0}^{T}\left[-\left(\ell(\boldsymbol{p}^t,\boldsymbol{x}^t)-\ell(\overline{\boldsymbol{p}},\boldsymbol{x}^t)\right)+O(\langle k^2\rangle^t)\right]=\sum_{t=0}^{T}\left[D(\overline{\boldsymbol{p}}||\boldsymbol{p}^{t+1})-D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)\right] \tag{2.C280}$$

$$-\sum_{t=0}^{T}\left[\ell(\boldsymbol{p}^t,\boldsymbol{x}^t)-\ell(\overline{\boldsymbol{p}},\boldsymbol{x}^t)\right]+\sum_{t=0}^{T}O(\langle k^2\rangle^t)=\sum_{t=0}^{T}\left[D(\overline{\boldsymbol{p}}||\boldsymbol{p}^{t+1})-D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)\right] \tag{2.C281}$$

$$-\left[\sum_{t=0}^{T}\ell(\boldsymbol{p}^t,\boldsymbol{x}^t)-\sum_{t=0}^{T}\ell(\overline{\boldsymbol{p}},\boldsymbol{x}^t)\right]+\sum_{t=0}^{T}O(\langle k^2\rangle^t)=\sum_{t=0}^{T}\left[D(\overline{\boldsymbol{p}}||\boldsymbol{p}^{t+1})-D(\overline{\boldsymbol{p}}||\boldsymbol{p}^t)\right] \tag{2.C282}$$

$$-\left[L^T-\bar{L}^T\right]+\sum_{t=0}^{T}O(\langle k^2\rangle^t)=D(\overline{\boldsymbol{p}}||\boldsymbol{p}^{T+1})-D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) \qquad \text{definitions of } L^T \text{ and } \bar{L}^T \tag{2.C283}$$

$$-\bar{R}^T+\sum_{t=0}^{T}O(\langle k^2\rangle^t)=D(\overline{\boldsymbol{p}}||\boldsymbol{p}^{T+1})-D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) \qquad \text{definition of regret } \bar{R}^T \tag{2.C284}$$

$$\bar{R}^T-\sum_{t=0}^{T}O(\langle k^2\rangle^t)=-D(\overline{\boldsymbol{p}}||\boldsymbol{p}^{T+1})+D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) \tag{2.C285}$$

$$\bar{R}^T=D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)-D(\overline{\boldsymbol{p}}||\boldsymbol{p}^{T+1})+\sum_{t=0}^{T}O(\langle k^2\rangle^t) \tag{2.C286}$$

$$\bar{R}^T\leq D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)+\sum_{t=0}^{T}O(\langle k^2\rangle^t) \qquad D(\overline{\boldsymbol{p}}||\boldsymbol{p}^T)\geq 0 \tag{2.C287}$$

We know that $O(\langle k^2\rangle^t)$ is decreasing monotonically when selection is approaching a stationary evolutionarily stable state, so $\sum_{t=0}^{T}O(\langle k^2\rangle^t)$ is sublinear in T. Therefore we can restate the bound on regret

$$\bar{R}^T\leq D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)+O(\langle k^2\rangle^0)\,, \tag{2.C288}$$

as was to be shown. The stationary evolutionarily stable state $\overline{\boldsymbol{p}}$ is an asymptotically stable equilibrium point, and the population's initial type distribution is within its basin of attraction as stated in the premise of the theorem. Therefore, $\boldsymbol{p}^T\to\overline{\boldsymbol{p}}$ as $T\to\infty$, and we have that

$$\lim_{T\to\infty}\bar{R}^T=D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)-D(\overline{\boldsymbol{p}}||\boldsymbol{p}^{T+1})+O(\langle k^2\rangle^0) \tag{2.C289}$$

$$\lim_{T\to\infty}\bar{R}^T=D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)-D(\overline{\boldsymbol{p}}||\overline{\boldsymbol{p}})+O(\langle k^2\rangle^0) \tag{2.C290}$$

$$\lim_{T\to\infty}\bar{R}^T=D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)+O(\langle k^2\rangle^0)\,. \qquad \text{KL divergence is zero for equal distributions} \tag{2.C291}$$

as was to be shown. □

When most individuals have similar relative fitnesses and thus the selection coefficients k_{ij} are small, the Jensen gap is also small. In the limit of weak selection where the selection coefficients k_{ij} are nearly zero for all types i in all conditions j, the Jensen gap is vanishing and the results stated in Proposition 6.1 and

Theorem 6 are equal to the corresponding results for continuous-time replicator dynamics (Proposition 5.1 and Theorem 5).

Corollary 6.1. *For any game matrix $\boldsymbol{G}$ and for any sequence of distribution of environmental conditionss $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$ such that the population's initial type distribution $\boldsymbol{p}^0$ is in the basin of attraction of an evolutionary stable state $\overline{\boldsymbol{p}}$ that remains stationary for all $t \in [0, T]$, the total regret $\bar{R}^T$ with respect to $\overline{\boldsymbol{p}}$ of the trajectory of type distributions $\boldsymbol{p}^0, \ldots, \boldsymbol{p}^T$ generated by discrete-time replicator dynamics in the limit of weak selection is bounded from above by the initial potential information*

$$\bar{R}^T \leq D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) \quad \forall\, T\,, \tag{2.C292}$$

with equality as $T \to \infty$

$$\lim_{T\to\infty} \bar{R}^T = D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)\,. \tag{2.C293}$$

2.C.3.2. ***Regret bounds for variable learning problems.*** Now we expand our consideration to the general case where no assumptions are made about the sequence of environments or stable states. The upper bounds established in this setting provide guarantees for the maximum possible regret a population can experience in any setting, including arbitrary or adversarial sequences of environments. Therefore these results describe the worst-case performance of natural selection as a learning process.

Discrete-time replicator dynamics are a special case of the Multiplicative Weights Updating (MWU) algorithm. MWU has been widely studied in computational learning theory and related fields, and various bounds have been established for this class of algorithms in different contexts. In some cases—such as when certain assumptions can be made about the learning problem, when the loss function has certain properties, or when the learning rate can be tuned in response to the learning problem—versions of MWU achieve very tight bounds on regret. In fact, it can be proven that no other learning process can possibly achieve tighter bound than some special instances of MWU. However, the instance of MWU that is equivalent to replicator dynamics does not fall into any of these special categories. Nevertheless, general regret bounds for MWU that hold for any learning rate and any loss function can be directly applied to replicator dynamics.

The main result of this section is the following theorem, which places an upper bound on the total empirical regret of discrete-time replicator dynamics in the general case.

Theorem 7. *For any game matrix $\boldsymbol{G}$ and for any sequence of distribution of environmental conditionss $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$, the total empirical regret $\widetilde{R}^T$ with respect to $\widetilde{\boldsymbol{p}}$ of the trajectory of type distributions $\boldsymbol{p}^0, \ldots, \boldsymbol{p}^T$ generated by discrete-time replicator dynamics is bounded from above by*

$$\widetilde{R}^T \leq \frac{1}{e-1}\left(\widetilde{L}^T + eD(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0)\right) \quad \forall\, T\,, \tag{2.C294}$$

where $\widetilde{L}^T$ is the cumulative loss of the empirically optimal strategy $\widetilde{\boldsymbol{p}}^T$, and where e denotes Euler's number.

This result follows from Freund and Schapire (1999), which in turn draws on the amortized analysis introduced by Kivinen and Warmuth (1993). We recite an adapted version of their proof here, which follows similar logic as those in previous sections (e.g., Equation 2.C178 - Equation 2.C181). The heart of the proof is the following proposition, which bounds the cumulative loss of the learner in terms of the loss of the empirically optimal strategy and the initial potential information. This proposition is presented as Proposition 4 in the main text.

Proposition 7.1. *For any game matrix $\boldsymbol{G}$ and for any sequence of distribution of environmental conditionss $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$, the total cumulative loss L^T of the trajectory of type distributions $\boldsymbol{p}^0, \ldots, \boldsymbol{p}^T$ generated by discrete-time replicator dynamics is bounded from above by*

$$L^T \leq \frac{e}{e-1}\left(\widetilde{L}^T + D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0)\right) \quad \forall\, T\,, \tag{2.C295}$$

where $\widetilde{L}^T$ is the cumulative loss of the empirically optimal strategy $\widetilde{\boldsymbol{p}}^T$, and where e denotes Euler's number.

Proof. As with proofs in the previous sections, we begin by seeking an expression for the change in the potential information in one update in terms of single round losses.

$$D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^{t+1}) - D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^t) = \sum_i \widetilde{p}_i^T \log \frac{\widetilde{p}_i^T}{p_i^{t+1}} - \sum_i \widetilde{p}_i^T \log \frac{\widetilde{p}_i^T}{p_i^t} \qquad \text{definition of KL divergence} \tag{2.C296}$$

$$= \sum_i \widetilde{p}_i^T \left(\log \frac{\widetilde{p}_i^T}{p_i^{t+1}} - \log \frac{\widetilde{p}_i^T}{p_i^t}\right) \tag{2.C297}$$

$$= \sum_i \widetilde{p}_i^T \left(\log \widetilde{p}_i^T - \log p_i^{t+1} - \log \widetilde{p}_i^T + \log p_i^t\right) \tag{2.C298}$$

$$= \sum_i \widetilde{p}_i^T \left(\log p_i^t - \log p_i^{t+1}\right) \tag{2.C299}$$

$$= \sum_i \widetilde{p}_i^T \log \frac{p_i^t}{p_i^{t+1}} \tag{2.C300}$$

$$= \sum_i \widetilde{p}_i^T \log \frac{p_i^t}{\frac{p_i^t e^{-\eta \ell(i,\boldsymbol{x}^t)}}{\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)}}} \qquad p_i^{t+1} = \frac{p_i^t e^{-\eta \ell(i,\boldsymbol{x}^t)}}{\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)}} \text{ is MWU (Equation 2.B103)} \tag{2.C301}$$

$$= \sum_i \widetilde{p}_i^T \log \frac{\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)}}{e^{-\eta \ell(i,\boldsymbol{x}^t)}} \tag{2.C302}$$

$$= \sum_i \widetilde{p}_i^T \left(\log \left(\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)}\right) - \log \left(e^{-\eta \ell(i,\boldsymbol{x}^t)}\right)\right) \tag{2.C303}$$

$$= \sum_i \widetilde{p}_i^T \left(\log \left(\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)}\right) - \left(-\eta \ell(i,\boldsymbol{x}^t)\right)\right) \tag{2.C304}$$

$$= \sum_i \widetilde{p}_i^T \left(\log \left(\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)}\right) + \eta \ell(i,\boldsymbol{x}^t)\right) \tag{2.C305}$$

$$= \sum_i \left(\widetilde{p}_i^T \log \left(\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)}\right) + \widetilde{p}_i^T \eta \ell(i,\boldsymbol{x}^t)\right) \tag{2.C306}$$

$$= \sum_i \widetilde{p}_i^T \log \left(\sum_k p_k^t e^{-\eta \ell(k,\boldsymbol{x}^t)}\right) + \eta \sum_i \widetilde{p}_i^T \ell(i,\boldsymbol{x}^t) \tag{2.C307}$$

$$= \log \left(\sum_i p_i^t e^{-\eta \ell(i,\boldsymbol{x}^t)}\right) + \eta \ell(\widetilde{\boldsymbol{p}}^T, \boldsymbol{x}^t) \qquad \text{def. of expected loss } \ell(\widetilde{\boldsymbol{p}}^T, \boldsymbol{x}^t) \text{ (Equation 2.B59)} \tag{2.C308}$$

Here we make use of the fact that, by convexity, $\beta^u \leq 1-(1-\beta)u$ for $\beta \geq 0$ and $u \in [0,1]$[6]. Thus if we let $\beta = e^{-\eta}$ and $u = \ell(i, \boldsymbol{x}^t)$, then $e^{-\eta\ell(i,\boldsymbol{x}^t)} \leq 1-(1-e^{-\eta})\ell(i,\boldsymbol{x}^t)$. Returning to our derivation at Equation 2.C308, we have

$$D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^{t+1}) - D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^t) = \log\left(\sum_i p_i^t e^{-\eta\ell(i,\boldsymbol{x}^t)}\right) + \eta\ell(\widetilde{\boldsymbol{p}}^T, \boldsymbol{x}^t) \tag{2.C309}$$

$$D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^{t+1}) - D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^t) \leq \log\left(\sum_i p_i^t \left(1-(1-e^{-\eta})\ell(i,\boldsymbol{x}^t)\right)\right) + \eta\ell(\widetilde{\boldsymbol{p}}^T, \boldsymbol{x}^t) \tag{2.C310}$$

$$D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^{t+1}) - D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^t) \leq \log\left(\sum_i p_i^t - (1-e^{-\eta})p_i^t\ell(i,\boldsymbol{x}^t)\right) + \eta\ell(\widetilde{\boldsymbol{p}}^T, \boldsymbol{x}^t) \tag{2.C311}$$

$$D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^{t+1}) - D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^t) \leq \log\left(1-(1-e^{-\eta})\ell(\boldsymbol{p}^t,\boldsymbol{x}^t)\right) + \eta\ell(\widetilde{\boldsymbol{p}}^T, \boldsymbol{x}^t) \qquad \text{def. of expected loss } \ell(\boldsymbol{p}^t,\boldsymbol{x}^t) \text{ (Equation 2.B59)} \tag{2.C312}$$

Now we simplify the first loss term using the fact that $\log(1-u) \leq -u$ for any $u < 1$. If we let $u = 1-(1-e^{-\eta})\ell(\boldsymbol{p}^t,\boldsymbol{x}^t)$, then we have

$$D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^{t+1}) - D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^t) \leq \log\left(1-(1-e^{-\eta})\ell(\boldsymbol{p}^t,\boldsymbol{x}^t)\right) + \eta\ell(\widetilde{\boldsymbol{p}}^T, \boldsymbol{x}^t) \tag{2.C313}$$

$$D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^{t+1}) - D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^t) \leq -(1-e^{-\eta})\ell(\boldsymbol{p}^t,\boldsymbol{x}^t) + \eta\ell(\widetilde{\boldsymbol{p}}^T, \boldsymbol{x}^t) \tag{2.C314}$$

$$D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^{t+1}) - D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^t) \leq \eta\ell(\widetilde{\boldsymbol{p}}^T, \boldsymbol{x}^t) - (1-e^{-\eta})\ell(\boldsymbol{p}^t,\boldsymbol{x}^t) \tag{2.C315}$$

By summing this inequality over a period of duration T and rearranging terms we obtain

$$\sum_{t=0}^{T}\left[D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^{t+1}) - D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^t)\right] \leq \sum_{t=0}^{T}\left[\eta\ell(\widetilde{\boldsymbol{p}}^T, \boldsymbol{x}^t) - (1-e^{-\eta})\ell(\boldsymbol{p}^t,\boldsymbol{x}^t)\right] \tag{2.C316}$$

$$\sum_{t=0}^{T}\left[D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^{t+1}) - D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^t)\right] \leq \eta\sum_{t=0}^{T}\ell(\widetilde{\boldsymbol{p}}^T, \boldsymbol{x}^t) - (1-e^{-\eta})\sum_{t=0}^{T}\ell(\boldsymbol{p}^t,\boldsymbol{x}^t) \tag{2.C317}$$

$$D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^{T+1}) - D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0) \leq \eta\widetilde{L}^T - (1-e^{-\eta})L^T \tag{2.C318}$$

$$(1-e^{-\eta})L^T + D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^{T+1}) \leq \eta\widetilde{L}^T + D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0) \tag{2.C319}$$

$$(1-e^{-\eta})L^T \leq \eta\widetilde{L}^T + D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0) \qquad D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^{T+1}) > 0 \tag{2.C320}$$

$$L^T \leq \frac{\eta}{1-e^{-\eta}}\widetilde{L}^T + \frac{1}{1-e^{-\eta}}D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0)\,. \tag{2.C321}$$

The learning rate implicit in the equivalence between replicator dynamics and MWU is $\eta = 1$. Plugging in this value for η we have

$$L^T \leq \frac{1}{1-e^{-1}}\widetilde{L}^T + \frac{1}{1-e^{-1}}D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0) \tag{2.C322}$$

$$L^T \leq \frac{1}{1-e^{-1}}\left(\widetilde{L}^T + D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0)\right) \tag{2.C323}$$

$$L^T \leq \frac{e}{e-1}\left(\widetilde{L}^T + D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0)\right)\,, \tag{2.C324}$$

as was to be shown. □

[6]Note that any loss function can be scaled to the range $[0, 1]$ without loss of generality.

Proof of Theorem 7. The bound on cumulative load given in Proposition 7.1 can be rearranged to express a bound on empirical regret

$$L^T \leq \frac{e}{e-1}\left(\widetilde{L}^T + D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0)\right) \tag{2.C325}$$

$$L^T \leq \frac{e}{e-1}\widetilde{L}^T + \frac{e}{e-1}D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0) \tag{2.C326}$$

$$L^T \leq \left(\frac{e}{e-1} - 1 + 1\right)\widetilde{L}^T + \frac{e}{e-1}D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0) \tag{2.C327}$$

$$L^T \leq \left(\frac{e}{e-1} - 1\right)\widetilde{L}^T + \widetilde{L}^T + \frac{e}{e-1}D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0) \tag{2.C328}$$

$$L^T - \widetilde{L}^T \leq \frac{1}{e-1}\widetilde{L}^T + \frac{e}{e-1}D(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0) \tag{2.C329}$$

$$L^T - \widetilde{L}^T \leq \frac{1}{e-1}\left(\widetilde{L}^T + eD(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0)\right) \tag{2.C330}$$

$$\widetilde{R}^T \leq \frac{1}{e-1}\left(\widetilde{L}^T + eD(\widetilde{\boldsymbol{p}}^T||\boldsymbol{p}^0)\right), \tag{2.C331}$$

as was to be shown. □

2.C.4. The cost of information gain.

2.C.4.1. ***Information gain in a single generation is bounded by look-ahead regret.*** The following result relates the information gained by selection in a single generation (as measured by the divergence $D(\boldsymbol{p}^{t+1}||\boldsymbol{p}^t)$) to the expected fitness loss due to type mismatch in the current and next generations.

Proposition 7.2. *Let $\mathbf{p}^t$ be the type frequency distribution of a population that evolves according to discrete-time replicator dynamics over a distribution of environmental conditions $\boldsymbol{x}^t$ at time t. Then the information gain in one selective update, as measured by the divergence between the population's updated and previous compositions, can be expressed*

$$D(\boldsymbol{p}^{t+1}||\boldsymbol{p}^t) = \left(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{p}^{t+1}, \boldsymbol{x}^t)\right) + O(\left\langle k^2\right\rangle^t) \tag{2.C332}$$

$$= \sum_{i,j}\left(p_i^t - p_i^{t+1}\right)x_j^t\left(\log W_{*j} - \log W_{ij}\right) + O(\left\langle k^2\right\rangle^t), \tag{2.C333}$$

where k_{ij} is the Wrightian selection coefficient of type i in condition j (Appendix 2.A.1.1) and $\left\langle k^2\right\rangle^t = \sum_{i,j} p_i^t x_j^t k_{ij}^2$ is the expected squared selection coefficient across types and conditions at time t.

PROOF. This result follows directly from Proposition 6.1 where the reference distribution $\boldsymbol{q}$ is taken to be the type frequency distribution in the next generation $\boldsymbol{p}^{t+1}$.

$$D(\boldsymbol{q}||\boldsymbol{p}^{t+1}) - D(\boldsymbol{q}||\boldsymbol{p}^{t}) = -\left(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{q}, \boldsymbol{x}^t)\right) + O(\left\langle k^2 \right\rangle^t) \qquad \text{Proposition 6.1} \tag{2.C334}$$

$$D(\boldsymbol{p}^{t+1}||\boldsymbol{p}^{t+1}) - D(\boldsymbol{p}^{t+1}||\boldsymbol{p}^{t}) = -\left(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{p}^{t+1}, \boldsymbol{x}^t)\right) + O(\left\langle k^2 \right\rangle^t) \qquad \text{Let } \boldsymbol{q} = \boldsymbol{p}^{t+1} \tag{2.C335}$$

$$-D(\boldsymbol{p}^{t+1}||\boldsymbol{p}^{t}) = -\left(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{p}^{t+1}, \boldsymbol{x}^t)\right) + O(\left\langle k^2 \right\rangle^t) \qquad \text{KL divergence is zero for equal distributions} \tag{2.C336}$$

$$D(\boldsymbol{p}^{t+1}||\boldsymbol{p}^{t}) = \left(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{p}^{t+1}, \boldsymbol{x}^t)\right) + O(\left\langle k^2 \right\rangle^t)\,, \tag{2.C337}$$

as was to be shown. □

The following corollary is presented as Proposition 5 in the main text.

Corollary 7.2.1.

$$D(\boldsymbol{p}^{t+1}||\boldsymbol{p}^{t}) = \ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{p}^{t+1}, \boldsymbol{x}^t) \tag{2.C338}$$

$$= \sum_{i,j} \left(p_i^t - p_i^{t+1}\right) x_j^t \left(\log W_{*j} - \log W_{ij}\right)\,. \tag{2.C339}$$

2.C.4.2. ***Information gain converges on ESS regret.*** The result of this section is the following theorem, which establishes that the evolving population's total information gain converges on the total ESS regret of discrete-time replicator dynamics in contexts characterized by a stationary ESS.

Theorem 8. *For any game matrix $\boldsymbol{G}$ and for any sequence of environmental conditions $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$ such that the population's initial type distribution $\boldsymbol{p}^0$ is in the basin of attraction of an evolutionary stable state $\overline{\boldsymbol{p}}$ that remains stationary for all $t \in [0, T]$, the total information gain I^T and the total regret $\bar{R}^T$ with respect to $\overline{\boldsymbol{p}}$ of the trajectory of type distributions $\boldsymbol{p}^0, \ldots, \boldsymbol{p}^T$ generated by replicator dynamics in the limit of weak selection both converge on the value of the initial potential information*

$$\lim_{T\to\infty} I^T = \lim_{T\to\infty} \bar{R}^T = D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)\,. \tag{2.C340}$$

PROOF. The stationary evolutionarily stable state $\overline{\boldsymbol{p}}$ is an asymptotically stable equilibrium point (Proposition 6.2), and the population's initial type distribution is within its basin of attraction as stated in the premise of the theorem. Therefore, $\boldsymbol{p}^T \to \overline{\boldsymbol{p}}$ as $T \to \infty$, and Corollary 6.1 of Theorem 6 tells us that

$$\lim_{T\to\infty} \bar{R}^T = D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0)\,, \tag{2.C341}$$

in the limit of weak selection (i.e., $k_{ij} \lll 1$ for all i, j) or continuously overlapping generations (Theorem 5).

The convergence of the information gain I^T on the same value—the initial potential information—follows from the definition of information gain and the convergence of $\boldsymbol{p}^T \to \overline{\boldsymbol{p}}$.

$$I^T = D(\boldsymbol{p}^T||\boldsymbol{p}^0) \qquad \text{definition of } I^T \tag{2.C342}$$

$$\lim_{T\to\infty} I^T = \lim_{T\to\infty} D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) \tag{2.C343}$$

$$\lim_{T\to\infty} I^T = D(\overline{\boldsymbol{p}}||\boldsymbol{p}^0) \qquad \text{using } \boldsymbol{p}^T \to \overline{\boldsymbol{p}} \text{ as } T \to \infty\,. \tag{2.C344}$$

Combining Equation 2.C341 and Equation 2.C344, we obtain the stated result. □

2.C.4.3. ***Information gain is bounded by empirical regret.*** The main result of this section is the following theorem, which establishes that the evolving population's information gain is always bounded by the empirical regret. This bound is fully general and holds in all environmental contexts.

Theorem 9. *For any game matrix $\boldsymbol{G}$ and for any sequence of distribution of environmental conditionss $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$, the total information gain $D(\boldsymbol{p}^T||\boldsymbol{p}^0)$ of the trajectory of type distributions $\boldsymbol{p}^0, \ldots, \boldsymbol{p}^T$ generated by discrete-time replicator dynamics is at all times bounded from above by the empirical regret*

$$D(\boldsymbol{p}^T||\boldsymbol{p}^0) \leq \widetilde{R}^T \; \forall \, T \, . \tag{2.C345}$$

Proof. Proposition 6.1 relates the instantaneous change in potential information to the instantaneous regret with respect to some reference distribution $\boldsymbol{q}$ for discrete-time replicator dynamics.

$$D(\boldsymbol{q}||\boldsymbol{p}^{t+1}) - D(\boldsymbol{q}||\boldsymbol{p}^t) = -\left(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{q}, \boldsymbol{x}^t)\right) + O(\langle k^2 \rangle^t) \qquad \text{Proposition 6.1} \tag{2.C346}$$

We will work from this result to establish the relationship between the total information gain and the total empirical regret. Summing over a period of selection of duration T we obtain

$$\sum_{t=0}^{T} \left[D(\boldsymbol{q}||\boldsymbol{p}^{t+1}) - D(\boldsymbol{q}||\boldsymbol{p}^t)\right] = \sum_{t=0}^{T} \left[-\left(\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{q}, \boldsymbol{x}^t)\right) + O(\langle k^2 \rangle^t)\right] \tag{2.C347}$$

$$\sum_{t=0}^{T} \left[D(\boldsymbol{q}||\boldsymbol{p}^{t+1}) - D(\boldsymbol{q}||\boldsymbol{p}^t)\right] = -\sum_{t=0}^{T} \left[\ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \ell(\boldsymbol{q}, \boldsymbol{x}^t)\right] + \sum_{t=0}^{T} O(\langle k^2 \rangle^t) \tag{2.C348}$$

$$\sum_{t=0}^{T} \left[D(\boldsymbol{q}||\boldsymbol{p}^{t+1}) - D(\boldsymbol{q}||\boldsymbol{p}^t)\right] = -\left[\sum_{t=0}^{T} \ell(\boldsymbol{p}^t, \boldsymbol{x}^t) - \sum_{t=0}^{T} \ell(\boldsymbol{q}, \boldsymbol{x}^t)\right] + \sum_{t=0}^{T} O(\langle k^2 \rangle^t) \tag{2.C349}$$

$$D(\boldsymbol{q}||\boldsymbol{p}^{T+1}) - D(\boldsymbol{q}||\boldsymbol{p}^0) = -\left[L^T - L^T_{\boldsymbol{q}}\right] + \sum_{t=0}^{T} O(\langle k^2 \rangle^t) \, . \qquad \text{definitions of } L^T \text{ and } L^T_{\boldsymbol{q}} \tag{2.C350}$$

Let us evaluate with respect to the reference distribution $\boldsymbol{q} = \boldsymbol{p}^{T+1}$.

$$D(\boldsymbol{p}^{T+1}||\boldsymbol{p}^{T+1}) - D(\boldsymbol{p}^{T+1}||\boldsymbol{p}^0) = -\left[L^T - L^T_{\boldsymbol{p}^{T+1}}\right] + \sum_{t=0}^{T} O(\langle k^2 \rangle^t) \tag{2.C351}$$

$$-D(\boldsymbol{p}^{T+1}||\boldsymbol{p}^0) = -\left[L^T - L^T_{\boldsymbol{p}^{T+1}}\right] + \sum_{t=0}^{T} O(\langle k^2 \rangle^t) \qquad \text{KL divergence is zero for equal distributions} \tag{2.C352}$$

$$D(\boldsymbol{p}^{T+1}||\boldsymbol{p}^0) = L^T - L^T_{\boldsymbol{p}^{T+1}} - \sum_{t=0}^{T} O(\langle k^2 \rangle^t) \tag{2.C353}$$

$$D(\boldsymbol{p}^{T+1}||\boldsymbol{p}^0) \leq L^T - L^T_{\boldsymbol{p}^{T+1}} \, . \tag{2.C354}$$

So far we have established that the information gain through $T+1$ generations is bounded by the cumulative loss of the evolving population minus the cumulative loss of the fixed reference distribution $\boldsymbol{p}^{T+1}$.

At this point, we recall that the empirically optimal strategy $\widetilde{\boldsymbol{p}}^T$ is defined as the fixed strategy with the minimum possible cumulative loss over the observed sequence of environments $\boldsymbol{x}^0 x^T$ in hindsight

$$\widetilde{\boldsymbol{p}}^T = \arg\min_{\boldsymbol{q}} L^T_{\boldsymbol{q}} = \arg\min_{\boldsymbol{q}} \sum_{t=0}^{T} \ell(\boldsymbol{p}^t, \boldsymbol{q}) \,. \tag{2.C355}$$

Therefore the cumulative loss of the empirically optimal strategy $\widetilde{L}^T$ is less or equal to the cumulative loss of any other fixed strategy by definition

$$\widetilde{L}^T \leq L^T_{\boldsymbol{q}} \; \forall \; \boldsymbol{q} \,, \tag{2.C356}$$

which implies that the empirically optimal cumulative loss is less than or equal to the cumulative loss of the fixed reference distribution $\boldsymbol{p}^{T+1}$

$$\widetilde{L}^T \leq L^T_{\boldsymbol{p}^{T+1}} \,. \tag{2.C357}$$

Using this fact with Equation 2.C354, we obtain

$$D(\boldsymbol{p}^{T+1}||\boldsymbol{p}^0) \leq L^T - L^T_{\boldsymbol{p}^{T+1}} \leq L^T - \widetilde{L}^T \,, \qquad \text{using } \widetilde{L}^T \leq L^T_{\boldsymbol{p}^{T+1}} \tag{2.C358}$$

and therefore

$$D(\boldsymbol{p}^{T+1}||\boldsymbol{p}^0) \leq L^T - \widetilde{L}^T \tag{2.C359}$$

$$D(\boldsymbol{p}^{T+1}||\boldsymbol{p}^0) \leq \widetilde{R}^T \,, \qquad \text{defintion of emprical regret } \widetilde{R}^T = L^T - \widetilde{L}^T \tag{2.C360}$$

as was to be shown. □

Information gain is bounded by substitution load. Substitution load is a special case of mismatch load that applies when the environment consists of a single unchanging condition (i.e., $m = 1$). In such a case, a single type has the highest fitness for all time and sweeps to fixation. Fixation of the optimal type is the empirically optimal strategy $\widetilde{\boldsymbol{p}}^T$ for all T, and the empirically optimal cumulative loss $\widetilde{L}^T = 0$ for all T. Therefore substitution load is equivalent to empirical regret $\widetilde{R}^T$ in applicable cases, and it follows from Theorem 9 that the population's information gain is always bounded by substitution load in such cases.

Corollary 9.1. *If the environment consists of a single unchanging environmental condition ($m = 1$), then the total information gain $D(\boldsymbol{p}^T||\boldsymbol{p}^0)$ of the trajectory of type distributions $\mathbf{p}^0, \ldots, \mathbf{p}^T$ generated by discrete-time replicator dynamics is at all times bounded from above by the substitution load*

$$D(\boldsymbol{p}^T||\boldsymbol{p}^0) \leq L^T_{\text{sub}} \; \forall \; T \,. \tag{2.C361}$$

Proof.

$$D(\boldsymbol{p}^T||\boldsymbol{p}^0) \leq \widetilde{R}^T \qquad \text{Theorem 9} \tag{2.C362}$$

$$D(\boldsymbol{p}^T||\boldsymbol{p}^0) \leq L^T - \widetilde{L}^T \qquad \text{definition of empirical regret} \tag{2.C363}$$

$$D(\boldsymbol{p}^T||\boldsymbol{p}^0) \leq L^T_{\text{sub}} \,. \qquad L^T = L^T_{\text{sub}} \text{ and } \widetilde{L}^T = 0 \text{ where } m = 1 \tag{2.C364}$$

□

Information gain is bounded by mismatch load in general. Another notable corollary of Theorem 9 is that the information gain of a population that evolves by natural selection is always bounded by the mismatch load in the general case.

Corollary 9.2. *For any game matrix $\boldsymbol{G}$ and for any sequence of distribution of environmental conditionss $\boldsymbol{x}^0, \ldots, \boldsymbol{x}^T$, the total information gain $D(\boldsymbol{p}^T||\boldsymbol{p}^0)$ of the trajectory of type distributions $\boldsymbol{p}^0, \ldots, \boldsymbol{p}^T$ generated by discrete-time replicator dynamics is at all times bounded from above by the empirical regret*

$$D(\boldsymbol{p}^T||\boldsymbol{p}^0) \leq L^T \; \forall \, T \, . \tag{2.C365}$$

Proof.

$$D(\boldsymbol{p}^T||\boldsymbol{p}^0) \leq \widetilde{R}^T \qquad \text{Theorem 9} \tag{2.C366}$$

$$D(\boldsymbol{p}^T||\boldsymbol{p}^0) \leq L^T - \widetilde{L}^T \qquad \text{definition of empirical regret} \tag{2.C367}$$

$$D(\boldsymbol{p}^T||\boldsymbol{p}^0) \leq L^T \, . \qquad \text{cumulative loss } \widetilde{L}^T \text{ is non-negative} \tag{2.C368}$$

□

Appendix 2.D. *Selection experiments.*

	Chromosomal Genotype			Plasmid			
Strain ID	*E. coli* strain	Mutations	Sel. Markers	Plasmid ID	Backbone	FP Gene Insert	Sel. Markers
WT	B (REL606)	*See Table 2.D.5*	-	pRMbcGFP	pBR322	pEB1-mGFPmut2	Tet^R
M1	B (REL606)	*See Table 2.D.5*	Rif^R	pRMbc1RFP	pBR322	pEB2-mScarlet-I	Tet^R
M2	B (REL606)	*See Table 2.D.5*	Rif^R	pRMbc2RFP	pBR322	pEB2-mScarlet-I	Tet^R
M3	B (REL606)	*See Table 2.D.5*	Rif^R	pRMbc3RFP	pBR322	pEB2-mScarlet-I	Tet^R

TABLE 2.D.5. Basic characteristics of bacterial strains used in selection experiments.

WT	M1	M2	M3	Position	Gene	Mutation	Annotation	Description
×	×	×	×	3,925,216	*gppA* ←	C→T	L272L (CTG→CTA)	Guanosine pentaphosphatase
×				458,185	*acrR* →	IS1 (+) +9 bp	Coding (358 366/648 nt)	Transcriptional repressor
	×			4,162,387	*rpoB* →	T→G	L511R (CTG→CGG)	RNA polymerase subunit β
		×		4,161,298	*rpoB* →	A→T	Q148L (CAG→CTG)	RNA polymerase subunit β
		×		4,162,382	*rpoB* →	C→A	S509R (AGC→AGA)	RNA polymerase subunit β
			×	4,161,292	*rpoB* →	T→A	V146D (GTT→GAT)	RNA polymerase subunit β
			×	4,162,540	*rpoB* →	A→C	E562A (GAA→GCA)	RNA polymerase subunit β

TABLE 2.D.5. Mutations present in each bacterial strain relative to the *Escherichia coli* B (REL606) reference genome (presence of a mutation in a given strain is indicated by an × in that strain's column).

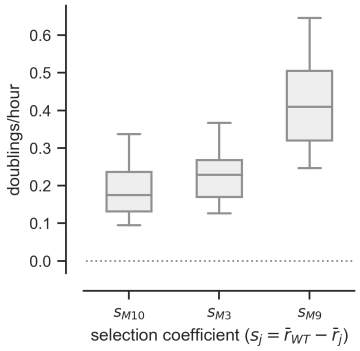


FIGURE 2.D.6. **Selection coefficients.** The selection coefficient is a measure of the fitness of a type (i.e., strain) relative to another. In Malthusian fitness terms, the selection coefficient of a type j is defined as the difference between the exponential growth rate of the optimal type and that of the jth strain; that is, $s_j = r_* - r_j$ (Appendix 2.A.1.2). In our experimental system, the WT strain had the fastest growth rate. Therefore the selection coefficient of a strain j was defined as $s_j = r_{\mathrm{WT}} - r_j$. A positive selection coefficient $s_j > 0$ indicates that the growth rate of the jth strain is less than that of the WT strain. By definition, the selection coefficient of the WT strain is zero ($s_{\mathrm{WT}} = 0$). Here we show the selection coefficients of the mutant strains (M1, M2, M3) relative to the WT strain as measured by an exponential growth assay. Individual strain cultures were sampled and transferred to fresh media to maintain exponential growth every 2 hours for a total of 8 hours. Cell enumeration was performed on culture samples using flow cytometry, and strain growth rates were calculated from changes in culture density over the course of the assay. Selection coefficients were then calculated according to the definition given above.

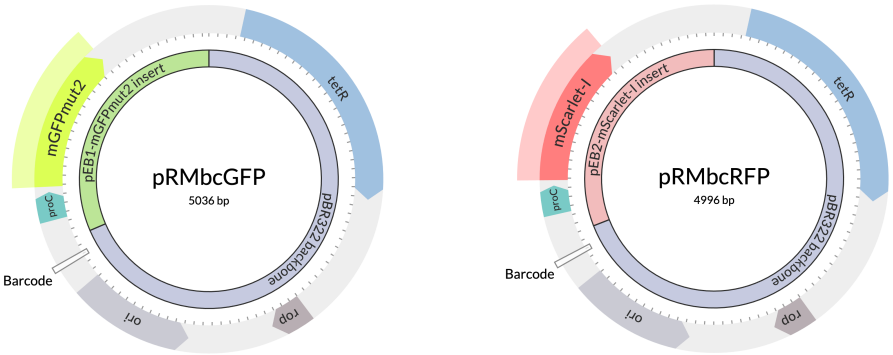


FIGURE 2.D.7. **Engineered plasmid maps.** Each strain was transformed with an engineered marker plasmid. The pBR322 plasmid, which carries a *bla* gene conferring ampicillin resistance (Amp^R) and a *tetA* gene conferring tetracycline resistance (Tet^R), served as a vector backbone. Using Gibson assembly, the *bla* gene and corresponding promoter region was removed and replaced with an insert carrying a fluorescent protein gene under a strong constitutive *proC* promoter. **Left:** The pRMbcGFP plasmid carries a region encoding the green fluorescent protein gene *mGFPmut2* under the *proC* promoter, which was taken from the pEB1-mGFPmut2 plasmid (Addgene plasmid #103980). **Right:** The pRMbcRFP plasmids carry a region encoding the red fluorescent protein gene *mScarlet-I* under the *proC* promoter, which was taken from the pEB2-mScarlet-I plasmid (Addgene plasmid #104007).

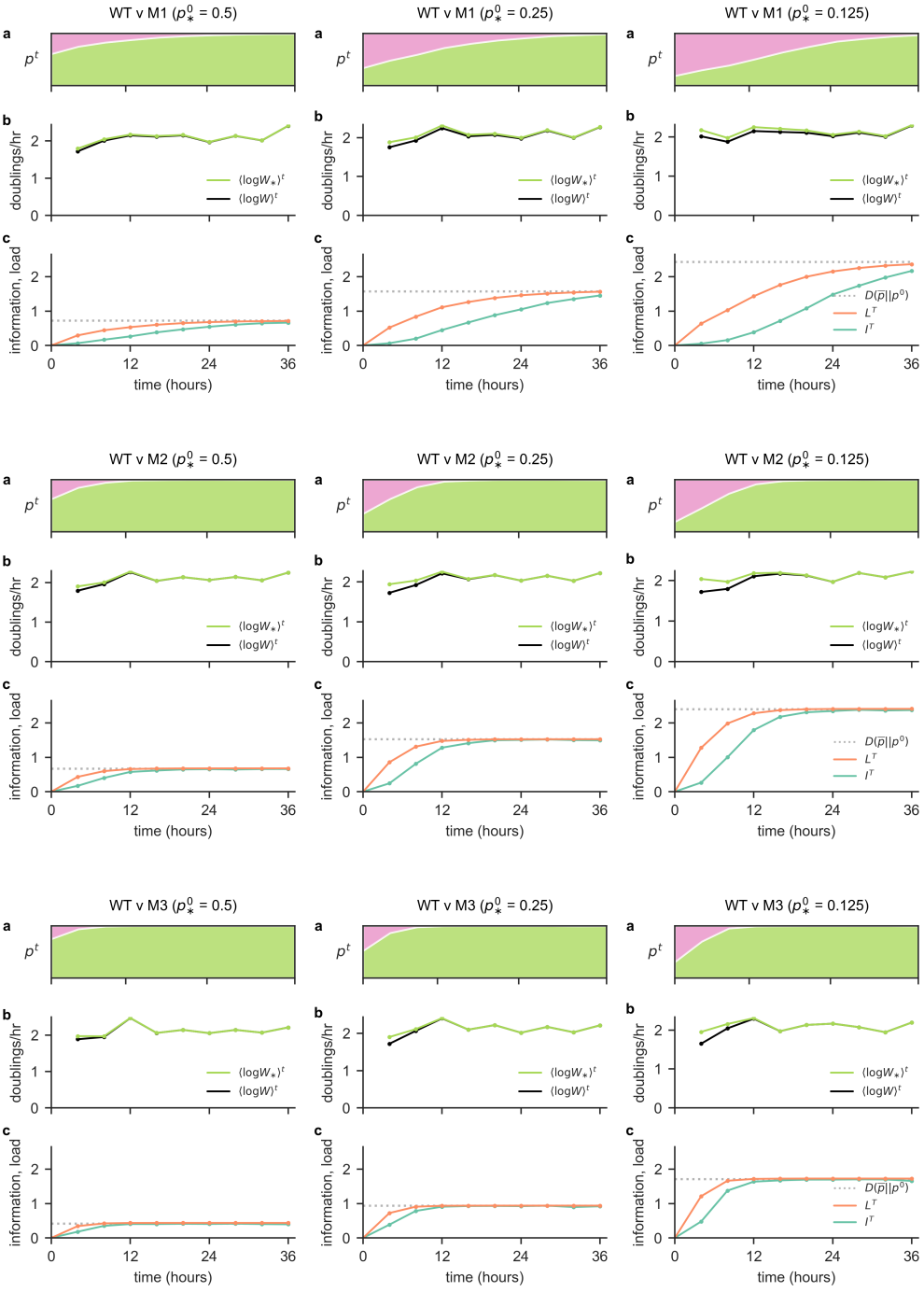


FIGURE 2.D.8. **Supplemental results from selection experiments.** Here we show the changes in strain frequencies (**a** plots) and growth rates (**b** plots), and information and load measures (**c** plots) over time for each of the 9 selection competitions presented in main text Figure 2.3. The title of each sub-panel indicates the strain combination and initial frequency of the optimal WT strain in the respective competition.

2.D.1. Supplemental methods.

2.D.1.1. ***Transfer protocol.*** Selection experiments involved incubating mixed batch cultures and tracking the densities and frequencies of the competing strains for a total of 36 hours, which was the time necessary to observe the optimal WT strain approaching fixation in all competitions. These strains enter a stationary phase as the culture approaches a saturation density of approximately 1×10^9 cfus/mL, so regular transfers to fresh media were necessary to maintain active growth and selection over this duration. Basic transfer protocols often involve repeatedly growing cultures to saturation and transfering a sample of the saturated culture to fresh media at low density (e.g., ~1×10^5 cfus/mL). However, we were interested in measuring the outcomes of selection due to differential exponential growth rates, and thus we wished to maintain cultures in exponential phase to avoid the potential confounding effects of cells entering and exiting lag phase. In addition, highly accurate density estimation using flow cytometry required a culture density of at least ~1×10^7 cfus/mL at the time of sampling. Therefore, in order to ensure that the culture would remain in exponential growth while maintaining a measurable density for the flow cytometer we required a transfer protocol that would reliably maintain culture densities between ~$1 \times 10^7 - 1 \times 10^8$ cfus/mL at the end of each growth interval. The following transfer protocol was developed to dynamically calculate the transfer volume for each competition culture in order to achieve this (note that all competitions were conducted simultaneously).

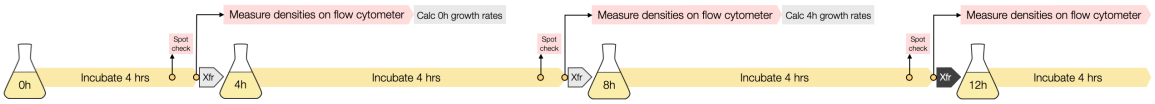


FIGURE 2.D.9. **Illustration of transfer protocol.**

To make things concrete, we outline the transfer protocol from the perspective of an experimenter determining transfer volumes at the 12 hour mark of the overall experiment (black-shaded block arrow in Figure 2.D.9), although the same process was used at all transfer points throughout the 36 hour experiment. Flasks are labeled by the time point at which they were inoculated (e.g., the '0h' culture flask was inoculated at the beginning of the experiment, the '4h' culture flask was inoculated via transfer at the 4 hour mark, and so on).

- **4 hours prior to transfer:** Data collection samples are taken from each competition, and competition cultures are transferred to fresh media.
  (1) Samples are taken from the previous set of culture flasks (e.g., the '4h' cultures) for density estimation on the flow cytometer.
  (2) The current set of flasks (e.g., the '8h' cultures) are inoculated via transfer and their incubation period begins.
  (3) Cytometry samples are prepared and measured over the following 1-2 hours.
  (4) The newly obtained cytometry measurements provide endpoint density estimates for the the previous set of culture flasks (e.g., the '4h' cultures). The initial density of these same cultures is estimated from the endpoint density of the prior set of cultures (e.g., obtained from cytometry measurements of the '0h' cultures) and the transfer volumes used at that time. The growth rates of each of the previous cultures (e.g., the '4h' cultures) is estimated using the initial and endpoint densities for this interval (e.g., the 4h-8h interval).

- **30 minutes prior to transfer:** A "spot check" is performed on a subset of the current flasks in order to obtain a rough estimate of the densities and growth rates of the currently incubating cultures.
  (1) Samples are quickly taken from the competition cultures with the lowest growth rate and the highest growth rate observed in the latest set of cytometry-based growth rate estimates (e.g., for the '4h' cultures). These flasks are immediately returned to the incubator to finish the rest of the incubation period.
  (2) These samples are enumerated on the flow cytometer and used to estimate the endpoint densities and growth rates of the presumed fastest and slowest growing competition cultures in the current incubation period. This provides a presumed minimum and maximum growth rate for all competitions in the current growth interval.
  (3) The endpoint densities for all competition cultures in the current growth interval are projected based on the initial densities of these cultures (using the previous cultures' endpoint densities and the transfer volumes involved in their inoculations) and the average of the growth rates estimated from the spot check measurement described above.
- **10 minutes prior to transfer:** The transfer volume to be used for each competition culture is calculated as follows (a spreadsheet was programmed to perform these calculations automatically using flow cytometry data collected previously in the experiment):
  – For each competition culture, working in order of highest projected endpoint density (from #3 above) to lowest:
    (1) If this is the first competition considered, let the "trial" transfer volume for this competition culture be a default volume of 250 μL (i.e., a 100-fold dilution into 25 mL of media); else, let the "trial" transfer volume be the actual transfer volume of the previously considered competition.
    (2) Calculate the projected initial density of the post-transfer culture based on the projected final density of the current culture and the "trial" transfer volume.
    (3) Calculate the projected endpoint density of this competition culture in the next incubation period based on the projected initial density and the *minimum* growth rate that has been observed for any culture so far. Likewise, calculate the projected endpoint density of this competition culture in the next incubation period using the *maximum* growth rate that has been observed for any culture so far. This gives the range of plausible endpoint densities the post-transfer culture may realize if the "trial" transfer volume is used.
    (4) If the range of plausible endpoint densities using the "trial" transfer volume falls within the acceptable range of endpoint densities (i.e., $1 \times 10^7 - 1 \times 10^8$ cfus/mL), then use the "trial" transfer volume to transfer this competition culture; else, do the following:
      * Backcalculate the initial density that would lead this culture to just reach the *minimum* acceptable endpoint density (i.e., 1×10^7 cfus/mL) if the culture were to grow at the *minimum* growth rate that has been observed for any culture so far. This is the smallest post-transfer initial density that is ensured to reach the minimum acceptable endpoint density in a worst-case, slow growth scenario (based on the range of growth rates observed to this point). (This heuristic is used because we deemed

it more important to guarantee that cultures reach accurately measurable densities than to guarantee that cultures do not approach saturation.)

- Calculate the transfer volume that would give this initial density based on the projected endpoint density for this culture in the current incubation period.
- Use this "backcalculated" transfer volume to transfer this competition culture.

- **At time of transfer:** Data collection samples are taken from each competition, and competition cultures are transferred to fresh media using the calculated transfer volumes. The cycle repeats for the following incubation period and transfer.

This transfer protocol successfully maintained endpoint densities for all competition cultures within the range $8.3 \times 10^6 - 2.7 \times 10^8$ cfus/mL throughout the entire 36 hour experiment, with only a few culture endpoints falling outside of the tight target range of $1 \times 10^7 - 1 \times 10^8$ cfus/mL. All culture samples had densitites that were accurately measurable on the flow cytometer, and no competition culture showed signs of cells exiting exponential phase to any appreciable degree.

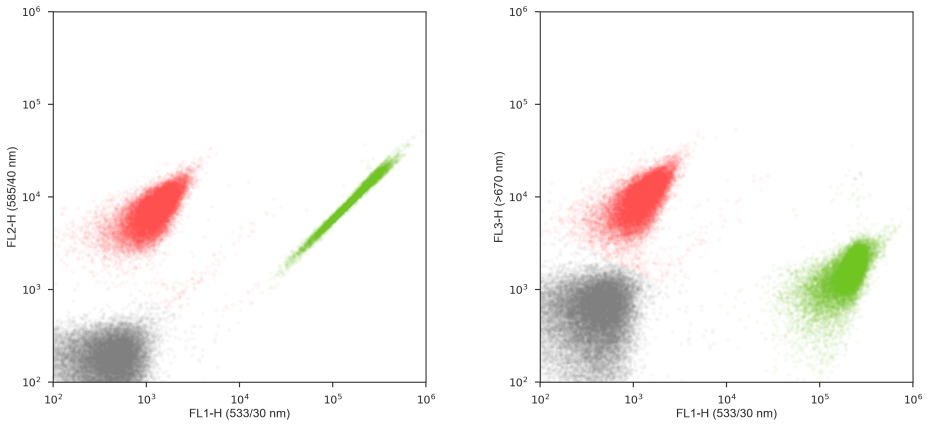


FIGURE 2.D.10. **Example flow cytometry data.** Culture densities were estimated with cell enumeration using a BD Accuri C6 Flow Cytometer. These plots present example data collected from a single sample taken from a mixed culture of GFP-marked WT cells and RFP-marked M1 cells. The x- and y-axes refer to fluorescence intensities in the wavelength bands indicated in parentheses. Each point represents the fluorescence intensities of a particle drawn from the culture sample. Three distinct clusters of points are immediately apparent, and fixed fluorescence gates were established to automatically differentiate GFP-marked cells, RFP-marked cells, and other particles. Cells that fall within the RFP gate are colored red above, cells that fall within the GFP gate are colored green, and other particles that fall outside of both FP gate regions are colored gray. Such data provides counts of GFP-marked cells, RFP-marked cells, and total cells, which can be used to calculate the strain densities and frequencies in the culture from which the cytometry sample was taken.

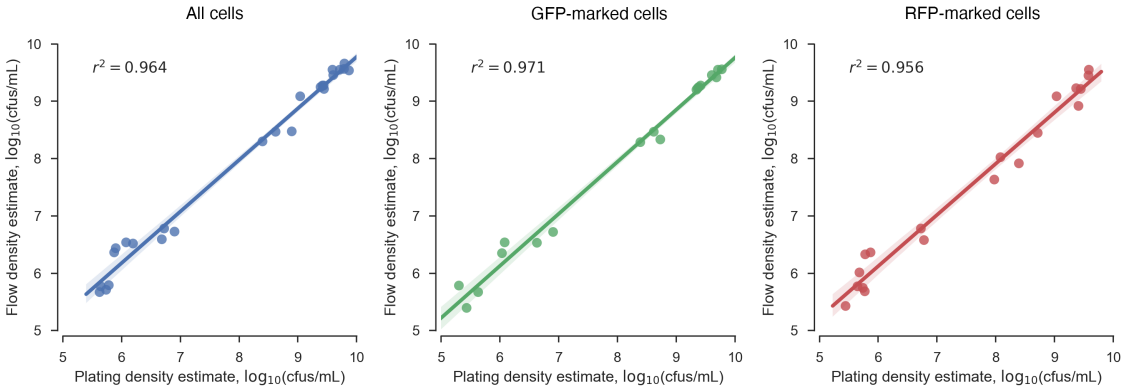


FIGURE 2.D.11. **Correspondence of cell enumeration using flow cytometry vs. agar plating.** Mixed cultures of GFP-marked cells and RFP-marked cells were grown from low density (~1×10^6 cfus/mL) to saturation (~2×10^9 cfus/mL). Samples of these cultures were taken throughout the growth process with densities spanning several orders of magnitude. Culture densities were estimated both with cell enumeration using a BD Accuri C6 Flow Cytometer and with traditional agar plating and colony counting (colonies of GFP-marked cells appear green, and colonies of RFP-marked cells appear pink). These plots show how density estimates from flow cytometry compare to those from plating. Each point depicts the respective density estimates for the same culture sample. The leftmost plot shows the correspondence of density estimates for all cells, the middle plot shows the correspondence of density estimates for GFP-marked cells alone, and the rightmost plot shows the correspondence of density estimates for RFP-marked cells alone. In all cases, the correlation of cytometry- and plating-based estimates of culture density is high.

Part 2

Stochastic dynamical network models of interventions for the COVID-19 pandemic.

CHAPTER 3

Model-driven mitigation measures for reopening schools during the COVID-19 pandemic.

As the COVID-19 pandemic accelerated in early 2020, schools around the world closed in an effort to preempt school-associated transmission and protect their students, teachers, and staff. By mid-April of that year, 195 countries had closed their schools in response to COVID-19, affecting more than 1.5 billion students (UNESCO 2020). In the United States (US), schools were among the first organizations to close, and many remained closed or transitioned to remote learning through the end of the 2019-20 school year. Some remain closed today. While remote learning affords students the opportunity to continue their education, it fails to provide many of the crucial benefits students typically receive through in-person schooling (Hanushek and Woessmann 2020). There is an urgent need to evaluate the effectiveness of evidence-based strategies that would allow children, teachers, and staff to safely return to in-person learning.

To date, widespread community transmission, conflicting public health guidance, and the emergence of new SARS-CoV-2 variants associated with higher transmissibility have compounded the challenges schools face when reopening (Leidman 2021, Nierenberg and Pasick 2020, Walensky et al. 2021). Numerous epidemiological models have been developed to forecast the spread of SARS-CoV-2 or compare the effectiveness of mitigation strategies in communities or large populations (Aleta et al. 2020, CDC 2021, Ferguson 2020, Giordano et al. 2020, Kucharski et al. 2020, Peak et al. 2020, Wang et al. 2020). However, only a few models have focused on the unique demographic and contact structures of primary and secondary school settings (Bershteyn et al. 2020, Bilinski et al. 2021, Bracis et al. 2021).

Case studies suggest that primary schools have a lower risk of transmission compared to secondary schools (Goldstein et al. 2020, Ismail et al. 2020a, Stein-Zamir et al. 2020). Two principal causes could be at play. First, younger children are less likely than adolescents or adults to become infected with SARS-CoV-2 (Viner et al. 2020), and less likely to experience symptomatic or severe disease (Assaker et al. 2020, Rajmil 2020). Second, primary and secondary schools have different contact structures. Primary school students have fewer contacts and typically spend the full day with a single teacher and the same group of students. By contrast, secondary school students move between classrooms and encounter multiple teachers and groups of students each day.

We have developed epidemiological models to simulate the spread of SARS-CoV-2 amongst students, teachers, and staff in both primary and secondary schools. Here, we use these models to better understand

the risks of reopening schools and to explore the effectiveness of different mitigation strategies: cohorting students, proactive testing, quarantine protocols, and vaccinating teachers and staff.

Model and methods

A stochastic network-based model of SARS-CoV-2 transmission. We use the SEIRS+ modeling framework (`https://github.com/ryansmcgee/seirsplus`) to study the dynamics of disease transmission in school populations. SEIRS+ builds upon classic SEIR compartment models that divide the population into susceptible (S), exposed (E), infectious (I), and recovered (R) individuals and track the transitions of individuals among these states (Keeling and Rohani 2011). The basic SEIR model is a deterministic model of a homogeneous population with well-mixed interactions. However, accounting for demographic heterogeneity and the structure of contact networks is particularly important when evaluating control strategies that perturb the contact network (e.g., social distancing) or make use of it (e.g., contact tracing)(Badham and Stocker 2010, Danon et al. 2012). For disease control, modeling stochasticity is crucial to understand the distribution of potential outcomes, especially in smaller populations.

To incorporate these important aspects of disease dynamics, we use the SEIRS+ modeling framework to implement an extended SEIR model of SARS-CoV-2 transmission on stochastic dynamical networks. Individuals are represented as nodes in a contact network. Parameters, interactions, interventions, and residence times in each compartment are specified on an individual-by-individual basis. This allows us to model realistic heterogeneities in disease, transmission, and behavioral parameters—which are particularly important when considering SARS-CoV-2 transmission dynamics in small, age-stratified school populations. The disease dynamics are summarized in Figure 3.1 and described in detail in Appendix 3.A. Parameter settings are outlined in Appendix 3.B.

We model infection as transmitted largely along a network of close contacts. Close contacts are individuals with whom one has repeated, sustained, or close-proximity interactions on a regular basis: classmates, friends, housemates, or other close relationships. Disease transmission can also occur among casual contacts—individuals who are not on one's contact network, but with whom one has incidental, brief, or superficial interactions. A network locality parameter sets the relative frequency and weight of transmission among close and casual contacts (Appendix 3.A.4). In both primary and secondary school settings, we assume that 80% of transmission occurs between close contacts specified by the networks (Mossong et al. 2008). Exposure to the community is modeled by randomly introducing new cases to the school population at a rate that corresponds to the community prevalence—see the Community prevalence and case introduction rate section below.

The likelihood that a susceptible individual becomes infected depends on the prevalence of infectious individuals among their contacts, the transmissibilities of these contacts, and their own susceptibility to infection (Appendix 3.A.4). Because individuals with large numbers of contacts—teachers, for example—are unlikely to interact as closely with each individual contact, we assume a logarithmic rather than linear scaling of transmission opportunity as a function of network degree (Appendix 3.A.4). Independent of connectivity, we assume an over-dispersed distribution of individual variation in biological transmissibility (Appendix 3.B.3), which corresponds to the observation that 80% of SARS-CoV-2 transmission may be attributable to 20% of infectious individuals (Adam et al. 2020, Althouse et al. 2020, Endo et al. 2020). This distribution of individual transmissibilities is calibrated to a nominal basic reproduction number R_0 for the population. While the

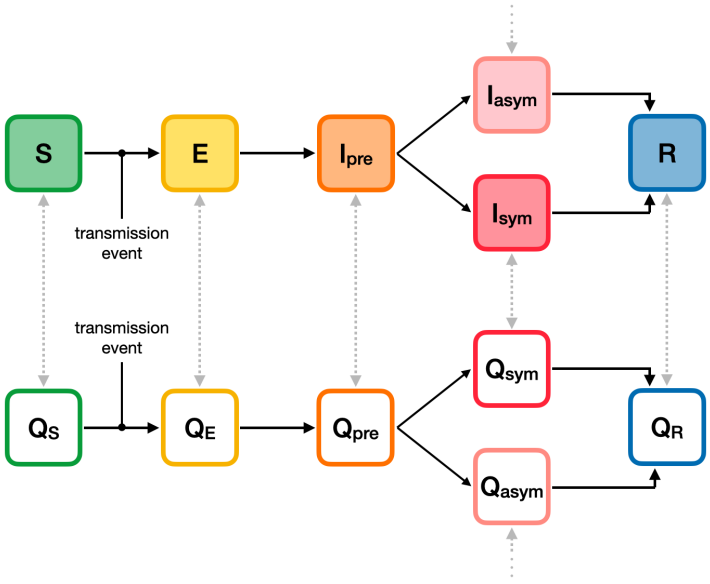


FIGURE 3.1. **Compartment model.** The progression of disease states in the Extended SEIR Network Model is represented by the compartments shown. Susceptible (S) individuals become infected (exposed) following transmissive contact with an infectious individual. Newly exposed (E) individuals undergo a latent period, during which time they are infected but not contagious. Infected individuals then progress to a pre-symptomatic infectious state (I_{pre}), in which they are contagious but not yet presenting symptoms. Some infectious individuals go on to develop symptoms (I_{sym}); while others will remain asymptomatic (I_{asym}). At the conclusion of the infectious period, infected individuals enter the recovered state (R) and are no longer contagious or susceptible to infection. The unshaded compartments represent quarantined individuals in the respective disease states. Individuals are moved into a quarantine compartment upon isolation due to symptoms or a positive test (gray arrows).

R_0 of SARS-CoV-2 varies, many estimates place R_0 upward of 2.5-3.0 without intervention, depending on the variants that are present in a given population (CDC 2020a, Galloway 2021, Li et al. 2020, Liu et al. 2020, Read et al. 2020). As a baseline, we assume that schools will implement sufficient mitigation measures, such as mask wearing, physical distancing, and increased ventilation, to reduce R_0 to 1.5 in the school population. More aggressive mitigation measures may bring the baseline R_0 in schools closer to 1.0. Results for other values of R_0 are discussed in Appendix 3.C.

Individuals in any disease state may enter quarantine due to symptoms or in response to a positive test result. The effect of isolating individuals is modeled by introducing compartments that represent quarantined individuals who do not make transmissive contact with others outside of the home (Figure 3.1, Appendix 3.B.4.3). Individuals remain in quarantine for 10 days (CDC 2020b), at which time they transition to the non-quarantine compartment corresponding to their present disease state. We assume that 20% of symptomatic individuals self-isolate upon the onset of symptoms (Poline et al. 2020, Appendix 3.C.4). At baseline, only the symptomatic or positive individual is isolated, but we go on to consider scenarios where classroom contacts of positive students are also isolated (Appendix 3.B.6.4).

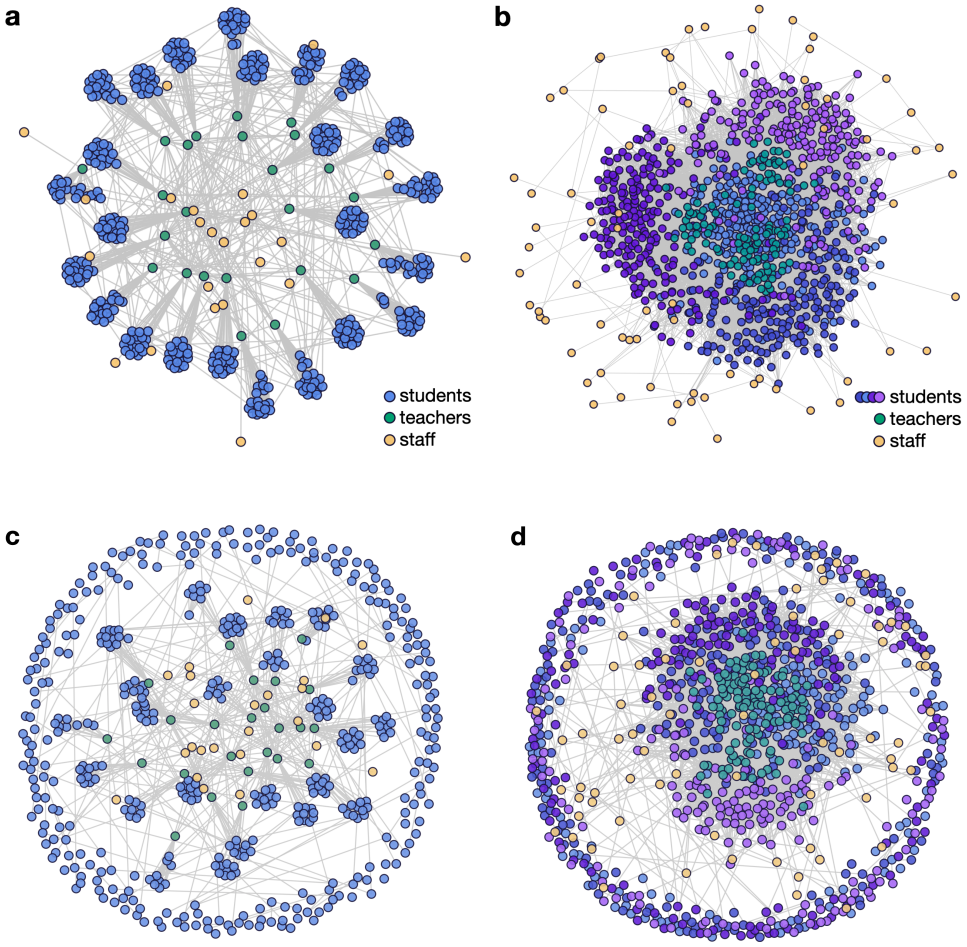


FIGURE 3.2. **Network structures for primary and secondary schools.** Each individual is represented by a circle, with grey lines connecting close contacts. (**a**) Primary school students (blue) are organized into classes with close contacts between all students in each classroom as well as a single teacher (green). School staff (yellow) interact with teachers and other staff. (**b**) Secondary school students (shades of blue and purple indicating grade levels) move between classrooms and have close contact with six teachers (green) each. School staff (yellow) interact with teachers and other staff. Secondary school students are clustered into loose social groups and are more likely to interact with other students in the same grade. (**c,d**) Example contact networks for primary and secondary schools, respectively, on a given day in a cohorting strategy in which students are divided into two groups that alternate in-person learning. Students that are in school on a given day (interior nodes) maintain the same school interactions as in the baseline networks. Students in the out-of-school cohort (peripheral nodes) make connections with any students that share their household (students in the same household are assigned to the same cohort in our model), but are disconnected from all other students and teachers. Students alternate between these interaction patterns according to a weekly or daily cohorting schedule.

Model considerations for primary schools versus secondary schools. We use distinct models for primary and secondary schools, with different contact networks reflecting the social structures in each setting (Appendix 3.B). We assume primary school children are 60% as susceptible as adults, while secondary school students are equally susceptible to adults (Goldstein et al. 2020, Viner et al. 2020). Our primary school model encompasses a school with 480 students, 24 teachers, and 24 additional staff. Primary school students have close contacts with their teacher, classmates, and other children in their household (e.g., siblings). For our secondary school model, we simulate a school with 800 students distributed across four grades, 125 teachers, and 75 additional staff. Secondary school students have close contacts with six teachers, with other students in their grade and social groups, and with other students in their households. Secondary school networks are parameterized such that connectivity statistics (e.g., mean degree, CV^2 of degree, clustering coefficient) are in line with empirical studies of secondary school contact networks (Barclay et al. 2014, Read et al. 2008, Salathe et al. 2010, Zhang et al. 2020, Appendix 3.B.4.2). Both settings feature a network of close contacts among teachers and staff. A new random network is generated for each simulation replicate. Example network diagrams for each school setting are shown in Figure 3.2. Detailed descriptions of the contact network structures and their generation are provided in Appendix 3.B.4.

Community prevalence and case introduction rate. To account for the effect of community prevalence on COVID-19 dynamics in schools, we model scenarios in which new cases are introduced into the school population stochastically at rates corresponding to daily, weekly, or monthly introductions on average (Appendix 3.B.5). When the effective community reproduction number R_{eff} is in the 1.0–2.0 range, these rates approximately correspond to the community prevalences shown in Table Appendix 3.B.5. We also consider the consequences of a single introduction; in this scenario, all replicates start off with the case introduction occurring on the first day of the simulation.

Simulations. To capture stochastic variability in outcomes, we report 1,000 replicates for each parameter set. Each replicate simulation tracks the progression of an outbreak that begins with the introduction of a single infected individual in an otherwise disease-free school population. The simulation begins on a random day of the week with the introduction of an initial case. Additional introductions may occur throughout the simulation at a Poisson rate reflecting the community prevalence. School is in session 5 days a week, and we assume that no close contacts are made outside of the household on weekends. Weekend transmission among casual contacts does occur. The simulation proceeds for 150 days to represent a school semester.

To allows comparisons across scenarios with different community prevalences, we report the percentage of cases attributable to transmissions within the school population (i.e., excluding introduced cases attributable to exogenous community exposure). These transmissions may occur either at school or among school-affiliated individuals while off campus, and are hereafter collectively described as "school transmission". We define "sizable outbreaks", as simulation runs where more than 5% of the population becomes infected in school over the course of the semester (150 days). While schools that experience sizable outbreaks are likely to stop in-person learning before very large case counts are realized, these data provide information about the probability of epidemic trajectories that could require such action.

Results

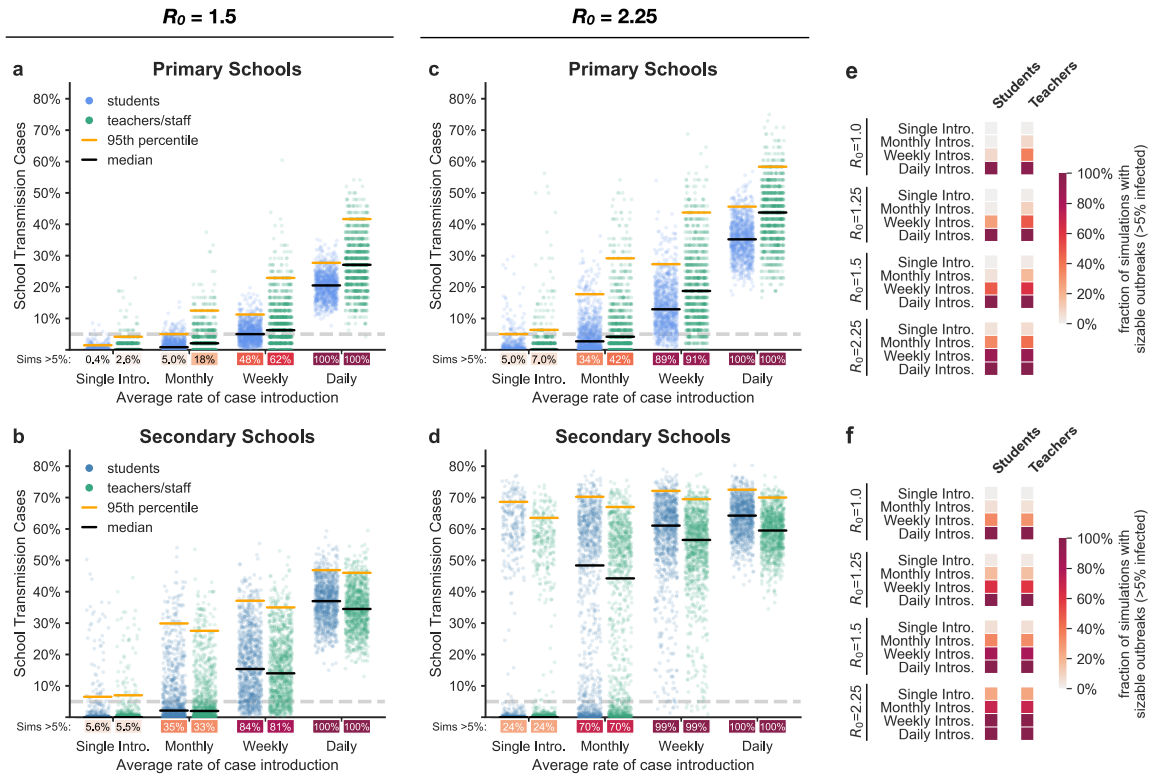


FIGURE 3.3. **Effect of community prevalence.** The distributions of school transmission cases as a percentage of the school population when new cases are introduced at different average rates. In these simulations, all students are in school five days a week and there is no proactive testing. (**a,b**) Outcomes for primary schools and secondary schools, respectively, with baseline transmission R_0=1.5. Black and orange lines represent median and 95th percentile outcomes, respectively. Under each jitter distribution we list the percentage of simulations where more than 5% (grey dashed line) of the population are infected in school. (**c,d**) Outcomes for primary and secondary schools in scenarios with heightened transmission R_0=2.25 due to the predominance of a highly-transmissible strain. (**e,f**) Heatmaps show the fraction of simulations where more than 5% of the student or teacher population are infected in primary and secondary schools, respectively, across a range of R_0 values and introduction rates.

The effect of community prevalence. The prevalence of COVID-19 in the community impacts the risk of transmission in schools. Figure 3.3 shows the percentage of the school population infected in primary and secondary schools over the course of a semester when only basic mitigation strategies (e.g. distancing, hygiene, and mask wearing) are in place. Higher COVID-19 prevalence in the surrounding community increases the probability of a sizable outbreak in primary and secondary schools alike. When community prevalence is so high that new introductions occur on a daily basis (between 0.25–1.0%), our simulations suggest that even aggressive mitigation strategies cannot prevent sizable outbreaks (Figure 3.5).

The probabilities of sizable outbreaks are higher in secondary schools than in primary schools, and the outbreaks tend to be larger in secondary schools. This difference holds across the range of parameters and interventions that we explore, and is primarily attributable to the difference in susceptibility between primary

and secondary school students (Appendix 3.C.3). Model outcomes are less sensitive to differences in the contact networks that characterize these settings (Appendix 3.C.3.2).

The effect of highly-transmissive variants. As of Spring 2021, several SARS-CoV-2 variants have evolved higher transmissibility relative to their ancestors (Galloway 2021, Hoffmann et al. 2021, Korber et al. 2020, Leung et al. 2021, Washington et al. 2021). For example, the B.1.1.7 lineage that emerged from the United Kingdom appears to be 30%-70% more transmissible than previous SARS-CoV-2 variants (Galloway 2021, Korber et al. 2020, Leung et al. 2021). To understand how highly transmissible variants may impact transmission dynamics where they become predominant, we look at the consequences of a 50% increase in transmissibility, which increases the assumed baseline R_0 for the school environment from R_0=1.5 to R_0=2.25. Results for more incremental increases in mean transmission rates that approximate intermediate penetrance of such strains can be found in Appendix 3.C.1.1.

Figure 3.3 illustrates how community prevalence, as modeled by introduction rate, influences school transmissions when schools are confronted by this more transmissible strain. Even under a monthly rate of new case introductions, schools face the risk of a major outbreak. With more frequent introductions, substantive outbreaks become the most likely outcome. Aggressive controls mitigate the risk somewhat, but are considerably less effective for a strain with R_0=2.25 than for a strain with R_0=1.5 (Figure 3.5).

The effects of interventions.

Cohorting. Cohorting, wherein students are divided into two or more groups that alternate in-person learning, is a common strategy for mitigating outbreaks in school settings (Kampe et al. 2020, Karin et al. 2020a, Zimmerman et al. 2021). In our model, we represent cohorting by shifting the contact networks according to which students are on campus (Appendix 3.B.4.5). While off campus, students are disconnected from the school network but maintain household connections and transmission to casual contacts (the latter representing out-of-school interactions among the student body). Teachers remain on campus across all cohorts.

?? shows the effects of three common cohorting strategies: (1) all students belong to a single cohort that is on campus full time, (2) students are divided into two cohorts, A and B, which are on campus on alternating days, and (3) students are divided into two cohorts which are on campus on alternating weeks (Appendix 3.B.6.3). We find that relative to no cohorting, both alternating day and alternating week strategies can improve outcomes substantially. Cohorting with alternating weeks generally outperforms cohorting with alternating days, but the marginal benefit of weekly cohorting is small when in-school transmission is limited (Appendix 3.C.1.1). In primary schools, student cohorting alone dramatically reduces the risk of outbreak amongst students. In secondary schools, cohorting is helpful but insufficient on its own to keep the likelihood of an outbreak low amongst students or amongst teachers and staff.

Proactive Testing. The purpose of proactive testing is to identify individuals who are infected but not currently showing symptoms, so that they can be quarantined (Denny et al. 2020, Larremore et al. 2020). We consider five proactive testing strategies, detailed in Appendix 3.B.6.2: (1) a baseline of no testing, (2) once-weekly testing amongst teachers and staff only, (3) twice-weekly testing amongst teachers and staff only, (4) once-weekly testing cadence amongst students, teachers, and staff, and (5) twice-weekly testing amongst students, teachers, and staff. We assume that 75% of students and 100% of teachers and staff are compliant with testing (Faherty et al. 2021). Previous work suggests that long test turnaround times severely curtail the value of

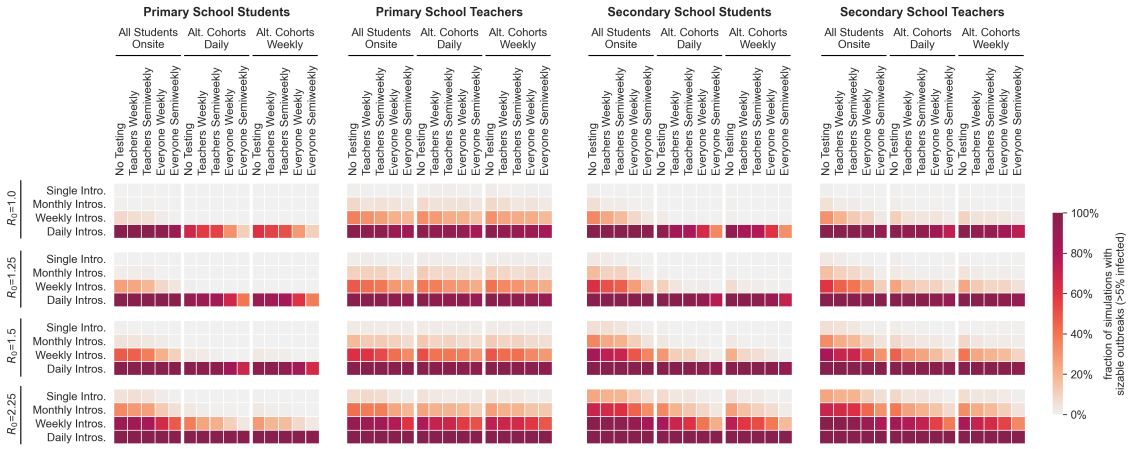


FIGURE 3.4. **Effects of cohorting and testing strategies.** Heatmaps illustrate the interactions of three student cohorting strategies and five proactive testing strategies (horizontal axis) across a range of transmission levels (R_0) and new case introduction rates (vertical axis). The color of each cell indicates the fraction of 1,000 simulations for the given parameter set that result in sizable outbreaks where more than 5% of the population is infected. Outcomes are shown for student and teacher populations in primary and secondary schools as indicated by the title above each heatmap.

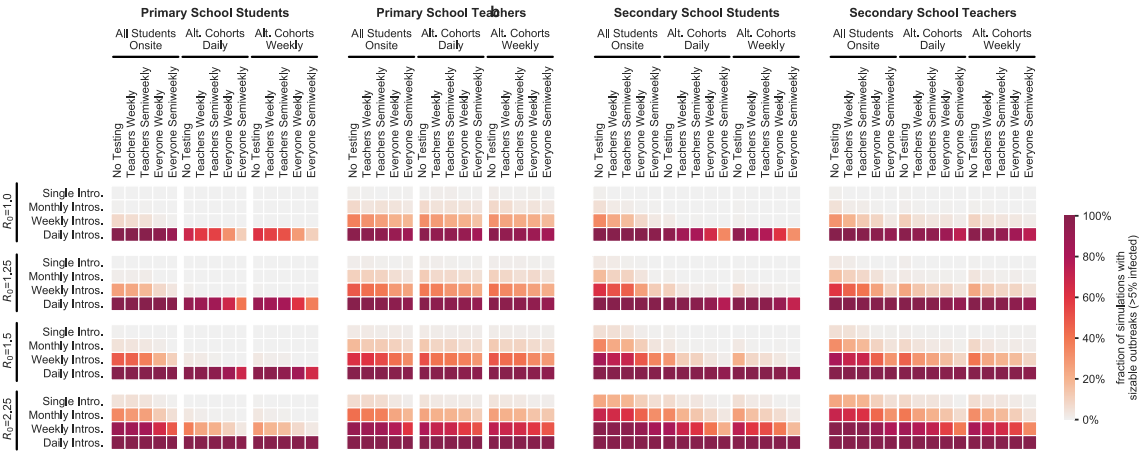


FIGURE 3.5. **Effects of cohorting and testing strategies.** Heatmaps illustrate the interactions of three student cohorting strategies and five proactive testing strategies (horizontal axis) across a range of transmission levels (R_0) and new case introduction rates (vertical axis). The color of each cell indicates the fraction of 1,000 simulations for the given parameter set that result in sizable outbreaks where more than 5% of the population is infected. Outcomes are shown for student and teacher populations in primary and secondary schools as indicated by the title above each heatmap.

testing (Bergstrom et al. 2020, Larremore et al. 2020). To account for this, we use a test turnaround time of 24 hours and assume positive individuals enter quarantine immediately. In our basic model, only those

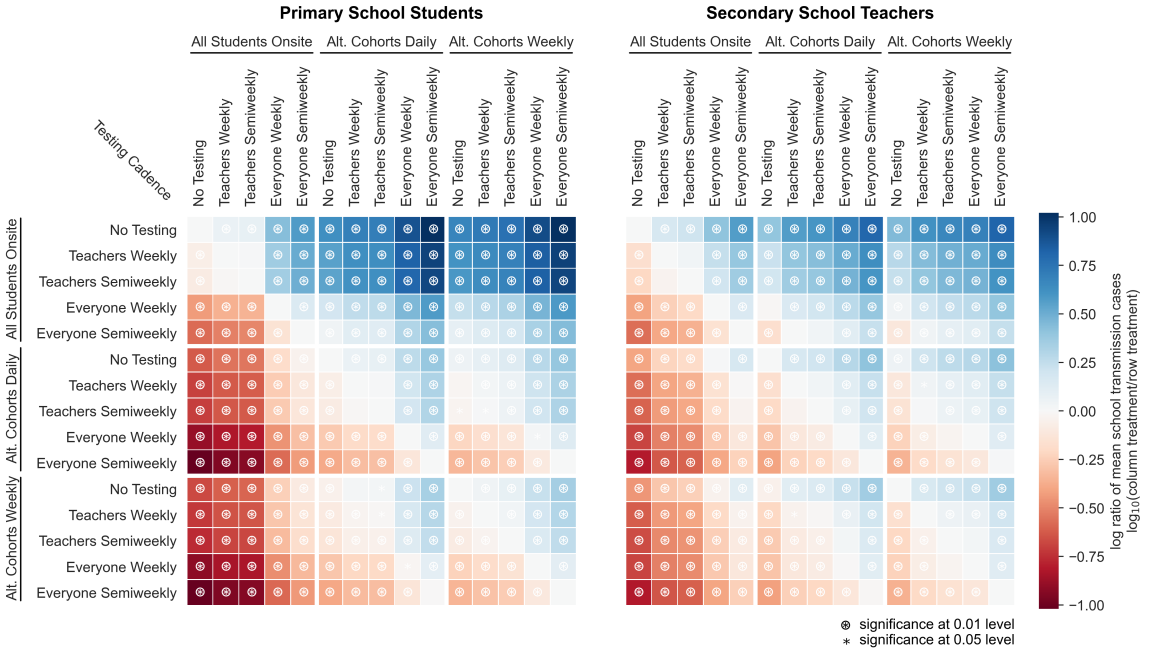


FIGURE 3.6. **Relative effects of testing and cohorting in a secondary school setting.** A heatmap of pairwise comparisons of testing and cohorting interventions illustrates the effects of various combinations on mean outbreak sizes. Each cell is colored according to the log-ratio of mean outbreak sizes for the two interventions, which represents the effect of the column intervention relative to the row intervention A **blue** cell indicates that the column intervention achieves a lower mean outbreak size than the row intervention, a **red** cell indicates that the column intervention has worse outcomes than the row intervention on average. Symbols in cells denote statistically significant differences in outbreak size distributions according to the Mann-Whitney U test at the 0.01 (⊛) and 0.05 (∗) levels. Results are shown for scenarios where R_0=1.5, case introductions occur weekly on average, and only the positive individual is quarantined when cases are detected.

who test positive are quarantined. We will later consider the consequences of quarantining all members of a primary school classroom when an individual therein tests positive.

Figure 3.5 illustrates the effects of cohorting and testing on the probability of sizable outbreaks. In our model, proactive testing consistently reduces the risk of outbreaks among teachers and students (Figure 3.5, Appendix 3.C). While cohorting alone does not completely mitigate the risk of sizable outbreaks in secondary schools, the combination of cohorting and testing can keep this risk in check when baseline transmissibility in the school is sufficiently low.

Figure 3.6 highlights the interactions between testing and cohorting measures in their effect on outbreak size in a secondary school environment. More aggressive testing helps reduce the size of outbreaks, as does cohorting. Testing and cohorting together outperform either measure alone. Interventions that help students also help teachers, and vice versa.

Isolation protocols. When an infected individual is identified by proactive testing, that person should be immediately isolated to prevent further transmission. In primary schools where classroom organization is stable, school administrators may additionally consider quarantining the entire classroom—students and teacher.

Our model indicates that classroom-level quarantine can reduce outbreak risk (Figure A4). For students and for teachers, with weekly introductions, the distribution of outcomes from isolating classrooms is stochastically smaller than the distribution of outcomes from isolating individuals (Mann-Whitney U test, $p \ll 0.01$). This means that if one takes the outcome of a randomly drawn simulation run with classroom-level isolation and another with individual-level isolation, the classroom-level simulation run is significantly more likely to have the better outcome. When introductions are less frequent, benefits of classroom isolation may not be statistically significant. One important consideration for quarantining at the classroom-level is that this approach imposes more quarantine days on the population. When the level of in-school transmission is relatively low (e.g., $R_0 = 1.5$), classroom-level isolation confers about the same amount of risk mitigation as moving from a weekly to semiweekly testing cadence, but classroom isolation leads to a large increase in the number of in-person learning days lost (Appendix 3.C.9).

Vaccination. Pfizer-BioNTech and Moderna have reported extremely successful results from their phase III and phase IV COVID-19 vaccine trials, with 90% or greater efficacy at blocking symptomatic disease (Baden et al. 2020, Dagan et al. 2021, Haas et al. 2021, Polack et al. 2020, Swift et al. 2021, Thompson 2021). The Johnson & Johnson vaccine has somewhat lower efficacy but requires only a single dose. Distribution of all three vaccines is well underway in the United States, with over a third of the US population fully vaccinated as of early May 2021.

Teachers and staff who have been vaccinated against COVID-19 are well-protected against infection (Figure 3.7). While initial phase III trial data focused only on diagnosis of symptomatic disease as a primary endpoint, current evidence suggests that the vaccines block transmission as well as symptomatic disease (Haas et al. 2021, Levine-Tiefenbrun et al. 2021, Petter et al. 2021, Thompson 2021). Vaccinating teachers can also reduce the risk of outbreaks among students, particularly when paired with cohorting. The combination of vaccinating teachers and cohorting students continues to substantially reduce the risk of outbreaks at higher levels of transmissibility, which suggests this strategy may offer a proactive defense against the spread of more transmissive variants.

Limitations

Like all epidemiological models, ours is a simplification of a complex, highly variable world. Our model is built on a series of assumptions and parameters; to the degree that these do not accurately reflect the real world, the model will be ineffective at predicting even the range of possible outcomes. We have attempted to account for uncertainty by embracing realistic heterogeneity and stochasticity in our model and by evaluating the sensitivity of outcomes across plausible ranges of values for critical parameters (Appendix 3.C). Still, in a novel pandemic where many epidemiological parameters remain uncertain, and social and behavioral factors are fluid, some mismatch is inevitable.

The basic reproduction number (R_0)—the average number of new cases generated by an infectious individual in a fully susceptible population—is a critical driver of disease dynamics. Estimates of R_0 at the community level reflect an average rate of transmission integrated over many contexts and behaviors, which may

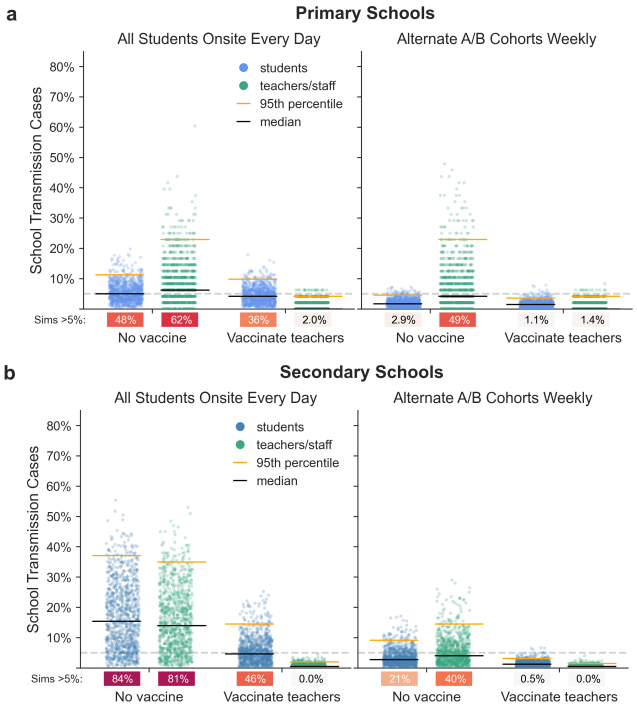


FIGURE 3.7. **Effects of vaccinating teachers.** The distributions of school transmission events as a percentage of school population for 1,000 simulations with either no vaccination or vaccination of all teachers and staff in (**a**) primary schools and (**b**) secondary schools. Results are shown for scenarios with R_0=1.5, approximately weekly new case introductions, and no testing. Because vaccination is only 90% effective in the model, some teachers and staff become infected even when all are vaccinated. Effective vaccinations block both disease and transmission.

include efforts to curtail transmission, such as social distancing, restricting large groups, and closing schools, businesses, and other gathering places. As schools reopen, larger numbers of individuals come together to interact, and average rates of transmission could be higher in schools than in the overall community. Still, basic in-school interventions such as mask wearing, physical distancing, and behavioral changes are expected to substantially reduce R_0.

In our model, we assume that these basic interventions can reduce R_0 to 1.5—roughly half of what it would be in the absence of intervention, depending on the transmissibility of variants circulating in the population. Previous studies suggest that transmission is relatively limited in schools (Heavey et al. 2020, Ismail et al. 2020b, Kampe et al. 2020, Macartney et al. 2020, Rajmil 2020, Zimmerman et al. 2021). In addition to the basic measures listed above, many of the schools described in these studies were already implementing one or more interventions along the lines of the ones we analyze here: cohorting, isolating groups, testing, contact tracing, reducing the number of people on campus, and so forth (Ismail et al. 2020a, Macartney et al. 2020, Panovska-Griffiths et al. 2020, Zimmerman et al. 2021). These studies largely corroborate our findings that school transmission is often kept in check when such mitigation strategies are used. Fewer studies have considered schools that are only using masks and other basic measures, but there is evidence that sizable school outbreaks can occur in these contexts (Stein-Zamir et al. 2020). We find that the probability and size

of outbreaks are strongly influenced by the underlying R_0, but the relative effects of mitigations are robust across a range of R_0 values (Appendix 3.C.1.1).

Here we have simulated a subset of the currently practiced strategies for returning to in-person learning (CDC 2020c). We assume that students, teachers and school staff adhere to testing cadences, cohorting schedules, and quarantine policies in addition to basic measures. In the absence of evidence to the contrary, we assume that—holding transmissibility and susceptibility constant—all forms of close contact are equally likely to result in transmission. In practice, the nature of interpersonal relations may make transmission from student to student or from teacher to teacher more likely than transmission between these groups, and could explain why some contact-tracing studies have reported disproportionately low student-to-teacher transmission (Gillespie DL 2021, Zimmerman et al. 2021).

We have modeled primary school children as being less susceptible to SARS-CoV-2 infection than teachers and staff. Recent evidence from seroprevalence and contact-tracing studies support this assumption (Bunyavanich et al. 2020, Carsetti et al. 2020, Goldstein et al. 2020, Leeb et al. 2020, Viner et al. 2020). However, because a high percentage of children develop asymptomatic disease, COVID-19 cases among children may be more likely to go undetected. Therefore, it is possible that the apparent decreased susceptibility to SARS-CoV-2 infection among primary school-age children is an artifact of underreporting. We find that school transmission is sensitive to the susceptibility of students, and if the susceptibility of primary school students is closer to that of adults then primary school outcomes will more closely resemble those of secondary schools (Appendix 3.C.3.1).

We also assume that there is no difference in infectiousness between pre-symptomatic, asymptomatic, and symptomatic individuals. Some studies suggest that pre-symptomatic individuals may contribute a disproportionately high number of cases relative to the duration of this disease state, and asymptomatic individuals may contribute disproportionately few cases relative to symptomatic individuals (Ferretti et al. 2020, Moghadas et al. 2020, Sayampanathan et al. 2021). However, these studies draw on contact tracing data, and it is unclear how much of these differences in transmission are attributable to differences in viral transmissibility (e.g., viral load and shedding) as opposed to changes in behavior or other factors associated with these disease states. Our assumption is conservative with respect to asymptomatic transmission in the sense that our results will err on the side of overestimating the number of school transmission events if asymptomatic or pre-symptomatic individuals are indeed less infectious than symptomatic individuals. However, our results may overstate the relative benefit of mitigations in such a case. If pre-symptomatic individuals are more infectious than symptomatic individuals, our results may underestimate the amount of transmission that is likely to occur in schools, but the importance of proactive mitigations would be even greater. The sensitivity of our model results to the relative infectiousness of pre-symptomatic and asymptomatic individuals is discussed in Appendix 3.C.1.3.

Over the course of the pandemic, community prevalence measurements have fluctuated substantially on a timescale of months, due to changing individual behaviors and societal interventions. Because these fluctuations have been largely unpredictable, we have elected to use a constant introduction rate throughout. In doing so we are effectively decoupling infection dynamics within the school from epidemic dynamics in the community. In our model, intervention choices that lead to a large number of school-related transmissions do not feed back on the community prevalence to influence the downstream hazard of community introduction back into the school. Similarly, in our model, mitigation choices that block school transmission do not reduce

the community introduction rate. This seems reasonable when schools are not important drivers of the community prevalence of SARS-CoV-2 infection, as appears to be the case especially for K-5 schools (Brandal et al. 2021, Ismail et al. 2020a, Zimmerman et al. 2021). Where schools are important drivers of community dynamics, however, our model risks underestimating the consequences of mitigation efforts. When schools drive community prevalence, planners must also consider the cost of the additional community infections that result from reopening schools—which we have not done here.

Summary

We have presented results from a simulation model of reopening schools during the COVID-19 pandemic. The purpose of this model is to provide a scenario-simulating tool that, when used in concert along with other credible sources of information and data, can aid decisions around school reopening policies.

We attempt to make reasonable assumptions about epidemiological parameters and aspects of human behavior that drive disease transmission. Our results tend to be robust to these choices, and the qualitative findings that we report—advantages to cohorting, testing, and vaccination—are expected to hold up more broadly. In Appendix 3.C, we provide detailed sensitivity analyses for a number of important parameters, including transmissibility, student susceptibilities, contact network structures, and compliance with intervention strategies. Our online webapp (`https://www.color.com/impact-of-primary-school-covid-19-testing`) provides a way to explore the range of parameters interactively, which can be used to assist in dynamic decision-making in response to uncertain and changing local circumstances.

Our model suggests that dividing students into cohorts that attend school in person on alternating schedules can be a powerful strategy for mitigating risk. Cohorting is effective in our model because students largely restrict in-person interactions to other individuals within their own groups, and this takes place only while at school. The cohorting strategy is fairly robust to students interacting off campus as well, provided that students continue to limit their contacts to students in their own cohort (Appendix 3.C.7). However, when students socialize beyond their close contacts and across cohort boundaries outside of school—as students are wont to do—the effectiveness of cohorting is reduced (Appendix 3.C.2). Schools could consider further efforts to reduce the mixing of the student body at school, which has a significant impact on the risk of transmission in all contexts (Appendix 3.C.2). This might include restructuring lunch periods, passing periods, transportation logistics, and other scenarios in which incidental transmission could occur between otherwise "unconnected" individuals.

Teachers and staff are more susceptible to the virus than primary school students and at higher risk of severe disease than students of any age. Moreover, teachers serve as conduits for outbreaks to move among classrooms within the school network. Frequent, proactive testing of teachers and staff can interrupt such transmission chains and further protect them from infection.

Vaccinating teachers and staff is a powerful tool for protecting this critical workforce. If vaccines effectively block SARS-CoV-2 transmission in addition to COVID-19 symptoms, vaccinating teachers and staff can significantly dampen outbreak dynamics in both primary and secondary schools. The result would be fewer cases among adults and students alike.

The success of reopening efforts will hinge on the amount of transmission that occurs in schools. The higher the transmissibility, parameterized here as R_0, the greater the chance of substantial outbreaks in a

school setting. Physical distancing, diligent use of masks, and other environmental controls offer a first-line approach to reducing transmission and will be an important component of reopening plans.

For both primary and secondary schools, the risk of an outbreak increases as cases in the surrounding community rise. One of the most effective ways to safely reopen schools is by controlling COVID-19 in the community. Surveillance should be in place to monitor levels of community transmission and schools should be prepared to respond flexibly.

Because highly transmissible variants such as B.1.1.7 pose increased risks for outbreaks, schools need to be vigilant on multiple fronts. First, where genomic surveillance is available, school districts and counties need to monitor the introduction and spread of these variants. Second, irrespective of the variants involved, it will be important to monitor epidemic dynamics within any given school and to respond quickly should uncontrolled spread take place. An additional virtue of testing is that it facilitates early detection of such events. We have not explicitly modeled surveillance testing and response, but general public health guidance should be followed. For example, schools could implement "tripwire" strategies, returning to distance learning for a period of time in response in-school outbreaks or rising community prevalence. In the event of isolated cases appearing at higher than expected rates, administrators should reconsider assumptions about the rate of community introduction and intensify control measures accordingly.

Our model suggests that under certain parameters, it may become difficult or impossible to keep the probability of outbreaks low across the schools of an entire district. Tripwire strategies may be necessary under these circumstances.

While gaps remain in our understanding of transmission in school settings, both real-world experience and models — including the one presented here — suggest a path forward for schools to reopen, particularly when community transmission is low and when it is possible to deploy and consistently implement the mitigation measures we have modeled here.

The authors thank Martin Rosvall for help in developing the contact network structures used in the SEIRS+ model. Ted Bergstrom, Natalie Dean, Bill Hanage, Michael Lachmann, and Marc Lipsitch provided valuable feedback in developing the model and adapting it to the school scenarios considered here.

Appendix 3.A. *SEIRS+ Extended SEIR Network Model*

SEIRS+ is an open source Python framework developed by McGee et al. that supports flexible parameterization and implementation of sophisticated epidemiological models (`https://github.com/ryansmcgee/seirsplus`). The models studied in this work are parameterizations of the stochastic Extended SEIR Network Model provided in the SEIRS+ framework. We simulate our models using the Interventions Simulation Loop provided in SEIRS+ with minor modifications for our particular school context. Extensive documentation for the models, simulation loops, and other features of SEIRS+ can be found on the SEIRS+ github wiki (`https://github.com/ryansmcgee/seirsplus/wiki`).

3.A.1. Heterogeneity. In the SEIRS+ Extended SEIR Network Model, individuals are represented as nodes in a contact network, and all parameters, interactions, and interventions can be specified on a node-by-node basis. Therefore, this model enables explicit representation of heterogeneity in disease characteristics, contact patterns, and behaviors, which are important for modeling small, age-stratified populations such as schools. Parameter choices and distributions for our school models are described in the School Models appendix section.

3.A.2. Compartments. The Extended SEIR Network Model extends the classic SEIR model of infectious disease to represent pre-symptomatic, asymptomatic, and symptomatic disease states, which are of particular relevance to the SARS-CoV-2 pandemic. The classic SEIR model divides the population into susceptible (S), exposed (E), infectious (I), and recovered (R) individuals. In this extended model, the infectious subpopulation is further subdivided into pre-symptomatic (I_{pre}), asymptomatic (I_{asym}), and symptomatic (I_{sym}) compartments, all of which represent contagious individuals (the full Extended SEIR Network Model includes a hospitalized infectious state, but we assume no hospitalization in this report and effectively ignore this compartment). Individuals transition from one compartment to the next at times determined by the disease characteristics (see Appendix 3.B.2 Disease progression parameters). A parameterizable fraction of infected individuals are deemed asymptomatic and will progress to the asymptomatic compartment when exiting the presymptomatic compartment, while the remainder of infected individuals will progress to the symptomatic compartment. The dynamics of compartment transitions are described further in the Dynamics appendix section.

The effect of isolating individuals in response to symptoms or testing is modeled by introducing compartments that represent quarantined individuals (Figure 4.1). An individual may be quarantined in any disease state, and every disease state has a corresponding quarantine compartment. Quarantined individuals follow the same progression through the disease states, but their set of close contacts are defined by a distinct quarantine contact network (Appendix 3.B.4 Contact Networks). In this work, individuals are moved into quarantine states by the Intervention Simulation Loop (Appendix 3.B.6.1), such as when a positive test result is returned, as opposed to according to a transition rate. Individuals remain in the quarantine compartment flow until the designated isolation period has been reached (10 days in this work), at which time they are moved into the non-quarantine compartment corresponding to their current disease state.

3.A.3. Dynamics. Transmission dynamics are simulated using the Gillespie algorithm, a common and rigorous method for simulating stochastic interaction dynamics. Briefly, the system's differential equations are adapted to compute the 'propensity' of the possible events (i.e., the expected amount of time until a given

event will take place) for all nodes at each time step. These propensities are then used to compute the probabilities of all possible state events normalized across the entire population. A random node and corresponding transition are selected to execute according to these probabilities in each time step. The propensities of transmission events (S to E transitions) are proportional to the product of the prevalence of infectious individuals among each node's contacts and the transmissibilities and susceptibilities of the interacting individuals (see Appendix 3.A.4 Transmission).

SEIRS+ supports calculating the propensities of disease progression transitions (e.g., E to I_{pre}, Q_{pre} to Q_{sym}) in two different ways: 1) using transition rates (standard Gillespie implementation), or 2) using compartment residence times. Running the model in the first mode results in exponentially distributed residence times in each compartment, as in classic mass action SEIR-like models. In the latter mode, each individual is assigned a residence time for each compartment, and the propensity for a given individual to transition out of the current state is 0 until this residence period has elapsed, at which time this propensity becomes large such that the next event will be this given individual transitioning to the next compartment with probability approaching 1. This results in a hybrid model where transmission events occur stochastically according to the Gillespie algorithm while other disease progression transitions occur in a clock-like manner in parallel. In reality, residence times in each disease state are not exponentially distributed, and the discrepancies can be particularly important when looking at early stages of an outbreak and when considering control strategies such as proactive testing.Feng et al. 2007, Krylova and Earn 2013 As such, we use the residence time propensity calculation mode in this work and assign heterogeneous residence times to individuals drawn from gamma distributions that better match empirical descriptions of the disease dynamics for COVID-19 (Appendix 3.B.2).

For more information about the propensity equations, refer to `https://github.com/ryansmcgee/seirsplus/wiki`.

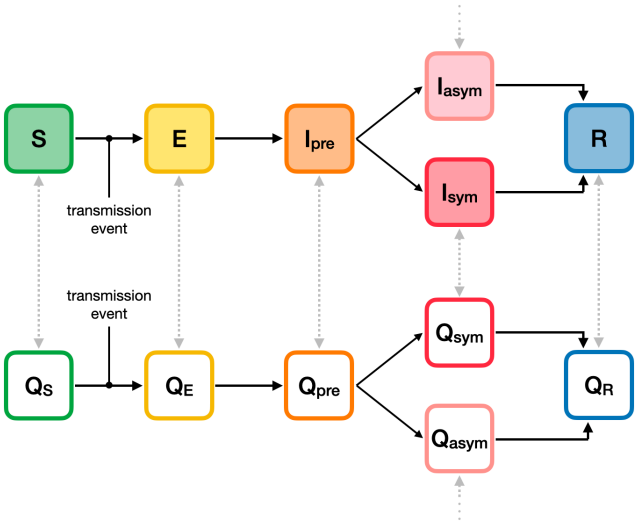


FIGURE 3.A.8. **Compartment model.** The compartment model that defines the progression of disease states in the Extended SEIR Network Model.

3.A.4. Transmission. The dynamics governing transmission events that cause susceptible individuals to become exposed are fundamental to any epidemiological model, so we break down the transmission dynamics of the Extended SEIR Network Model in detail here.

In general, the propensity $P^{(i)}(S \to E)$ of a given susceptible individual i becoming infected is proportional to the product of the prevalence of infectious individuals among their contacts, the average transmissibility of their infectious contacts $\overline{\beta}^{(\text{contacts})}$, and their own susceptibility to infection $\alpha^{(i)}$.

$$P^{(i)}(S \to E) \propto \alpha^{(i)} \times \overline{\beta}^{(\text{contacts})} \times (\text{prevalence among contacts})$$

An individual's transmissibility $\beta^{(i)}$ (i.e., transmission rate) is equal to the expected number of cases that this individual would generate in a fully-susceptible population (i.e., the reproduction number for the individual, $R_0^{(i)}$ divided by the length of their infectious period $\gamma^{(i)}$.

$$\beta^{(i)} = \frac{R_0^{(i)}}{\gamma^{(i)}}$$

For the purposes of the models considered in this work, the propensity of a given individual i becoming infected is calculated using the following equation[1], which we will break down in the rest of this section

$$P^{(i)}(S \to E) = \alpha^{(i)} \left[p \underbrace{\left(\frac{\bar{\beta}\left(I_{\text{pre}} + I_{\text{sym}} + I_{\text{asym}}\right)}{N} \right)}_{\text{global transmission}} + (1-p) \underbrace{\left(\frac{\sum_{j \in C_G^{(i)}} \delta^{(ji)} \left(\beta^{(j)} \mathbf{1}_{X^{(j)} \in \{I_{\text{pre}}, I_{\text{sym}}, I_{\text{asym}}\}} \right)}{|C_G^{(i)}|} \right)}_{\text{local transmission}} \right]$$

In this model, disease transmission may occur either from close contacts defined by the contact network structure or from casual contacts. Close contacts are individuals with whom one has repeated, sustained, or close proximity interactions on a regular basis: classmates, friends, housemates, or other close relationships. In contrast, casual contacts are individuals with whom one has incidental, brief, or superficial contact on an infrequent basis and to whom one is not connected directly on the network. A network locality parameter p sets the relative frequency and weight of transmission among close (local network) and casual (global) contacts in the model population.

3.A.4.1. *Global transmission.* A fraction p of a given individual's interactions are with casual contacts, which are assumed to be individuals randomly sampled from the population at large, irrespective of the contact network. With respect to these global interactions, every node in the population is equally likely to come into contact with every other node, and the population can be considered well-mixed. Thus the propensity of global transmission is calculated in the same way as mass action compartment models that assume a well-mixed population. The propensity for a given susceptible individual to become exposed due to global transmission is proportional to the product of that individual's susceptibility $\alpha^{(i)}$, the population mean transmissibility of infectious individuals $\overline{\beta}$, and the prevalence of infectious individuals in the overall population $(I_{\text{pre}} + I_{\text{sym}} + I_{\text{asym}})/N$.

[1] This equation is simplified from the general equation implemented in the Extended SEIR Network Model, which includes parameters and terms that are not used here and are thus zeroed out.

3.A.4.2. ***Local transmission.*** A fraction $1-p$ of a given individual's interactions are with individuals from their set of "close contacts." An individual's close contacts are defined as the nodes adjacent to the given node in the contact network. $C_G^{(i)}$ denotes the set of close contacts for individual i: the nodes adjacent to node i in the contact network graph G. $|C_G^{(i)}|$ denotes the size of this set: the number of close contacts that i has.

With respect to local transmission, transmissibility is considered on a pairwise basis. That is, every directed edge of the contact network representing transmission from infected node j to susceptible node i is assigned a transmissibility weight $\beta^{(j)}$ that depends on the transmissibility of the infected individual j alone. The propensity for a given susceptible individual to become exposed due to local transmission is calculated as the product of that individual's susceptibility and the sum transmissibility of their infectious close contacts ($\mathbf{1}_{X^{(j)}\in\{I_{\text{pre}},I_{\text{sym}},I_{\text{asym}}\}}$ is an indicator function that takes the value 1 when the state $X^{(j)}$ of the contact node j is one of the infectious states and 0 otherwise), scaled by the size of their local network as described below ($|C_G^{(i)}|$ denotes the size of the set of close contacts for individual i).

This amounts to the propensity of exposure for node i being proportional to the product of their susceptibility and the transmissibility-weighted prevalence of infectious individuals in their local network. Thus, propensity for exposure due to local transmission is frequency dependent and analogous to the propensity contribution from global transmission. Implicit in this formulation is an assumption that all individuals have limited interaction budgets and individuals with more close contacts (i.e., higher degree) interact less with each contact and are therefore less likely to become exposed by any single individual. A factor $\delta^{(ji)}$ appears in the calculation of propensity for exposure due to local transmission. This pairwise factor is used to re-weight the transmissibility of interactions according to the connectivity of the interacting individuals. While it is reasonable to think that individuals (e.g., secondary school teachers) who have many contacts (e.g., students) do not interact as closely with each of their contacts as another individual who only has a handful of contacts, we do not assume that the propensity of infection decreases linearly with degree for SARS-CoV-2 transmission. We define the degree scaling factor $\delta^{(ji)}$ as

$$\delta^{(ji)} = \frac{\log(D^{(i)}) + \log(D^{(j)})}{2\log(\overline{D})},$$

where $D^{(j)}$ and $D^{(i)}$ are the degrees of nodes j and i, respectively, and $\overline{D}$ is the mean degree of the network. Thus, the propensity for infection by a single infectious contact is lower for highly-connected individuals compared to low connectivity individuals, but not proportionally so.

Appendix 3.B. *School Models*

The following sections describe the specific assumptions and parameter values used to define the primary and secondary school models studied in this work. These models were implemented using the SEIRS+ framework's Extended SEIR Network Model (see Appendix Appendix 3.A SEIRS+ Extended SEIR Network Model).

3.B.1. Parameters Overview. Table 4.B.2 gives an overview of parameter values used in the primary and secondary school models. More information about these parameters can be found throughout Appendix 4.A. Exploration of the sensitivity of the models to these parameters can be found in Appendix 3.C.

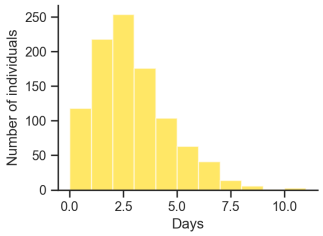
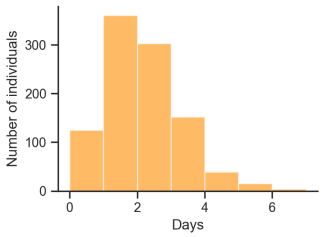
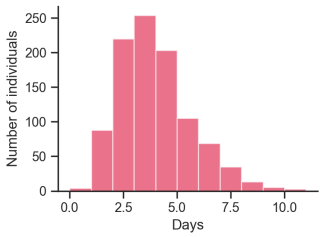
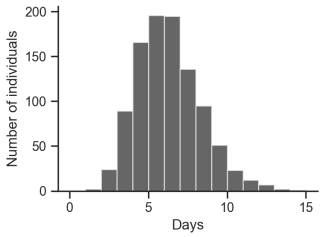
TABLE 3.B.1. Overview of parameter values for the primary and secondary school models.

Parameter	Primary School Model	Secondary School Model
Basic reproduction number R_0	1.5 or 2.25 (highly transmissible variant) *See Appendix 3.C.1.1*	1.5 or 2.25 (highly transmissible variant) *See Appendix 3.C.1.1*
Individual $R_0^{(i)}$ coefficient of variation	2.0 (overdispersed)Adam et al. 2020, Althouse et al. 2020, Endo et al. 2020 *See Table 3.B.3* *See Appendix 3.C.2*	2.0 (overdispersed)Adam et al. 2020, Althouse et al. 2020, Endo et al. 2020 *See Table 3.B.3* *See Appendix 3.C.2*
Presymptomatic transmissibility *(relative to symptomatic transmissibility)*	100% *See Appendix 3.C.1.3*	100% *See Appendix 3.C.1.3*
Asymptomatic transmissibility *(relative to symptomatic transmissibility)*	100% *See Appendix 3.C.1.3*	100% *See Appendix 3.C.1.3*
Susceptibility of students *(relative to susceptibility of adults)*	60% Goldstein et al. 2020, Viner et al. 2020 *See Appendix 3.C.3.1*	100% Goldstein et al. 2020, Viner et al. 2020 *See Appendix 3.C.3.1*
Fraction of asymptomatic cases	40% of students, Hurst et al. 2020, Poline et al. 2020, Rajmil 2020 30% of teachers/staff Buitrago-Garcia et al. 2020, Oran and Topol 2020	30% of students, Buitrago-Garcia et al. 2020, Oran and Topol 2020 30% of teachers/staff Buitrago-Garcia et al. 2020, Oran and Topol 2020
Disease state periods	*See Table 4.B.4*	*See Table 4.B.4*
Compliance with isolation upon symptoms	20% *See Appendix 3.C.4*	20% *See Appendix 3.C.4*
Compliance with testing	75% of students,Faherty et al. 2021 100% of teachers/staff	75% of students,Faherty et al. 2021 100% of teachers/staff
Test sensitivity	RNA-based time seriesLevine-Tiefenbrun et al. 2020 *See Appendix 3.B.6.2*	RNA-based time seriesLevine-Tiefenbrun et al. 2020 *See Appendix 3.B.6.2*
Vaccine uptake	0% of students, 0 or 100% of teachers/staff	0% of students, 0 or 100% of teachers/staff
Vaccine effectiveness	90% Baden et al. 2020, Dagan et al. 2021, Haas et al. 2021, Polack et al. 2020, Swift et al. 2021, Thompson 2021 *See Appendix 3.C.8*	90% Baden et al. 2020, Haas et al. 2021, Polack et al. 2020, Swift et al. 2021, Thompson 2021 *See Appendix 3.C.8*
Rate of disease with effective vaccine	0% Baden et al. 2020, Dagan et al. 2021, Haas et al. 2021, Polack et al. 2020, Swift et al. 2021, Thompson 2021	0% Baden et al. 2020, Haas et al. 2021, Polack et al. 2020, Swift et al. 2021, Thompson 2021
Rate of transmission with effective vaccine	0% Levine-Tiefenbrun et al. 2021, Petter et al. 2021	0% Levine-Tiefenbrun et al. 2021, Petter et al. 2021
Contact network parameters	*See Table 3.B.4*, *Table 4.B.3*	*See Table 3.B.6*, *Table 3.B.7*
Proportion of global transmission	20%Mossong et al. 2008 *See Appendix 3.C.2*	20%Mossong et al. 2008 *See Appendix 3.C.2*

3.B.2. Disease progression parameters. As described in Appendix 3.A.3 Dynamics, individuals remain in each compartment (excluding Susceptible) for a designated period of time before progressing to the next disease state. The population is heterogeneous for each disease state period, with each individual being assigned disease state periods drawn from gamma distributions that are informed by empirical studies of COVID-19 progression. Refer to Table 4.B.4 for more information about each distribution. We assume that the distributions of disease state periods are the same for all age groups and for both quarantined and non-quarantined individuals. The same gamma distribution parameters are used to define the period probability distributions in every simulation, but the period values are randomly drawn and assigned in each replicate.

Additionally, we assume that 30% of adults and secondary school students are asymptomatic, and that 40% of primary school students (young children) are asymptomatic. In the initialization of each simulation, each individual in the population is randomly assigned a symptomatic or asymptomatic status according to these probabilities. If an individual becomes infected, they will progress to the symptomatic (I_{sym}) or asymptomatic (I_{asym}) state when exiting the pre-symptomatic (I_{pre}) state according to this assigned status. 20% of symptomatic individuals self-isolate upon entering the symptomatic state, but there are no other parameter differences between symptomatic and asymptomatic individuals in our model.

TABLE 3.B.2. A representative distribution of period values drawn for a secondary school with 1,000 individuals is shown for each parameter in the center column below. Statistics across all replicate distributions in our analysis are shown in the rightmost column.

Disease state period	**Distribution**	**Statistics**
Latent period (time in E state)	 `gamma(mean=3.0, CV=0.6)`	mean 3.0 days std 1.8 days 95% CI (0.6, 7.4) References: Backer et al. 2020, Guan et al. 2020, He et al. 2020a, Lauer et al. 2020, Li et al. 2020, Tindale et al. 2020
Pre-symptomatic period (time in I_{pre} state)	 `gamma(mean=2.2, CV=0.5)`	mean 2.2 days std 1.1 days 95% CI (0.6, 4.8) References: He et al. 2020a, Tindale et al. 2020
Symptomatic period (time in I_{sym} or I_{asym} state)	 `gamma(mean=4.0, CV=0.4)`	mean 4.0 days std 1.6 days 95% CI (1.5, 7.6) References: Ganyani et al. 2020, He et al. 2020a, Wölfel et al. 2020, Young et al. 2020
Total infectious period (total time in I_{pre}, I_{sym}, and I_{asym} states)	 `gamma(mean=2.2, CV=0.5)`	mean 6.2 days std 1.9 days 95% CI (3.0, 10.5) References: Ganyani et al. 2020, He et al. 2020a, Tindale et al. 2020, Wölfel et al. 2020, Young et al. 2020

3.B.3. Transmission parameters. As described in Appendix 3.A.4 Transmission, the propensities of transmission events depend on the transmissibility and susceptibility parameters of interacting individuals. Each individual is assigned an individual reproduction number $R_0^{(i)}$, which is the expected number of secondary cases that the individual generates when infectious in a fully susceptible population. Each individual reproduction number is converted to an individual transmissibility (i.e., transmission rate) parameter using the following standard formula

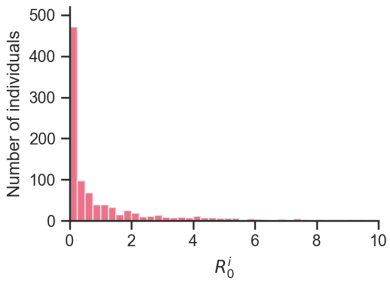
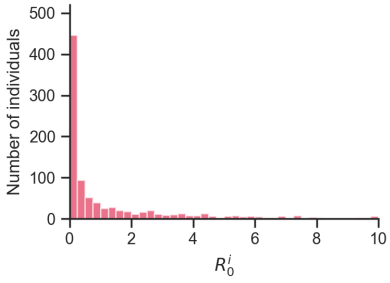
$$\beta^{(i)} = \frac{R_0^{(i)}}{\gamma^{(i)}},$$

where $\gamma^{(i)}$ is the total infectious period for individual i. We assume that individual transmissibility is heterogeneous and follows an overdispersed (long-tailed) distribution that corresponds approximately to 20% of individuals contributing 80% of the total expected number of secondary cases (the "80/20 rule").Adam et al. 2020, Althouse et al. 2020, Endo et al. 2020 We calibrate the individual reproduction number distribution such that its mean corresponds to a chosen average basic reproduction number R_0 for the population (R_0=1.5 and R_0=2.25 are considered in the main text) and so that 80% of the weight falls in the upper 20th percentile of individuals in the tail of the distribution. Therefore, for any R_0 considered in this paper, many individuals are expected to generate fewer than 1 secondary case while a minority of individuals are expected to contribute a large number. Refer to Table 3.B.3 for more information about these distributions. We assume that all age groups have transmissibilities drawn from the same distribution. In addition, we assume there is no difference in transmissibility between the pre-symptomatic, symptomatic, and asymptomatic states (i.e., the same individual transmissibility is used while an individual is in any one of these states).

Additionally, individuals are assigned a susceptibility parameter value, which weights the propensity that they become infected by any infectious contacts they may have (see Appendix 3.A.4 Transmission). Adults and secondary students are assigned the baseline susceptibility value of 1.0, and thus their propensity of infection is based on the unweighted transmissibilities of their contacts. In the main text, primary school students (young children) are assumed to be 60% as susceptible as adults. Therefore, primary school students are assigned a susceptibility value of 0.6, and their propensity of infection is only 60% of that of an adult in the same infectious contact context.

Global transmission can be thought to represent both casual interactions among members of the school population while on campus as well as relatively infrequent interactions among members of the school population while off campus (e.g., on weekends and off-cohort days). We assume that 80% of transmission is attributable to transmission between close contacts and 20% is attributable to global transmission among casual contacts (see Appendix 3.A.4 Transmission). This parameterization is supported by the extensive contact survey conducted by Mossong et al. (2008), which finds that approximately 20-30% of interactions where transmission is possible are with first-time contacts or last less than 15 minutes.Mossong et al. 2008

TABLE 3.B.3. A representative distribution of drawn individual reproduction number values for a secondary school with 1,000 individuals is shown for the basic reproduction numbers considered in the main text below. Statistics across all replicate distributions in our analysis are shown in the rightmost column.

Population R_0	**Distribution of individual reproduction numbers $R_0^{(i)}$**	**Statistics**
R_0 = 1.5	 `gamma(mean=1.5, CV=2.0)`	80% of weight in top 20th percentile.Adam et al. 2020, Althouse et al. 2020, Endo et al. 2020 mean 1.5 std 3.0 median 0.26 95% CI (0, 10.2) 80th percentile: 2.2 31% of values > 1.0
R_0 = 2.25	 `gamma(mean=2.25, CV=2.0)`	80% of weight in top 20th percentile.Adam et al. 2020, Althouse et al. 2020, Endo et al. 2020 mean 2.25 std 4.5 median 0.39 95% CI (0, 15.45) 80th percentile: 3.3 38% of values > 1.0

3.B.4. Contact Networks. The SEIRS+ Extended SEIR Network Model allows arbitrary graphs to be used to specify the contact network that defines close contacts for local transmission (see Appendix 3.A.4 Transmission). Here we define distinct networks representing the contact structure of a primary school and a secondary school.

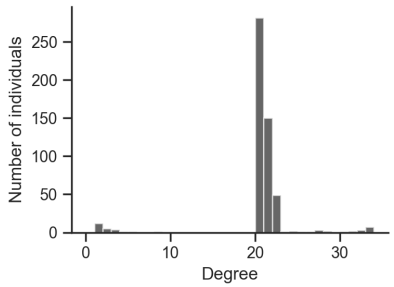
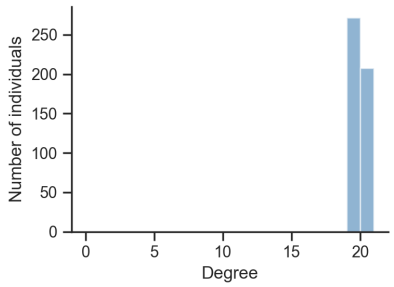
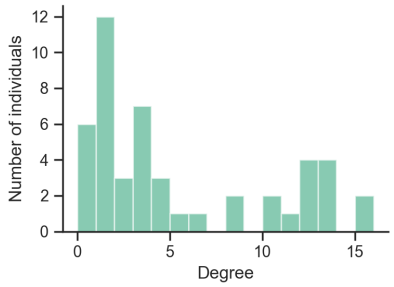
3.B.4.1. ***Primary school contact networks.*** For our primary school model, we simulate a medium-sized school of 480 students with 24 teachers and 24 additional staff. Each class comprises one teacher and 20 students in mutual contact. That is, the students and teacher for each classroom are strongly connected. Additionally, each teacher interacts with a handful of other teachers and staff, and students that share the same household are connected. The percentage of primary school aged children that share a household with another primary school aged child is calibrated by US census data. Most of the contacts that an individual makes in the school population are with the students and teacher in their own class, and disease transmission within a class is more likely than between classes. The FARZ algorithm generates random networks with built-in community structure and broad, heavy-tailed degree distributions that are realistic for human contact networks.Badham and Stocker 2010, Danon et al. 2012, Fagnan et al. 2018, Newman and Park 2003, Read et al. 2008 Refer to Table 3.B.4 for more information about the parameterization of these networks, and see Table 4.B.3 for more information about their degree distributions and other network properties.

3.B.4.2. ***Secondary school contact networks.*** For our secondary school model, we consider a medium-sized school with 800 students (200 per graduating class), 125 teachers, and 75 staff. We generate network layers for students and teachers and staff using the FARZ network generation algorithm, which allows us to calibrate epidemiologically-important network properties (e.g., cluster structure, assortativity, and clustering coefficient) to values consistent with studies of secondary school contact networks.Barclay et al. 2014, Salathe et al. 2010 A FARZ network layer is generated for each grade, with students belonging to one or more social groups (i.e., network clusters) of about 10 individuals each. 80% of each student's contacts are with students in the same grade, and 80% of those within-grade contacts are with students in their own social groups. Students

Table 3.B.4. Parameters for the generation of primary school contact networks.

Parameter	Value	Additional description
Number of grades	6 (K-5)	
Number of classes per grade	4	
Number of students per class	20	
Number of teachergroups	1	FARZ parameter for teacher/staff layers: Number of network clusters in teacher/staff layer
Teacher/staff mean degree	5	Average number of connections each teacher/staff makes with other teachers/staff
`alpha`	5	FARZ parameter for teacher/staff layers: Strength of common neighbor's effect on edge formation (tunes transitivity, clustering)
`gamma`	5	FARZ parameter for teacher/staff layers: Strength of degree similarity effect on edge formation (tunes assortativity)

TABLE 3.B.5. Degree distribution plots for a representative primary school network and network property statistics averaged across all primary school contact networks used in our analysis.

Network	Degree distribution	Network properties
Overall network		Degree mean: 20.1 Degree std: 4.5 Degree CV^2: = 0.5 Degree assortativity: 0.20 Clustering coeff.: 0.91 Average path length: 3.6
Student-Student layer		Degree mean: 19.5 Degree std 0.6 Degree CV^2 = 0.0 Degree assortativity: 0.05 Clustering coeff.: 0.95 Average path length: 3.9
Teacher-Staff layer		Degree mean 5.0 Degree std 4.8 Degree CV^2 = 0.8 Degree assortativity: 0.46 Clustering coeff.: 0.49 Average path length: 3.6

that share a household are connected as well. The percentage of secondary school aged children that share a household with another secondary school aged child is calibrated by US census data. Interactions between teachers and staff are represented by another FARZ network layer. Finally, students are connected with six random teachers with whom they have classes. Each teacher is associated with a grade level, and students take classes with teachers in their own grade level 75% of the time, which leads to students in the same grade being more likely to share teachers. A unique random network is generated as described for each simulation replicate. Refer to Table 3.B.6 for more information about the parameterization of these networks, and see Table 3.B.7 for more information about their degree distributions and other network properties.

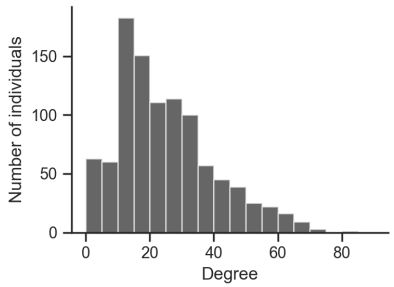
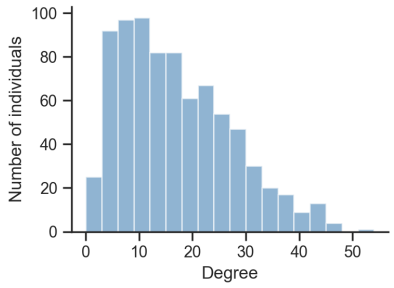
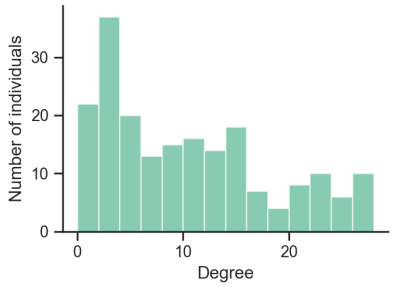
TABLE 3.B.6. Parameters for the generation of secondary school contact networks.

Parameter	Value	Additional description
Number of grades	4	
Number of students per grade	200	
Number of teachers	125	
Number of staff	75	
Number of classes per student	6	Number of classes that each student takes each day; i.e., number of teachers with whom each student is connected
Percentage of student-teacher grade level matches	75%	Probability that a student takes a class with, i.e., connects to, a teacher that is associated with their own grade level
Number of student social groups per grade	20	FARZ parameter for student layers: Number of network clusters in each student layer
Student mean intra-grade degree	16	Average number of connections each student makes with students in the same grade
Student percent inter-grade contacts	20%	Percent of each student's total student connections that are with students in another grade
Number of teacher/staff groups	10	FARZ parameter for teacher/staff layers: Number of network clusters
Teacher/staff mean degree	10	Average number of connections each teacher/staff makes with other teachers/staff
`alpha`	5	FARZ parameter for all layers: Strength of common neighbor's effect on edge formation (tunes transitivity, clustering)
`gamma`	5	FARZ parameter for all layers: Strength of degree similarity effect on edge formation (tunes assortativity)
`beta`	0.8	FARZ parameter for all layers: Probability of edges formation within clusters (strength of cluster structure)
`r`	2	FARZ parameter for all layers: Maximum number of clusters each node can belong to
`q`	0.5	FARZ parameter for all layers: Probability of a node belonging to the multiple clusters

3.B.4.3. ***Quarantine contact networks.*** When individuals are in quarantine, a separate quarantine contact network is referenced when calculating propensities of transmission involving that individual. Here we define the contact network as the school contact network with all edges removed except for those between housemates. That is, a quarantined individual makes contact with their housemates (e.g., siblings) but no one else from the school population. Global transmission is set to 0 for individuals in quarantine.

3.B.4.4. ***Weekend contact networks.*** The contact network that is in effect on weekends is the same as the quarantine network. That is, individuals only have direct contacts with housemates on weekends. However, global transmission is left at 20% for non-quarantined individuals on weekends to represent general mixing among the school population when out of school.

TABLE 3.B.7. Degree distribution plots for a representative secondary school network and network property statistics averaged across all primary school contact networks used in our analysis.

Network	Degree distribution	Network properties
Overall network		Degree mean: 24.1 Degree std: 15.0 Degree CV^2: = 0.39 Degree assortativity: -0.10 Clustering coeff.: 0.16 Average path length: 2.6 References: Barclay et al. 2014, Read et al. 2008, Salathe et al. 2010, Zhang et al. 2020
Student-Student layer		Degree mean: 16.0 Degree std 10.1 Degree CV^2 = 0.39 Degree assortativity: 0.16 Clustering coeff.: 0.22 Average path length: 2.9 References: Barclay et al. 2014, Read et al. 2008, Salathe et al. 2010, Zhang et al. 2020
Teacher-Staff layer		Degree mean 10.0 Degree std 8.2 Degree CV^2 = 0.64 Degree assortativity: 0.39 Clustering coeff.: 0.40 Average path length: 2.7

3.B.4.5. ***Cohort contact networks.*** One of the mitigations we consider is student cohorting, in which students are divided into two groups, only one of which attends school on any given day (see Appendix 3.B.6.3 Cohorting). In our model, cohorting is implemented by alternating between two modified school contact networks. Students are divided into two cohorts, A and B. Primary students are divided such that exactly half of each classroom is in each cohort. Secondary school students are arbitrarily divided (even and odd node indexes). A modified contact network is then generated to represent when cohort A is onsite, and one is generated to represent when cohort B is onsite. Each cohort network removes all edges from offsite students, except for their household connections, while maintaining the edges of onsite students. These networks are alternated according to the given cohorting schedule, as applicable.

. The degree-based pairwise transmissibility factors $\delta^{(ji)}$ (See Appendix 3.A.4.2 Local transmission for details) are calculated according to the connectivities of individuals in the baseline, "everyone onsite" network. The same set of factors derived from this baseline are used to calculate the propensities of local transmission at all times (i.e., for all school days, weekend days, and cohorting days), regardless of which cohort or weekend network is being used to define the structure of close contacts. This reflects an assumption that, for example, the interactions between individuals who are on campus don't become more intense under cohorting just because fewer students are on campus.

3.B.5. Case Introductions. Exposure to the community is modeled by randomly introducing new cases to the school population according to a Poisson process with an average introduction rate that corresponds to the community prevalence. Each day of the simulation, the number of introductions is drawn from a Poisson distribution using the given introduction rate as the Poisson parameter λ. Then for each exposure that is to be introduced (if greater than zero), an individual is drawn randomly from the population with selection probabilities proportional to the relative susceptibility of each individual. If the selected individual(s) are susceptible, they become exposed (infected)—otherwise they have been previously infected and their state is left unchanged. This process is handled within the simulation loop adapted from the SEIRS+ Intervention Simulation Loop (Appendix 3.B.6.1).

We consider monthly, weekly, and daily introduction rates, as well as single introduction scenarios. These rates roughly correspond to the community prevalences shown in Table 3.B.8. These associations between the community prevalence and the rate of introduction to the school population are approximated using the following method. The expected number of new cases to be generated in the overall community is approximated using the equation for the change in the number of infected individuals from the classic SIR model

$$dI_c = \frac{\beta_c S_c I_c}{N_c} = \beta_c S_c \pi_c \,,$$

where dI_c gives the expected number of new infections in the community per day, N_c is the size of the community, β_c is the average community transmission rate, S_c is the number of susceptible individuals in the community, and $\pi_c = I_c/N_c$ is the community prevalence (the subscript c denotes a community value). Then the number of these new cases that will land in the school population is assumed to be proportional to the ratio of the size of the school population to the overall community population.

$$\text{expected school introduction rate} = \frac{\beta_c S_c I_c}{N_c} \frac{N}{N_c} \,.$$

When the numbers of current and prior cases in the community (I_c and R_c, respectively) are small relative to the size of the community (i.e., $S_c \approx N_c$; this estimation will tend to overestimate the school introduction rate when there is significant susceptible depletion in the community), this can be simplified to a reasonable approximation that does not depend on the size of the overall community

$$\begin{aligned}
\text{expected school introduction rate} &= \frac{\beta_c S_c I_c}{N_c} \frac{N}{N_c} \\
&= \frac{\beta_c (N_c - I_c - R_c) I_c}{N_c} \frac{N}{N_c} \\
&\approx \frac{\beta_c N_c I_c}{N_c} \frac{N}{N_c} \\
&= \frac{\beta_c I_c}{N_c} N \\
&= \beta_c \pi_c N \,.
\end{aligned}$$

Thus the expected rate of introductions to the school population approximately equal to the product of the community transmission rate β_c, the community prevalence π_c, and the size of the school population N. The community transmission rate is equal to the effective reproduction number R_{eff} for the community divided by the average infectious period of the disease. Given estimates for these values, the introduction rate can be estimated. This method was used to estimate introduction rates for primary schools (N=528) and secondary

schools (N=1,000) for R_{eff} in the range (1.0, 2.0), a mean infectious period of 6.2 days, and a range of community prevalence values. The community prevalence ranges for each introduction rate listed in Table 3.B.8 are those prevalences for which the expected number of new cases per day in the school population is approximately equal to the listed introduction rate (monthly, weekly, or daily) for some R_{eff} in (1.0, 2.0) using this method and these parameters.

	Corresponding Community Prevalence	
Introduction rate (Poisson λ)	Primary school (528 individuals)	Secondary school (1000 individuals)
Monthly (λ = 1/30)	0.02 - 0.04%	0.01 - 0.02%
Weekly (λ = 1/7)	0.08 - 0.17%	0.04 - 0.09%
Daily (λ = 1)	0.6 - 1.2%	0.3 - 0.6%

For R_{eff} in the 1.0–2.0 range.

TABLE 3.B.8. **Introduction rates and community prevalences.** Given community transmission in the range of R_{eff} = 1.0–2.0, this table relates the prevalence of disease in the community to the frequency at which new cases are introduced into a school. Details of how these ranges are estimated are provided in Appendix 3.B.5.

3.B.6. Interventions. We model several interventions for mitigating the spread of SARS-CoV-2. The SEIRS+ framework provides code for a simulation loop that can implement several interventions, including testing, tracing, and isolation. We make use of a subset of the features in this simulation loop (with minor modification) to implement the mitigation strategies studied in this work.

3.B.6.1. ***Simulation loop.*** The simulation loop repeatedly calls a function that iterates the Gillespie dynamics of the Extended SEIR Network Model, which determines the next compartment transition (transmission event or disease progression) that will take place, advances the simulation time to the time of that event, and executes the state update. Every time the simulation time crosses an integer value (i.e., a new day is reached), the simulation loop interfaces with the model and its nodes, states, and parameters to implement various intervention procedures. If the Gillespie time to the next event is greater than a day, the simulation advances by a maximum time step that is a fraction of a day to ensure that intervention days are not skipped or irregularly timed. The simulation loop performs the following updates each iteration:

(1) Advance the Gillespie compartment transition dynamics
(2) If a new day has been reached, execute the following; else Return to (1):
    i Update active contact networks and parameters according to the weekend and cohorting schedule when applicable (see Appendix 3.B.6.3 Cohorting and Appendix 3.B.4.5 Cohort contact networks).
    ii Introduce new community exposure cases (see Appendix 3.B.5 Case Introductions).
    iii Isolate symptomatic individuals who are compliant with self-isolation upon symptom onset (see Appendix 3.B.6.4 Isolation).
    iv If the current day is part of the testing cadence, test individuals who are eligible for testing when applicable (see Appendix 3.B.6.2 Testing); else skip.

v Isolate individuals who have received a positive test result (following the test result lag time) and their classmates when applicable (see Appendix 3.B.6.4 Isolation)

vi Return to (1)

More details about these interventions are provided in the following sections.

3.B.6.2. ***Testing.*** We consider proactive testing that is executed according to one of several testing cadences (including no testing) shown in Table 3.B.9. These cadences define which groups of individuals are tested and on which days of the week. On a designated testing day, all individuals who are eligible to be in the testing pool are tested. An individual is considered part of the testing pool when they:

- Are a member of one of the groups designated in the testing cadence
- Are not currently in isolation
- Have not already had a positive test result
- Have not already recovered from the disease
- Have not been vaccinated
- Are compliant with testing

We assume that 100% of teachers and staff are compliant with testing, but 25% of students are non-compliant and thus never get tested. Students are assigned a compliance status randomly according to this probability when the model is initialized.

We model realistic temporal test sensitivities consistent with RNA-based tests. We assume 0% sensitivity for individuals in the exposed (latent) state. We assume 75% sensitivity for individuals in the first 2 days of their pre-symptomatic period and 80% sensitivity for any pre-symptomatic days beyond that. Sensitivities for symptomatic and asymptomatic individuals alike follow the time course shown in Figure 4.B.6, which follows data for RNA-based tests from Levine-Tiefenbrun et al.Levine-Tiefenbrun et al. 2020 We assume there are no false positives. We evaluate the sensitivity of model outcomes to the testing method in Appendix 3.C.6, where we consider test sensitivities in line with data for antigen-based tests.

There is a 1 day lag (exactly 24 hours) between administering a test and receiving the result. Individuals that receive a positive result enter isolation (i.e., move into a quarantine compartment) immediately upon receiving the positive result, thus 1 day after being tested. We assume that all individuals in the school population are compliant with entering isolation upon a positive test result.

Proactive Testing Cadence	Mon	Tue	Wed	Thu	Fri	Sat	Sun
No Testing							
Teachers Weekly	● ●						
Teachers Semiweekly	● ●			● ●			
Everyone Weekly	● ● ●						
Everyone Semiweekly	● ● ●			● ● ●			

● Teachers ● Staff ● Students

TABLE 3.B.9. **School testing cadences.** We explore the consequences of five testing cadences, from no testing to testing all members of the school community twice a week. Once-weekly testing takes place every Monday, and Semiweekly testing takes place on Mondays and Thursdays.

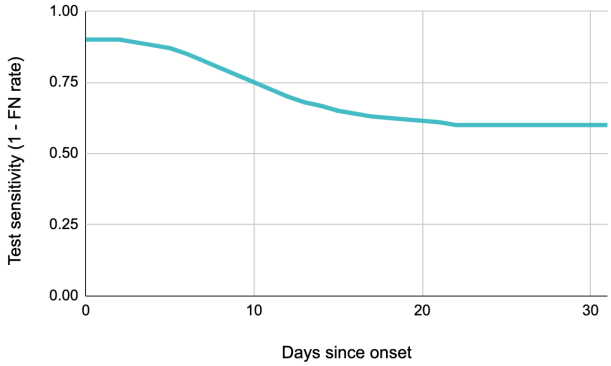


FIGURE 3.B.9. **Test sensitivities.** The probability of returning a positive test results when testing an individual in a symptomatic (I_{sym} or asymptomatic I_{asym} infectious state as a function of the number of days since entering that state (i.e., days since onset of symptoms). The test sensitivity is equivalent to 1 minus the false negative rate.

3.B.6.3. ***Cohorting.*** Cohorting consists of dividing students into two groups and following an alternating schedule in which only one group is on campus at a time. We consider two cohorting schedules, one where the group of students that is onsite alternates daily, and one where the onsite group alternates weekly (in addition to no cohorting). The cohorting schedules are summarized in Figure 3.B.10.

Cohort Schedule		Mon	Tue	Wed	Thu	Fri	Sat	Sun
All Students Onsite	Week 1	AB	AB	AB	AB	AB		
	Week 2	AB	AB	AB	AB	AB		
Alternate A/B Cohorts Daily	Week 1	A	B	A	B	A		
	Week 2	B	A	B	A	B		
Alternate A/B Cohorts Weekly	Week 1	A	A	A	A	A		
	Week 2	B	B	B	B	B		

A Cohort A on campus B Cohort B on campus

FIGURE 3.B.10.

Cohorting is implemented by alternating between different versions of the contact network in which one group of students or the other has their connections with the school population removed, except for any connections to individuals in their own household. Global transmission remains active for students who are offsite due to the cohort schedule, which can be thought of as students having some propensity to interact with other members of the school population outside of school. We assume that all students comply with the cohorting schedule. Students who are offsite due to cohorting are not considered to be in isolation, and these students are still part of the testing pool when otherwise applicable.

3.B.6.4. ***Isolation.*** When an individual enters isolation, they are moved into the quarantine compartment that corresponds to their disease state at the time of isolation. This compartment transition is executed "manually" by the simulation loop, separate from the Gillespie or residence time-based transition dynamics (Appendix 3.A.3 Dynamics). While in isolation, individuals transition between quarantine compartments according

to the same disease state residence times that are used when not in isolation. The set of close contacts for isolated individuals is given by a distinct quarantine contact network, which includes connections to members of the quarantined individual's household but no other members of the school population. In addition, quarantined individuals make no casual contacts (i.e., no global transmission). Isolated individuals remain in the quarantine sequence of compartments until their total isolation period has elapsed, at which time they are moved into the non-quarantine compartment that corresponds to their current disease state. We use a 10 day isolation period for all individuals, following the current CDC recommendation.CDC 2020d

In this model, individuals may enter isolation upon the onset of symptoms (if compliant), after receiving a positive test result, or when another member of their classroom has tested positive (for primary schools, when applicable).

We assume that 20% of all individuals elect to self-isolate upon the onset of symptoms. The symptomatic isolation compliance status of individuals is assigned randomly according to this probability when the model is initialized. For compliant individuals, there is a 1 day lag between transitioning into the symptomatic compartment and entering isolation. Individuals who are asymptomatic and thus enter the asymptomatic compartment rather than the symptomatic compartment never self-isolate. Symptomatic self-isolation does not trigger isolation of the symptomatic individual's contacts. Note that the rate of asymptomatic disease is assumed to be different in primary school-aged children (40%) and adults/adolescents (30%) while the rate of symptomatic isolation compliance is constant, so the effective rate of symptomatic self-isolation is lower in young children.

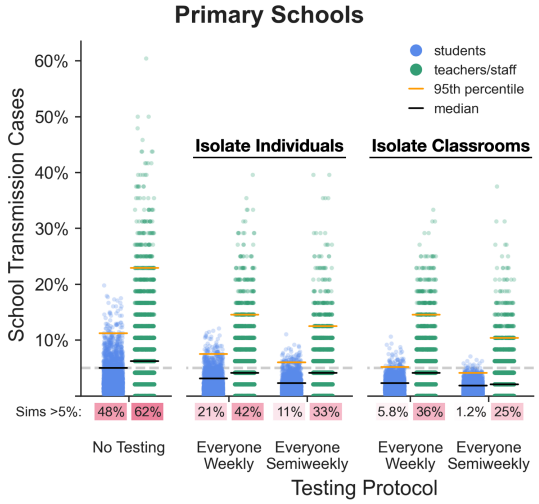


FIGURE 3.B.11. **Effect of isolating classrooms.** We consider two quarantine strategies for primary schools: (1) isolate single individuals who receive a positive test result, and (2) isolate the entire classroom (students and teacher) associated with an individual who receives a positive test. The distributions of school transmission cases as a percentage of the school population are shown for 1,000 simulations of each testing and isolation strategy with R_0=1.5 and approximately weekly new case introductions. Under each jitter distribution we list the percentage of simulations that result in outbreaks affecting more than 5% of the population. Black and orange lines represent median and 95th percentile outcomes respectively.

In primary schools, where classroom groupings are stable, we also consider scenarios where entire classrooms (i.e., all students and the teacher) isolate when any one member of the classroom tests positive. In such a case, the entire group enters isolation at the same time immediately after the index case receives their test result (i.e., 1 day after being tested).

3.B.6.5. ***Vaccination.*** We also evaluate the effectiveness of vaccinating teachers and staff on mitigating transmission in schools. We use the following working definitions with regard to vaccination in this model:

- Uptake: The percentage of individuals who receive a vaccine
- Effectiveness: The percentage of individuals receiving the vaccine that have an efficacious response. An efficacious response is characterized by an immune response that protects the vaccinated individual from falling ill and prevents that individual from transmitting disease.
- Reduction in transmissibility: The factor by which individual transmissibility is reduced for individuals with an efficacious response to vaccination

We model the scenario where 100% of teachers and staff are vaccinated with a vaccine that has 90% effectiveness (and no students are vaccinated). As noted above, an efficacious response also blocks 100% of transmission (i.e., the vaccinated individual's transmissibility is reduced to 0). The individuals that are to have an effective response are chosen randomly according to the probability of effectiveness when the model is initialized. Teachers are vaccinated and individuals with effective responses have their transmissibilities reduced before the simulation time begins. Individuals with effective immune responses are not included in case counts due to their immunity to the disease. Individuals with ineffective responses have no change in transmissibility or other parameters, can still contract and transmit the virus, and are included in case counts.

Appendix 3.C. *Sensitivity of Outcomes to Model Parameters*

School administrators and public health officials need ways of evaluating and mitigating the risk of disease transmission in school populations. Their challenge is exacerbated by ongoing uncertainty about disease characteristics, changing transmission and behavioral conditions, and variability from one population to the next. Even within a relatively circumscribed population such as a school, many variables influence the likelihood and severity of outbreaks. One benefit of a mechanistic modeling study is that the relative influence of parameters can be explored systematically.

In the main text we focus on two specific models intended to capture the dynamics of disease transmission within primary and secondary school populations, respectively. These models are characterized by the disease course, transmission, and contact network parameterizations described in detail in Appendix Appendix 3.B. Our choices of parameters in those models are data driven, where quality data are available, and are intended to capture the general features that are of epidemiological relevance for primary and secondary school populations; see Appendix Appendix 3.B for more information. In the real world, no two schools are alike, and few if any will exactly match our choice of parameters. In this section, we explore the effects of changes in parameter values on epidemic outcomes. This analysis provides information about the sensitivity of the model to various parameters, and offers insights about which parameters may merit special attention from pandemic response planners.

In each of the following sections, we present a sensitivity analysis across one or two referenced parameters at a time, while all other parameters are held at their baseline values as outlined in Table 4.B.2.

3.C.1. Transmissibility.

3.C.1.1. ***Basic reproduction number*** R_0**.** The basic reproduction number R_0 is a critical parameter of disease dynamics. In our model, each individual i is assigned an *individual reproduction number* $R_0^{(i)}$, which is the expected number of secondary cases that the individual generates when infectious in a fully susceptible population (Appendix 3.A.4). The *basic reproduction number* R_0 is the average number of secondary cases to be expected from an infectious individual in a wholly susceptible population. Any case has the potential to initiate a chain of transmission. There is stochasticity and individual variation associated with transmission events, but the greater the basic reproduction number the greater the number of secondary cases per infected individual on average—and the more likely it is that ongoing transmission occurs. In addition, the higher the value of R_0, the more likely that any single introduction seeds an outbreak, and the higher is the herd immunity threshold.

Figure 3.C.12 and Figure 3.C.13 show the average total number of school transmission cases and the mean cluster size per introduction, respectively, for primary and secondary schools across a range of R_0 values. As R_0 increases, the mean cluster size increases and the average number of school transmission cases increases as well. For a given R_0, primary schools have smaller clusters and fewer total school transmission cases on average, which is due in large part to the decreased susceptibility of primary school-aged children (Appendix 3.C.3.1). The relative benefits of mitigations are seen in the differences between curves in the same plot. All mitigations are beneficial even at low values of R_0, and the absolute benefits of all mitigations increase along with R_0. While the marginal benefits of testing semiweekly rather than weekly testing and of alternating cohorts on a

daily rather than weekly basis are relatively small when R_0 is low, these advantages become sizeable at higher levels of R_0.

The value of R_0 in a given population depends on many factors that influence the likelihood of transmission, including but not limited to individual behavior, interaction patterns, inherent transmissibilities of strains in the population, and features of the school environment. Of the parameters considered in our model, changes in R_0 have the greatest marginal effects on outcomes. This can make quantitative prediction difficult, because R_0 is seldom known precisely and is subject to fluctuations over time. But the qualitative trends are consistent: everything gets worse with higher R_0 values, and planners should make every effort to keep R_0 to a minimum using control measures such as physical distancing, mask wearing, and proper ventilation. Proactive mitigation strategies such as testing, cohorting, and vaccinating teachers can reduce the effective reproduction number further, but these strategies cannot sufficiently mitigate risk if the baseline level of transmission is too high.

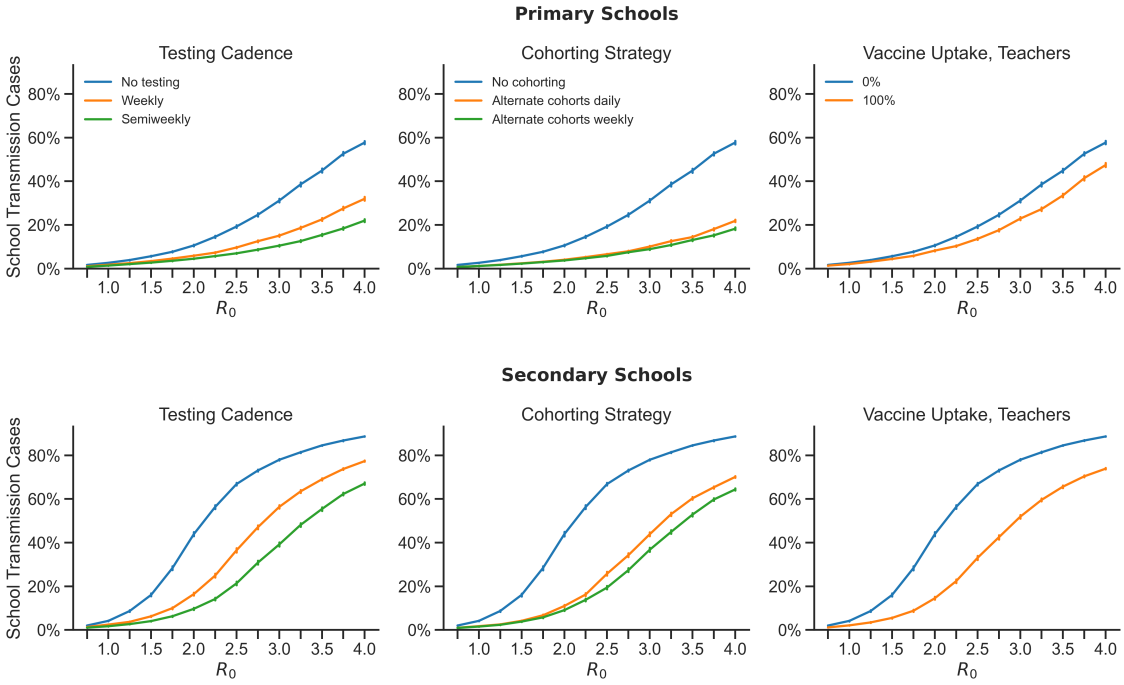


FIGURE 3.C.12. **Cases increase dramatically with increasing basic transmission number R_0.** The average number of school transmission cases are shown as a percentage of the total school population for R_0 values from 0.75 to 4.0. In this figure and in those that follow throughout this appendix, cases are introduced into the school once a week on average, and baseline values are used for all other parameters as outlined in Table 4.B.2. Error bars give the 95% confidence interval around the mean for 1000 replicates. Outcomes are shown for primary schools (top) and secondary schools (bottom). As a baseline we consider scenarios with no mitigation (blue lines), and to this we compare weekly or semiweekly testing of students and teachers (left column), alternating student cohorts on a weekly or daily schedule (center column), and vaccination of teachers (right column).

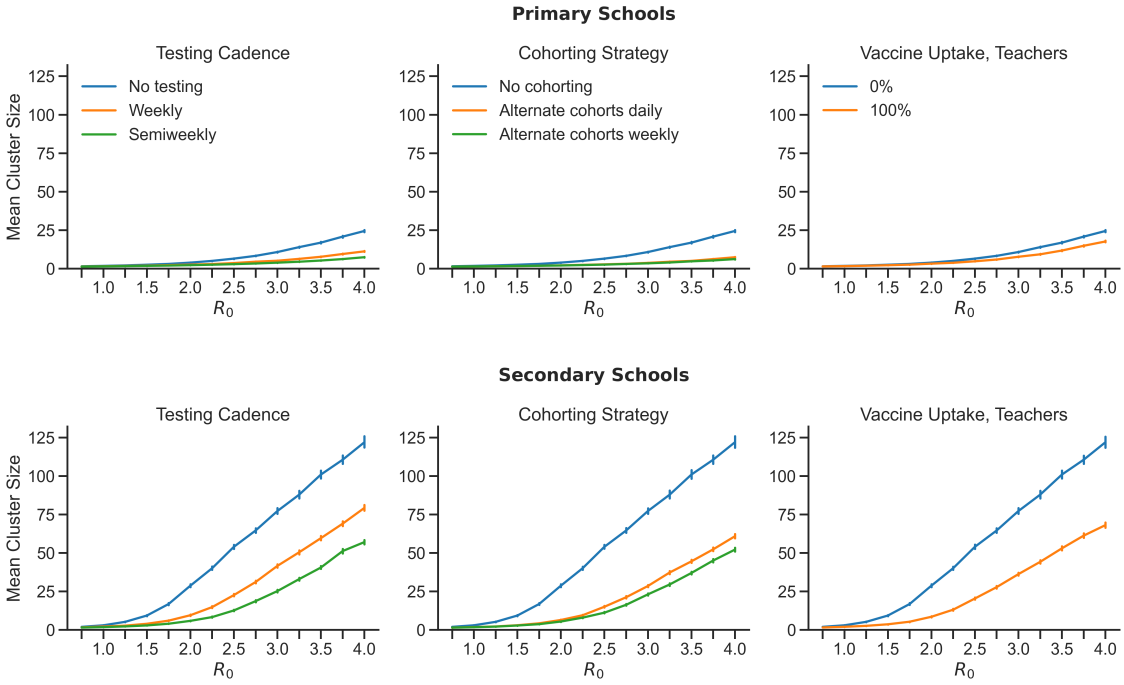


FIGURE 3.C.13. **Average cluster size increases with increasing basic reproductive number R_0.** The average cluster sizes per introduction in primary schools (top) and secondary schools (bottom) are shown for a range of R_0 values.

3.C.1.2. ***Overdispersion of individual transmissibility.*** In our model, individuals vary in their propensity to transmit disease. Each individual is assigned an individual reproduction number $R_0^{(i)}$ (i.e., the expected number of secondary cases that the ith individual would generate when infectious in a fully susceptible population). Individual reproduction numbers are drawn from a gamma distribution with a mean equal to the baseline R_0 and a shape parameterized by a coefficient of variation. In main body of this paper we assume that individual transmissibilities follow an overdispersed (long-tailed) distribution, with approximately 20% of individuals contributing 80% of the total expected number of secondary cases (the "80/20 rule").Adam et al. 2020, Althouse et al. 2020, Endo et al. 2020 Under such a distribution, the majority of individuals are expected to generate fewer than one secondary case while a minority of individuals are expected to contribute many cases. Evidence from contact tracing studies suggest that transmission of SARS-CoV-2 follows this 80/20 rule,CDC 2020a, Li et al. 2020, Liu et al. 2020, Read et al. 2020 but we might also consider cases with reduced variation in individual transmissibility, such that most individuals are expected to generate a number of cases close to the value of R_0. A gamma distribution with a coefficient of variation of 2.0 gives an overdispersed distribution where 80% of individuals fall in the upper 20th percentile of the distribution. A gamma distribution with a coefficient of variation of 0.2 gives a bell-shaped, mean-centered distribution with less variation in transmissibility.

Figure 3.C.14 shows the effect of the amount of variation in individual transmissibility on the number of transmissions in school populations. Populations with overdispersed distributions of individual transmissibilities exhibit fewer school transmission cases in principle,Althouse et al. 2020, HÃ©bert-Dufresne et al.

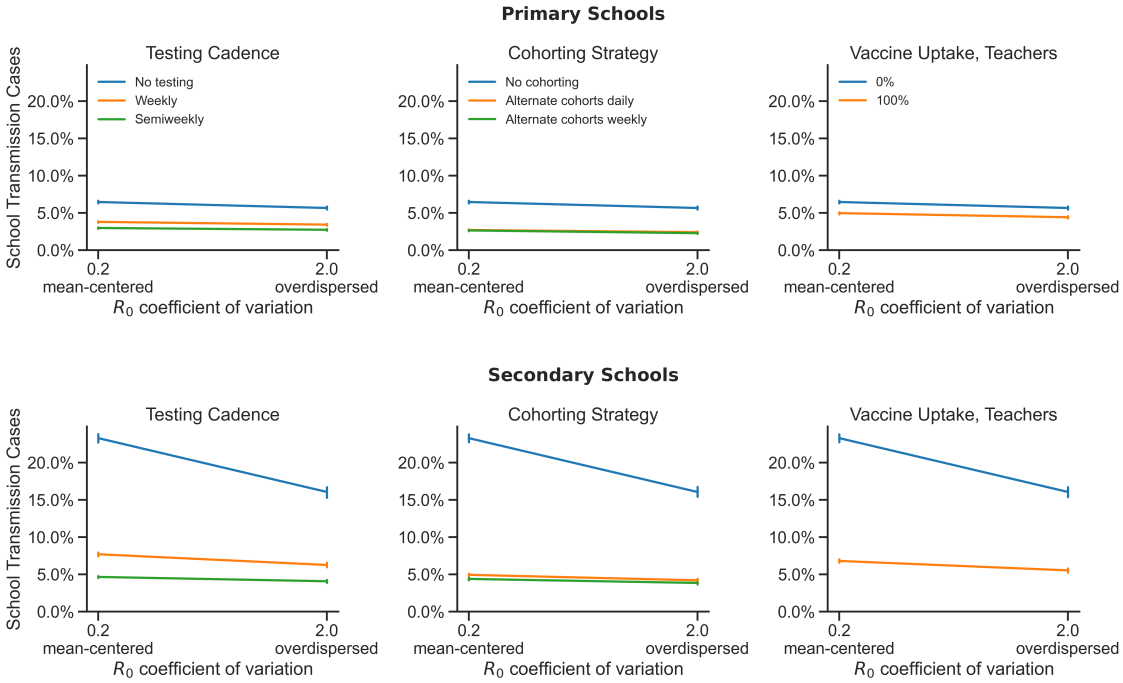


FIGURE 3.C.14. **Variation in individual transmissibility has at most a modest effect on school transmission.** The average number of school transmission cases are shown as a percentage of the total school population for mean-centered (gamma coeff. of variation = 0.2) and overdispersed (gamma coeff. of variation = 2.0) distributions of individual reproduction numbers. In primary schools under all conditions and in secondary schools with any testing, cohorting, or vaccination in place, variation in individual transmissibility has no effect on trasmission in schools. In secondary schools with no mitigations in place, overdispersed transmissibilities reduce the mean transmission somewhat.

2020, Lloyd-Smith et al. 2005 as is especially apparent in secondary schools without mitigation (blue lines in bottom row). In overdispersed populations, a small number of individuals may generate a large number of cases should they become infected, but a majority of the population have individual reproduction numbers $R_0^{(i)} < 1$ and are expected to contribute less than a single secondary case on average. Some outbreaks that go through highly-transmissible individuals are large, but the average transmission chain is short lived. On the other hand, most individuals in a population with a mean-centered distribution of individual transmissibility are expected to generate at least one secondary case if the mean reproduction number R_0 is greater than 1. Super spreading is less common, but the average number of transmissions is higher. We also see in Figure 3.C.14 that once mitigations are introduced, observed outcomes are insensitive to the amount of variation in individual transmissibility. This is presumably because when transmission chains are broken early, individual variation in transmissibility does not come much into play.

3.C.1.3. ***Pre-symptomatic and asymptomatic infectiousness.*** Some studies suggest that pre-symptomatic individuals may contribute a disproportionately high number of cases relative to the duration of this disease

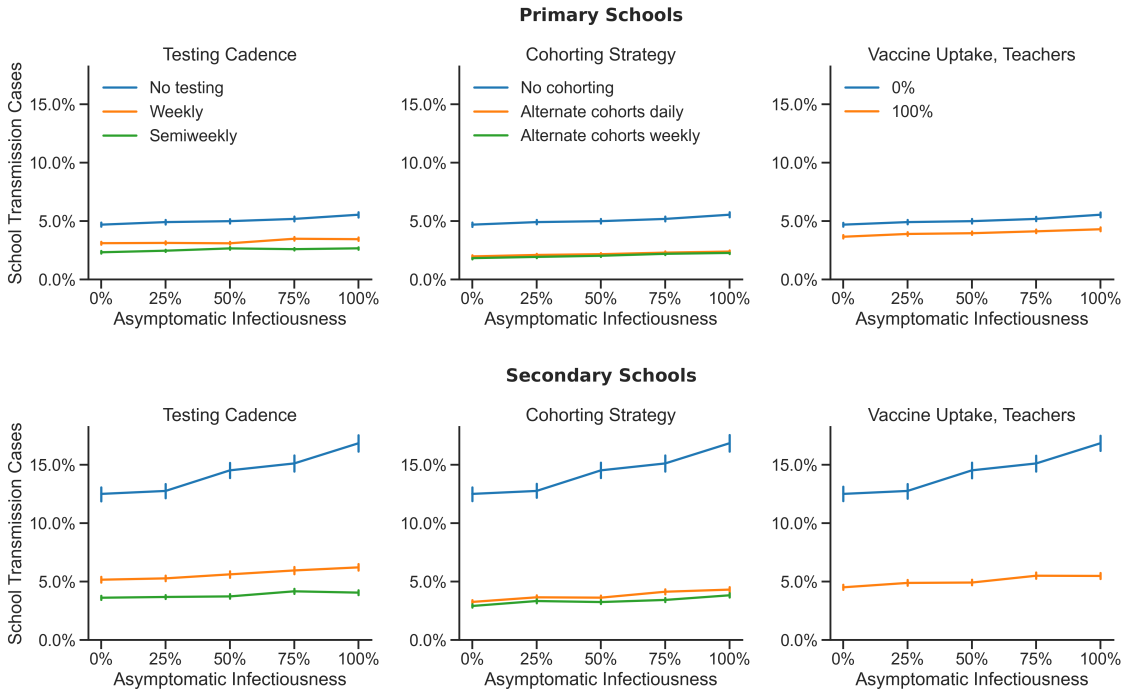


FIGURE 3.C.15. **Increasing asymptomatic transmissibility drives increased transmission in secondary schools when interventions are not used, but has minimal effects under other other conditions** The average number of school transmission cases are shown as a percentage of the total school population for a range of asymptomatic transmissibily values, relative to the transmissibility of symptomatic individuals.

state, and asymptomatic individuals may contribute disproportionately few cases relative to symptomatic individualsFerretti et al. 2020, Moghadas et al. 2020, Sayampanathan et al. 2021. However, these studies draw on contact tracing data (e.g., incidence rates and attack rates associated with individuals in different disease categories), and it is unclear how much of the differences in transmission between pre-symptomatic, symptomatic, and asymptomatic individuals are attributable to differences in biological transmissibility (e.g., viral load and shedding) as opposed to differences in behavior or other factors. Without more data regarding the relative viral loads and viral shedding of pre-symptomatic and asymptomatic individuals, we opt to assume that there is no difference in biological transmissibility between the pre-symptomatic, symptomatic, and asymptomatic states. This assumption is conservative with respect to asymptomatic transmission in the sense that our results will err on the side of overestimating the number of school transmission events if asymptomatic or pre-symptomatic individuals are indeed less infectious than symptomatic individuals. However, our results may overstate the relative benefit of mitigations in such a case. If pre-symptomatic individuals are more infectious than symptomatic individuals, our results may underestimate the amount of transmission that is likely to occur in schools, but the importance of proactive mitigations will be even greater.

Figure 3.C.15 shows the average number of school transmission cases for primary and secondary schools across a range of asymptomatic transmissibilities relative to symptomatic cases. Particularly in the absence of

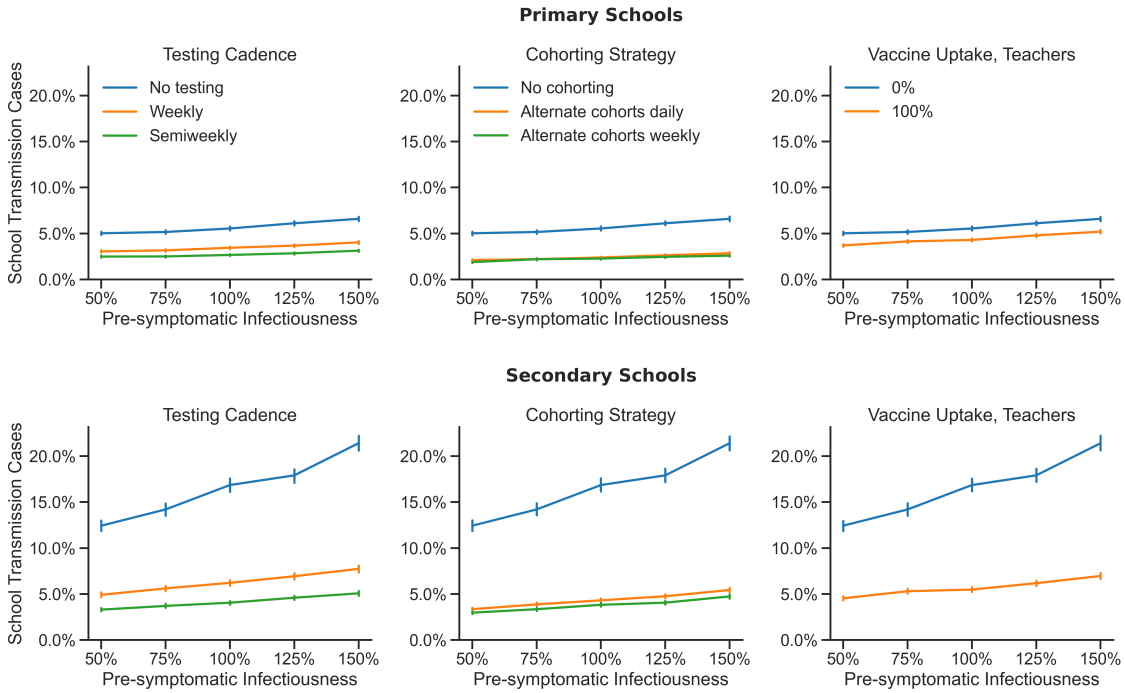


FIGURE 3.C.16. **Increasing presymptomatic transmissibility drives increased transmission in secondary schools when interventions are not used, but has modest effects under other other conditions.** The average number of school transmission cases are shown as a percentage of the total school population for a range of pre-symptomatic transmissibility values relative to the transmissibility of symptomatic individuals.

interventions, the average outbreak size is reduced somewhat when asymptomatic cases have reduced transmissibility. This effect is more pronounced in secondary schools than in primary schools. The marginal benefits of mitigations are greater at higher levels of asymptomatic infectiousness, but there is still significant benefit to all interventions in both settings even when asymptomatic individuals are not infectious at all.

Figure 3.C.16 shows the effect of varying pre-symptomatic infectiousness. As with asymptomatic infectiousness, the average number of school transmission cases and the benefits of mitigations are greater at higher levels of pre-symptomatic transmissibility. If pre-symptomatic individuals are indeed more infectious than symptomatic individuals (independent of changes in behavior, etc.), then proactive interventions such as testing, cohorting, and vaccination have even greater value.

Figure 3.C.17 illustrates the effect of varying pre-symptomatic and asymptomatic infectiousness together. In the absence of testing or other mitigations (top), model outcomes are approximately equally sensitive to changes in either parameter, although changes in asymptomatic infectiousness have a slightly greater impact. The effect of increasing pre-symptomatic infectiousness is approximately offset by a commensurate decrease in asymptomatic infectiousness. When semiweekly testing is implemented, many asymptomatic cases are detected and isolated, and outcomes become insensitive to the level of asymptomatic infectiousness. However, outbreak size is still sensitive to pre-symptomatic infectiousness even with frequent testing, because some individuals will not be tested during their presymptomatic period and because testing will fail to detect some

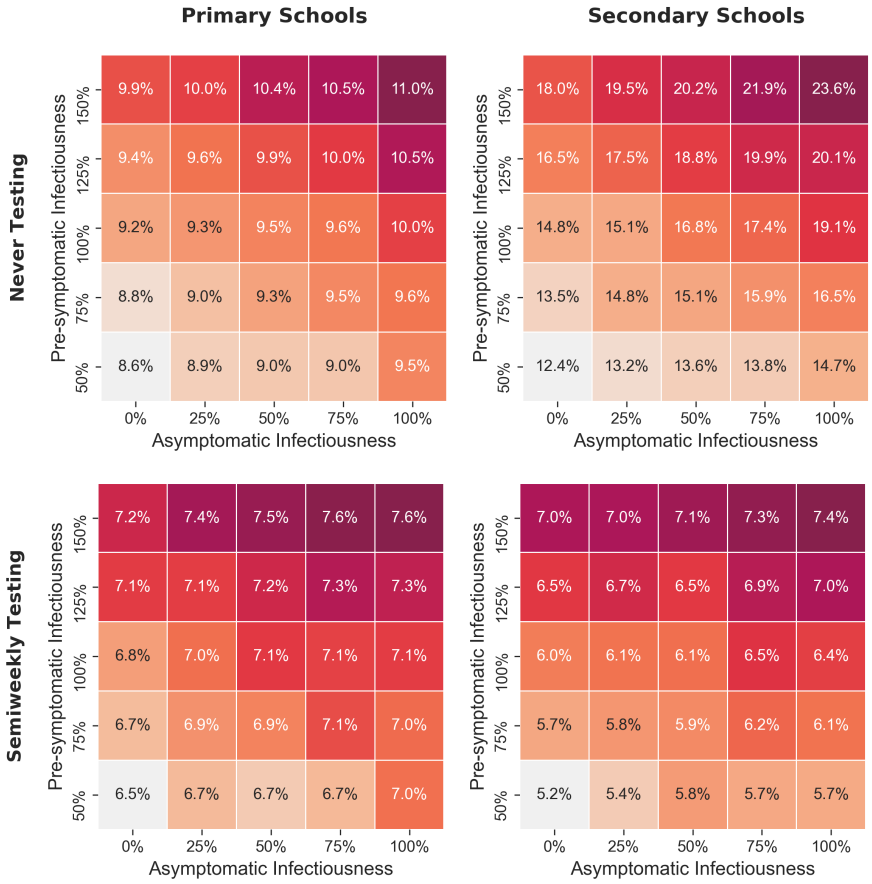


FIGURE 3.C.17. **Increasing asymptomatic and presymptomatic trasnsmissibilities compound to increase overall transmission.** The heatmap cells give the average number of school transmission cases shown as a percentage of the total school population for combinations of pre-symptomatic and asymptomatic infectiousness relative to symptomatic infectiousness. Outcomes are shown for primary schools (left column) and secondary schools (right column) in scenarios with no testing (top row) and semiweekly testing (bottom row). Cases are introduced into the school once a week on average, no other mitigations are in place, and baseline values are used for all other parameters as outlined in Table 4.B.2. The color maps of each heatmap have been independently scaled to highlight within-treatment sensitivity to pre-symptomatic and asymptomatic infectiousness.

cases within the pre-symptomatic period due to the false negative rate of the test being higher during this period.

3.C.2. Population mixing (global transmission). Disease transmission may occur either from close contacts (local transmission) defined by the contact network structure or from casual contacts (global transmission). Close contacts are individuals with whom one has repeated, sustained, or close proximity interactions on a regular basis, while casual contacts are individuals with whom one has incidental, brief, or superficial contact. Casual contacts are assumed to be randomly sampled from the population as in a well-mixed population. In the context of a school population, a student's close contacts might include their teacher(s), other students in

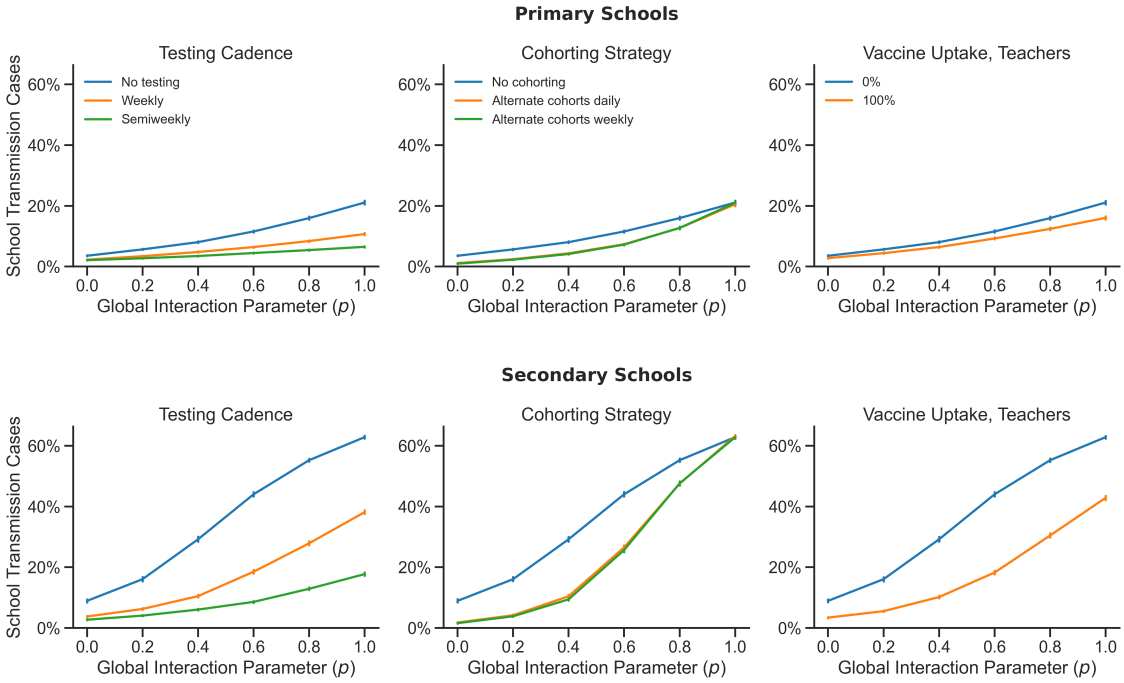


FIGURE 3.C.18. **Case increase as transmission shifts from primarily along the contact network to more globally among weak contacts.** The average number of school transmission cases are shown as a percentage of the total school population for a range of global transmission values.

their class(es) or activities, their friends, and their siblings. Global transmission may represent transmission among students, teachers, or staff that interact incidentally in class, in the cafeteria, in the hall, on the bus, etc. A network locality parameter p sets the relative weight of interaction among casual (global) contacts in the population. This parameter can be seen as setting the amount of mixing in the population, where higher values of p—more global interaction—correspond to more random encounters among individuals throughout the population. In the paper we assume that 80% of transmission occurs between close contacts on the contact network and 20% of transmission occurs globally (i.e., $p = 0.2$).

Figure 3.C.18 shows the average number of school transmission cases for primary and secondary schools across a range of global interaction weights, p. When $p = 0$ there is no global interaction, and all transmission occurs between close contacts on the contact network. As p increases, the population becomes increasingly mixed as more and more transmission occurs globally rather than on the network. When $p = 1$ the population is well-mixed, and the dynamics of transmission reduce to a stochastic implementation of a standard mass action compartment model. As the population becomes more mixed, the effect of stochasticity on the fates of transmission chains decreases, and the probability increases that a single case introduction will seed an outbreak that proceeds all the way to the herd immunity threshold. Since the influence of the contact network is diminished in a relatively mixed population, the efficacy of cohorting, which is defined as a partitioning of close contacts among students, decreases as the extent of global interaction increases.

These results suggest that schools should make efforts to limit the amount of mixing and opportunities for global transmission in their populations. This might include restructuring lunch periods, passing periods, transportation logistics, etc. in order to limit congregation and interaction among groups of otherwise "unconnected" individuals at school.

3.C.3. Differences between primary and secondary schools. Across parameter ranges and interventions, we find that compared to secondary schools, primary schools consistently have fewer school transmission cases and lower probabilities of sizable outbreaks. In our model, the key characteristics that differentiate primary and secondary schools are (1) the susceptibility of students and the (2) the sturcture of contact networks. In this section we investigate the relative contribution of these factors to the differences in outcomes across these settings.

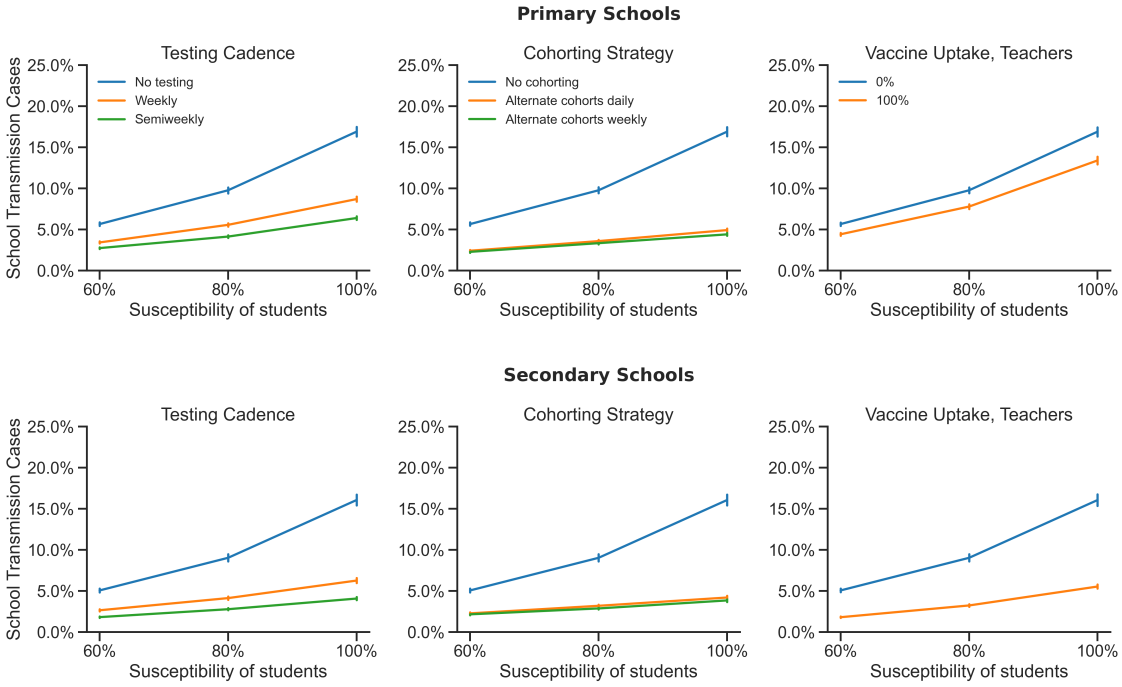


FIGURE 3.C.19. **When primary and secondary school students are equally susceptible to infection, primary schools and secondary schools exhibit similar case rates.** The average number of school transmission cases are shown as a percentage of the total school population, for a range of student susceptibilities relative to adults, using primary school (upper row) and secondary school (lower row) network structures. Baseline values are used for all other parameters as outlined in Table 4.B.2. The strong similarities between the primary and secondary school plots for any given level of susceptibility indicate that the differences in network structure have minimal effects.

3.C.3.1. ***Susceptibility of students.*** A fundamental difference between primary and secondary school populations is the age of the students. Evidence suggests that primary school-aged children are less susceptible to COVID-19 than adults, while secondary school-aged adolescents are approximately equally susceptible to

adults.Goldstein et al. 2020, Viner et al. 2020 In the main text, we have assumed that primary school students are 60% as susceptible as adults, and secondary school students are 100% as susceptible. Given the lower susceptibility of the majority of individuals in primary school populations, we would expect primary schools to have fewer and smaller outbreaks. This is borne out in our model as well as in empirical data.Goldstein et al. 2020, Ismail et al. 2020a, Stein-Zamir et al. 2020

Figure 3.C.19 shows the effect of varying student susceptibility from 60% to 100% that of adults in both primary and secondary schools. We see that given that same level of student susceptibility, primary and secondary schools have similar outcomes and show similar relative benefits to mitigation measures. When primary school students are 100% as susceptible as adults, outcomes in primary schools closely resemble those of secondary schools. When the susceptibility of secondary school students is reduced to 60% that of adults, the outcomes in secondary schools resemble those of primary schools. These patterns suggest that student susceptibility is a key determinant of the epidemiological outcomes in schools. In the next section we will see that outcomes are relatively insensitive to the differences in contact network structure between these two school settings, which supports the conclusion that student susceptibility is the an important driver of epidemiological differences between primary and secondary schools.

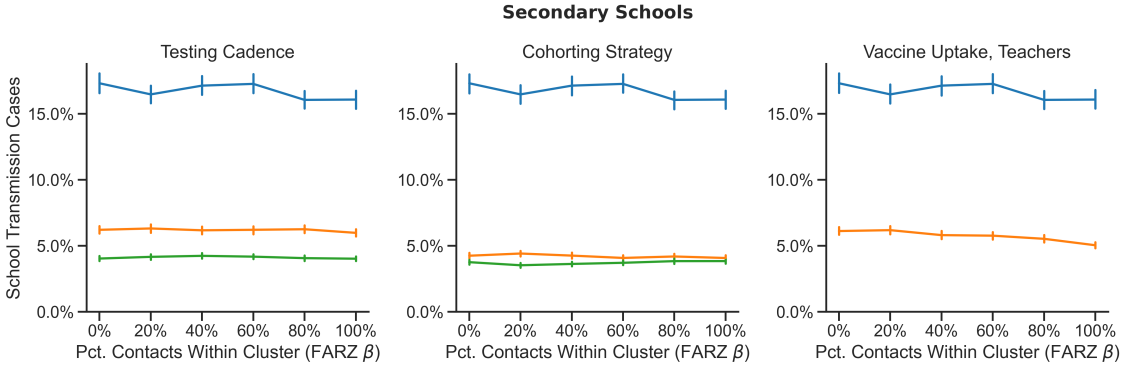


FIGURE 3.C.20. **In secondary schools, transmission is not sensitive to the network structure of within-grade interactions.** In the secondary school model, the FARZ network generator is used to generate a student contact network layer for each grade level. Students are assigned to network clusters, and the FARZ β parameter sets the probability of within-cluster edge formation. The average number of school transmission cases are shown as a percentage of the total school population for a range of β values, with higher values representing more modular networks.

3.C.3.2. ***Modularity of contact networks.*** Another major difference between primary and secondary schools is the structure of the contact network that represents the patterns of interaction in these populations. Perhaps the most significant difference between these two characteristic contact networks is their modularity. Primary schools are organized into classroom groups that interact closely throughout the day, which represents a highly modular contact structure. In secondary schools, students attend multiple classes each day, teachers interact with multiple groups of students, and students participate in activities outside of class—all of which contribute to a less modular interaction network. Indeed, the contact network used in our primary school

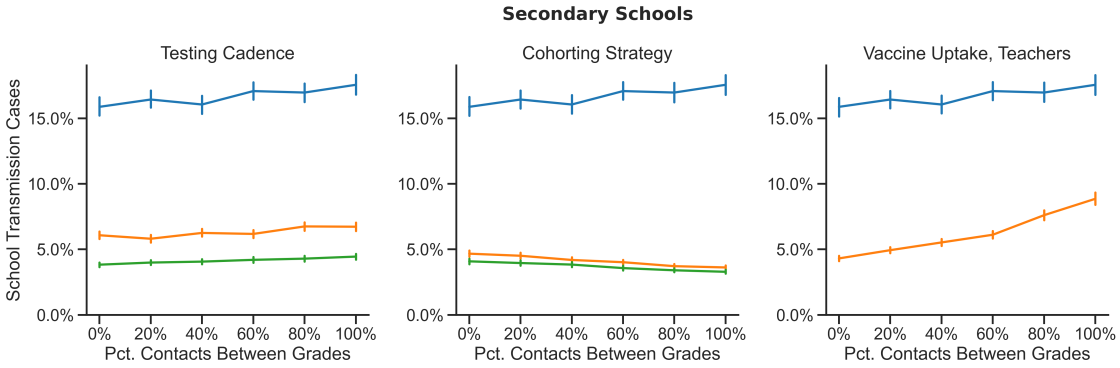


FIGURE 3.C.21. **In secondary schools, transmission is only slightly sensitive to the extent of between-grade interactions.** In the secondary school model, a student contact network layer is generated for each grade level. A parameterized fraction of each student's total number of edges are designated for between-grade interactions. The average number of school transmission cases are shown as a percentage of the total school population for a range of between-grade connectedness values, with lower values representing more modular student contact networks.

model is significantly more modular than the secondary school network (mean clustering coefficients of 0.91 and 0.16, respectively). In general, population modularity tends to dampen the severity of epidemics.Badham and Stocker 2010, Danon et al. 2012, Read et al. 2008 When individuals belong to more insular groups, they have fewer sources of possible exposure when susceptible and a more restricted and more correlated pool of potential infectees when contagious. This expectation is consistent with our findings that primary schools tend to have fewer and less severe outbreaks than secondary schools, but how much of the difference that we observed is attributable to network structure rather than other factors?

To answer this question we probe changes to the secondary school network that make it more modular and thus more similar to the primary school network. The secondary school contact network consists of five interconnected FARZ network layers: a student layer for each grade level and a teachers-and-staff layer (Appendix 3.B.4). The FARZ network generator assigns each student to 1-2 network clusters, which may be thought of as groups of students that take the same classes, participate in the same activities, belong to the same social circle, or are otherwise more likely to interact and share contacts with each other than with other students. The FARZ β parameter sets the expected fraction of each student's close contacts that are within their own network cluster. Therefore this parameter modulates the modularity of the student layer for each grade level. When $\beta = 1$ students belong to highly modular groups that interact amongst themselves, and when $\beta = 0$ these groups are degenerate as students interact exclusively with students from other groups. The secondary school contact network is also parameterized by the fraction of each student's contacts that are between grade levels. Decreasing the amount of inter-grade interaction makes the overall network more modular, and vice versa. Our baseline secondary school model features a contact network where 80% of each student's contacts are with students in the same grade, and 80% of those within-grade contacts are with students within their own network cluster.

Figure 3.C.20 and Figure 3.C.21 show the effect of varying these modularity parameters on the average number of school transmission cases in secondary schools. Network modularity is increased by increasing the

FARZ β parameter or by decreasing the fraction of inter-grade contacts. We find that the model outcomes are largely insensitive to changes in these parameters. This suggests that differences in the modular structure of students' interaction networks are not a substantive driver of the differences in outcomes in primary and secondary schools. Even when the student-student contact network is highly modular, other forms of interaction forge transmission chains across groups, including teacher-student interactions, household transmission, and global transmission due to incidental interactions. We have previously seen that model outcomes are highly sensitive to the degree of population mixing associated with global transmission. Therefore, efforts to restructure secondary school students into modular groups may not be as fruitful as other strategies, such as limiting general mixing of the student population (e.g., in the cafeteria, hallways, common spaces, buses) or cohorting strategies that limit the number of students on campus at one time.

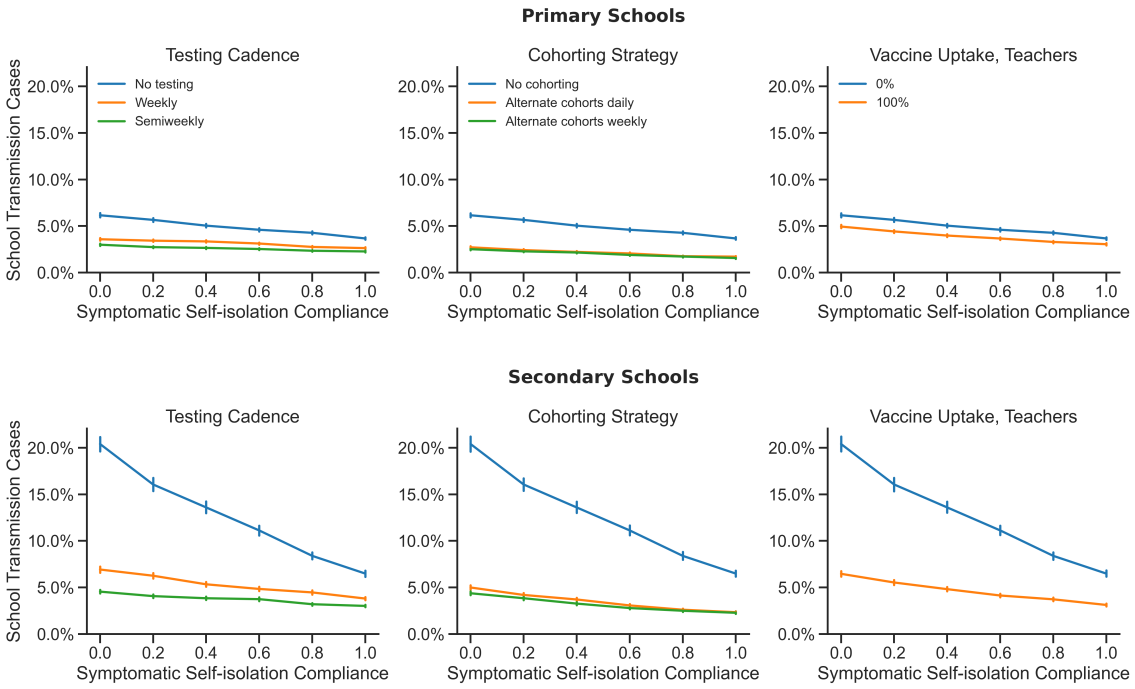


FIGURE 3.C.22. **In the absence of testing, self-isolation compliance is critical to reduce cases, particularly in secondary schools.** Testing detects individuals who are infected but non-compliant, and thereby reduces the importance of the compliance rate. The average number of school transmission cases are shown as a percentage of the total school population for a range of symptomatic self-isolation compliance rates.

3.C.4. Compliance with isolation upon symptoms. The key to mitigating the risk of large outbreaks is to isolate cases early in order to limit the number of days that any infectious individual is at large and potentially exposing new cases. If individuals self-isolate upon the onset of symptoms, this will help considerably. However, some individuals may not recognize COVID symptoms, or may choose not to comply with a mandated isolation period. Figure 3.C.22 shows the effect of the individual compliance rate on the average number of school transmission cases. In the absence of other interventions (blue lines), the rate of self-isolation has a

substantial impact on outcomes. In the main text we assume that 20% of individuals self-isolate at the onset of symptoms. This value follows from a finding that professional symptom screening fails to identify nearly half of cases in children with viral loads comparable to symptomatic individuals,Poline et al. 2020 and a further assumption that a fraction of individuals that do recognize their symptoms will nevertheless attend school for various reasons. Our results may overestimate the number of school transmission cases if compliance with self-isolation upon symptoms is substantially higher in a given population. In any case, once any proactive mitigation strategy (i.e., testing, cohorting, or vaccination) is introduced, the model outcomes become largely insensitive to the rate of self-isolation compliance. These results suggest that decision makers would be wise to incentivize compliance with symptomatic self-isolation, but they also highlight the fact that coordinated, proactive strategies reduce the risk of outbreaks without relying on individual behavior.

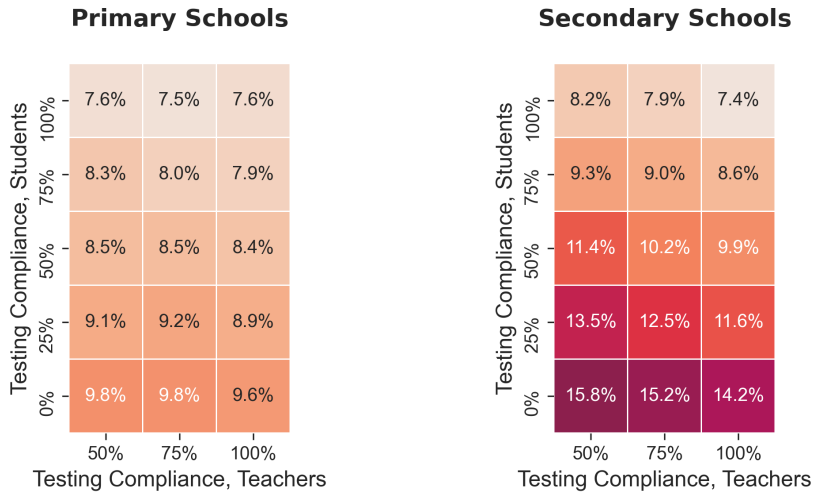


FIGURE 3.C.23. **Compliance with testing by teachers and by students improves outcomes.** Heatmap cells give the average number of school transmission cases. Cases are introduced to the school once a week on average, and baseline values are used for all other parameters as outlined in Table 4.B.2.

3.C.5. Compliance with testing. Our results show that frequent proactive testing is effective at mitigating the risk of sizable outbreaks, but testing is most effective if a large percentage of the population complies with the testing cadence. In the paper we assume that 100% of teachers and staff are tested and that 75% of students comply with testing. We further assume that compliance with testing is a longitudinal property of individuals, rather than determined on a test-by-test basis. Figure 3.C.23 shows the effect of varying the rates of student and/teacher compliance with testing on the average number of school transission cases. Unsurprisingly, there are fewer school transmissions on average when a larger percentage of individuals are tested. Outcomes appear to be more sensitive to the level of testing compliance among students than teachers, but this is because students make up a larger fraction of the school population, so drops in student compliance remove a larger absolute number of individuals from the testing pool. The effectiveness of proactive testing is optimized by maximizing the number of individuals that are screened as scheduled. Providing students with adequate incentives to comply with testing forms a necessary component of this strategy.

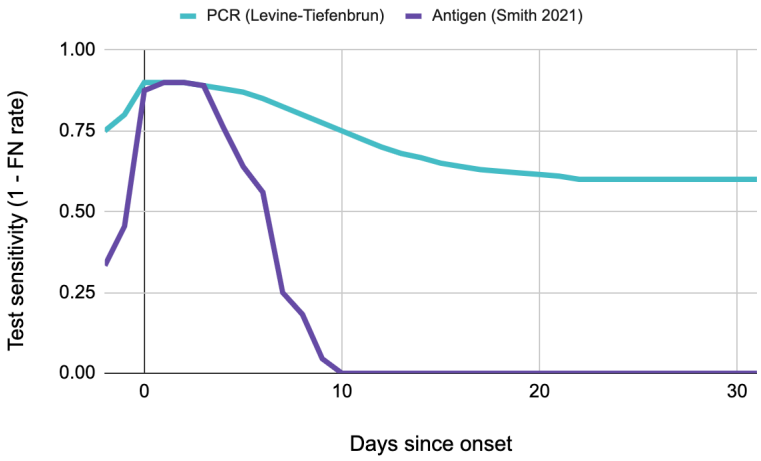


FIGURE 3.C.24. **Test sensitivities vary over the course of infection for RNA-based and antigen methods.** In modeling test sensitivity, we use the following curves for the probability of returning a positive test result for RNA-based tests (teal line, Levine et al. 2020Levine-Tiefenbrun et al. 2020) and antigen tests (purple line, Smith et al. 2021Smith et al. 2021) for infectious individuals as a function of days until or since the onset of the symptomatic or asymptomatic period. The test sensitivity is equivalent to 1 minus the false negative rate.

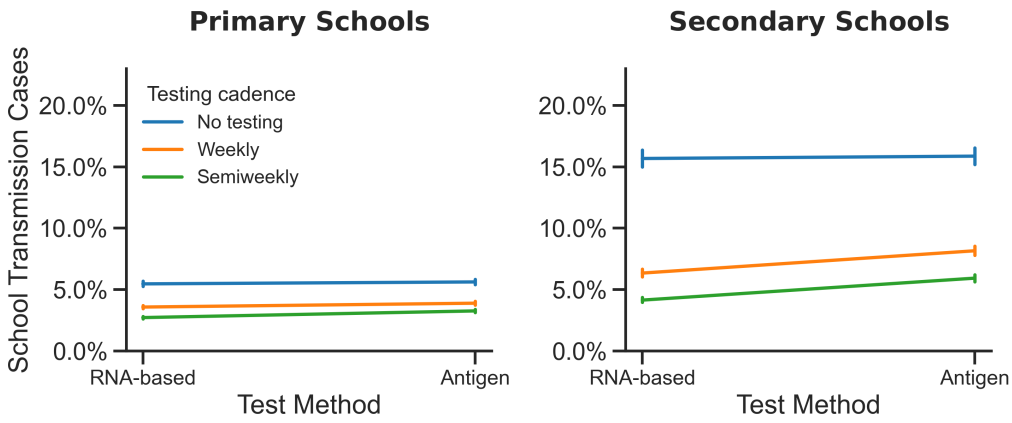


FIGURE 3.C.25. **The benefits of testing are relatively insensitive to testing method**. The average number of school transmission cases are shown as a percentage of the total school population where test sensitivities vary over time according to data corresponding to RNA-based testsLevine-Tiefenbrun et al. 2020 or rapid antigen tests.Smith et al. 2021 RNA-based tests show a modest advantage in a secondary school setting such that weekly RNA-based testing is comparable to twice-weekly antigen-based testing. This is largely because RNA-based tests are more effective at detecting infection early on during the presymptomatic period. .

3.C.6. Test sensitivity (false negative rate). The success of a proactive testing strategy depends in part on the sensitivity (false negative rate) of the testing methodology used. Test sensitivity changes over time according

to disease state of the individual and the amount of time spent in a given disease state (Appendix 3.B.6.2). In the main text we base our temporal test sensitivities on data collected for RNA-based tests.Levine-Tiefenbrun et al. 2020 Schools may also consider using rapid antigen tests, which are less sensitive than RNA-based tests but are faster and less expensive. Antigen tests have similar peak sensitivities to RNA-based tests around the onset of symptoms, but their sensitivity drops off more quickly after onset and they are less likely to detect pre-symptomatic cases.Smith et al. 2021

Figure 3.C.25 compares proactive testing strategies using RNA- and antigen-based testing methods with sensitivity curves drawn from Levine et al. 2020Levine-Tiefenbrun et al. 2020 and Smith et al. 2021,Smith et al. 2021 respectively (Figure 3.C.24). Smith et al. report sensitivities for antigen tests based on their agreement with results from RNA-tests administered at the same time. Here we cap the antigen peak sensitivity values at the corresponding RNA-based sensitivity values for a standardized comparison. Both testing methods significantly reduce school transmission when used weekly or semiweekly. Antigen tests lead to slightly worse outcomes because they are less likely to detect cases during the infectious pre-symptomatic period or late into an individual's infectious period. This difference is more significant in secondary schools where there are generally more cases to detect. In secondary schools, outcomes for weekly RNA-based testing are about equal to those for semiweekly antigen testing.

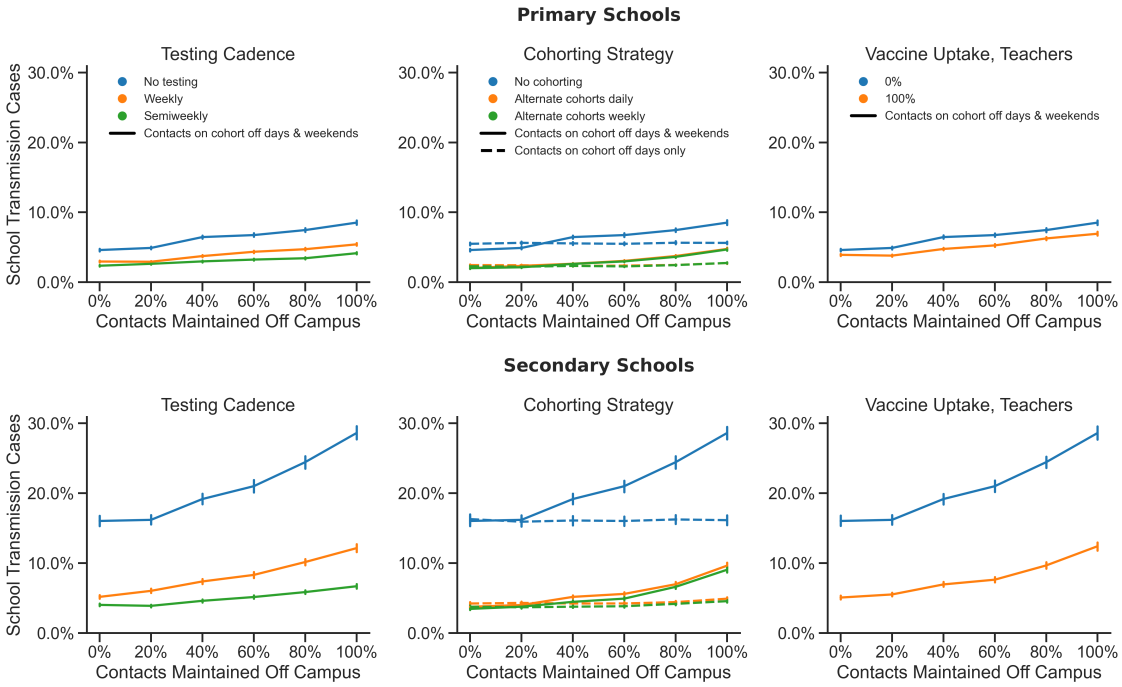


Figure 3.C.26. **Weekend interactions among students increase the size of outbreaks.** The average number of school transmission cases are shown as a percentage of the total school population for a range of fractions of school contacts maintained by students while off campus. Outcomes for scenarios where students maintain contacts on weekends as well as cohort off days (where applicable, center column) are shown with solid lines, and outcomes where students maintain contacts on cohort off days only are shown in dashed lines.

3.C.7. Out-of-school contacts. When some or all students are out of school, the number of interactions among the school population decreases. If a student is out of school during their infectious period, they are limited in the number of secondary cases they can generate and may fail to transmit disease entirely. Weekends provide a break in exposures between members of the school population, and many other interventions, including cohorting and isolation upon symptoms or positive test, involve keeping students out of school at strategic times. However, if students continue to interact with members of the school population while not in school, then the effectiveness of these exposure breaks will be reduced.

In Figure 3.C.26 we consider the effect that out-of-school interactions among students have on the number of transmission cases in the school population. In the main text we assume that a student makes no close contact with any member of the school population (with the exception of any students that share their household) when they are off campus on weekends, cohort off days, or isolation days. Students make incidental contacts with members of the general school population (i.e., global transmission) at the same rate while off campus (e.g., at the grocery store). Here we allow students to maintain some fraction of their close contacts on weekends and on cohort off days where applicable. On off-days, students interact with a fixed random subset of the close contacts they have at school. On weekends, students may maintain contact with any student they interact with in school. Students that are off campus due to cohorting only maintain contact with other students in their cohort. When students continue to interact at high levels on weekends, the number of school transmission cases increases, which is to be expected given that the total potential exposure time increases by 40%. Interestingly, out-of-school contacts on cohort off days alone do not have a significant effect on the benefit of the cohorting strategy (dashed lines, center column in Figure 3.C.26). This suggests that the division of students into two or more groups that never make close contact is sufficient to significantly limit transmission even if these groups continue intra-group interactions while out of school, provided that the cohort boundaries are respected.

3.C.8. Vaccine effectiveness. The success of a vaccination strategy depends on the uptake of the vaccine and its effectiveness in preventing disease and transmission. Uptake and effectiveness have a product relationship, so evaluating a range of values for one parameter or the other gives a good idea of the sensitivity of the model to the effectiveness of a vaccination program in general. We consider the impact of vaccinating all teachers and staff on mitigating transmission in the school population as a whole. Figure 3.C.27 shows the influence of vaccine effectiveness on the success of this strategy. Vaccine effectiveness refers to the probability that a vaccinated individual will mount an immune response that consequently blocks disease and infectiousness (Appendix 3.B.6.5). Here, we assume that an effective vaccination blocks 100% of an individual's propensity for transmission.Levine-Tiefenbrun et al. 2021, Petter et al. 2021 More effective vaccines prevent more school transmission cases on average, but substantive reductions in the mean outbreak size can be achieved even with vaccines that have effectiveness of 50% or lower. Similarly, this suggests that good results can be achieved even with moderate uptake of a highly effective vaccine.

Vaccinating teachers and staff has a greater effect in secondary schools. This is likely because teachers and staff make up a larger fraction of their school populations, interact with more students, and are more likely to transmit to students given the increased susceptibility of secondary school students relative to primary school students.

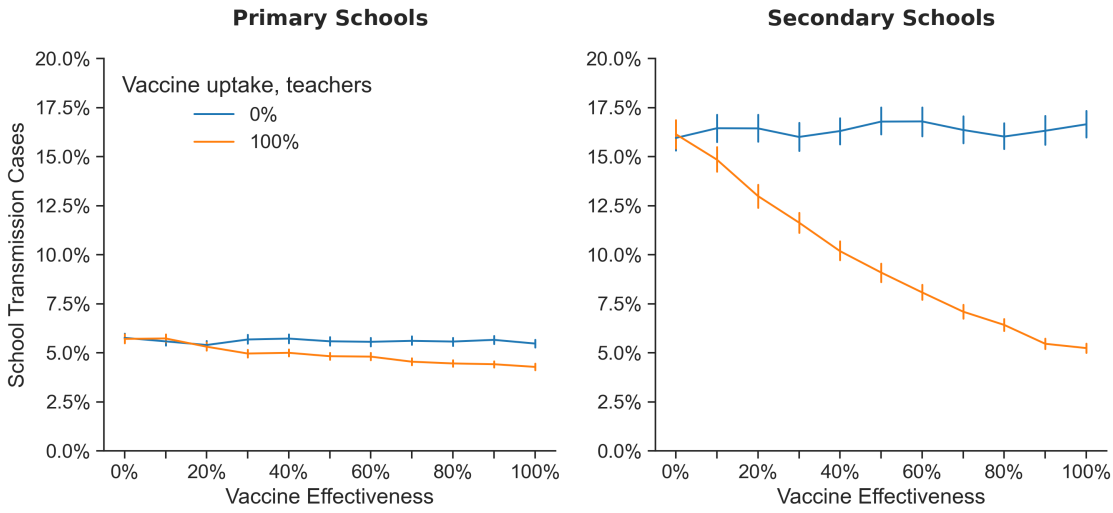


FIGURE 3.C.27. **The benefit of vaccinating teachers in a secondary school setting is sensitive to the efficacy of the vaccine.** The average number of school transmission cases are shown as a percentage of the total school population for a range of vaccine effectiveness values.

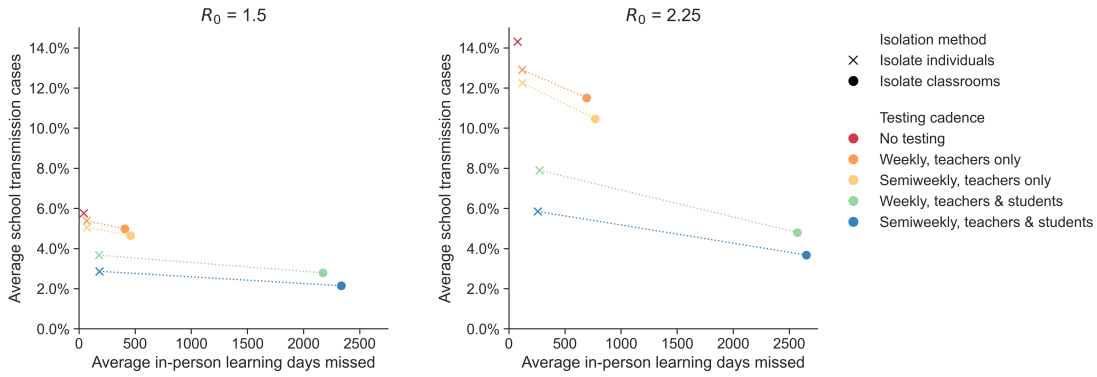


FIGURE 3.C.28. **Isolating entire classrooms can reduce cases, at the expense of missed days of school.** The average number of school transmission cases (as a percentage of the total school population) and the average number of in-person learning days missed (i.e., isolation person-days) are shown for testing and isolation strategies that involve individual-level isolation (crosses) or classroom-level isolation (circles). Results for both isolation policies are shown for five testing cadences, including no testing. A dashed line connects the points for individual-level and classroom-level isolation for each cadence; the slopes of these lines illustrate the relative impact on one outcome variable versus the other. Outcomes are shown for primary schools with weekly introductions and no cohorting or vaccination. Baseline values are used for all other parameters as outlined in Table 4.B.2.

3.C.9. Classroom isolation. We find that isolating all of an infected student's classroom contacts can mitigate the risk of sizable outbreaks in primary schools (Figure 3.B.11), where students are organized into stable classroom units. While this strategy can prevent school transmission, it also requires that additional in-person

learning days are lost upon each case detection. When choosing a quarantine strategy, the costs of in-person learning days lost, both by students and by teachers, are likely to be of interest to school decision makers.

Figure 3.C.28 shows how both school transmission and in-person learning days missed are impacted by adopting a classroom-level, rather than individual-level, isolation policy. Classroom-level isolation leads to fewer school transmission cases on average, as seen in the negative slope of the lines connecting the points for these two isolation methods for all testing cadences. However, the total number of in-person learning days lost does go up markedly when isolating classrooms rather than individuals.

Figure 3.C.28 also shows that when school transmission is relatively low ($R_0 = 1.5$) the marginal reduction in school transmission that is conferred by adopting classroom-level isolation with weekly testing (e.g., moving from the green cross to the green circle) is nearly equal to the reduction that is achieved by switching to semiweekly testing and isolating only positive individuals (e.g., moving from the green cross to the blue cross). Semiweekly testing leads to a negligible increase in in-person learning days missed over weekly testing, so increasing the testing cadence may be more attractive than adopting a classroom-level isolation policy in some contexts. When transmissibility is higher (right panel, $R_0 = 2.25$), we observe a greater benefit to isolating classrooms, as indicated by the steeper slopes of the lines. .

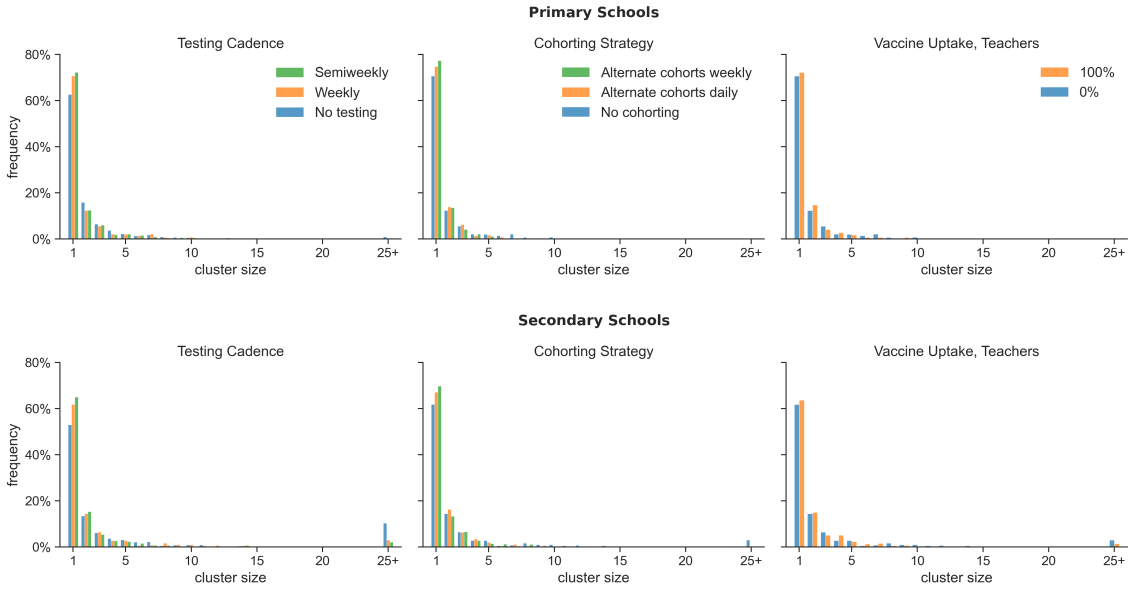


FIGURE 3.C.29. **Distribution of case clusters sizes for $R_0 = 1.5$.** Histograms give the frequency of case clusters of different sizes in 1000 replicate simulations for each treatment. Data are shown for primary schools (top) and secondary schools (bottom) in scenarios with no mitigation (blue bars), weekly or semiweekly testing of students and teachers (left column), alternating student cohorts on a weekly or daily schedule (center column), and vaccination of teachers (right column). Cases are introduced to the school weekly on average, and baseline values are used for all other parameters as outlined in Table 4.B.2.

3.C.10. Cluster Sizes. Throughout this paper, we report the epidemiological outcomes for schools in terms of the total number of school transmission cases and the probability of outbreaks affecting more than 5% of

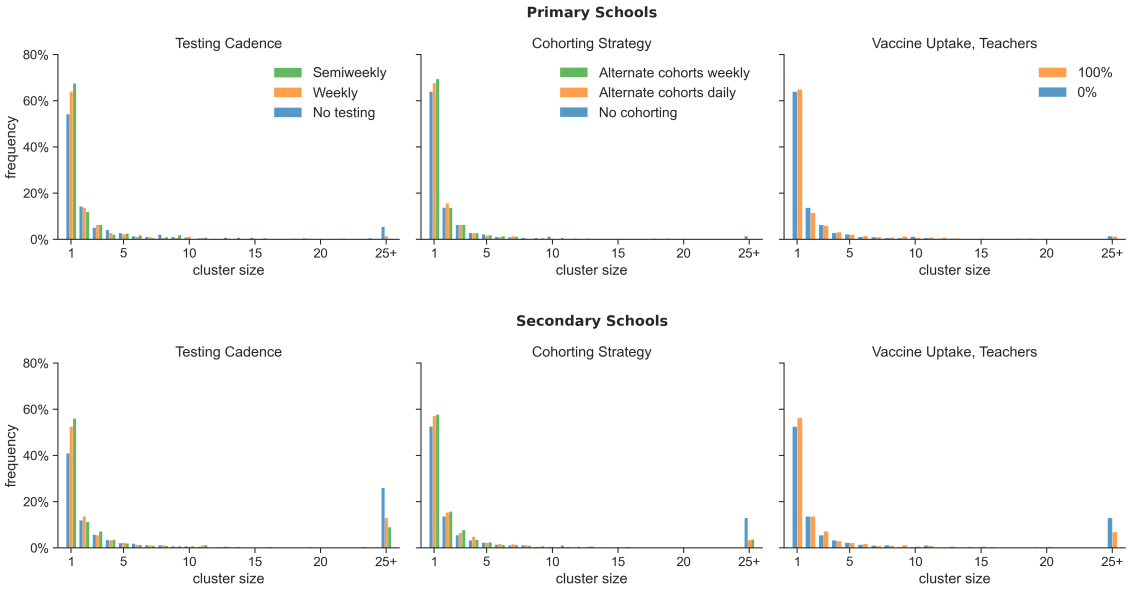


FIGURE 3.C.30. **Distribution of case cluster sizes for $R_0 = 2.25$.** Histograms give the frequency of case clusters of different sizes in 1000 replicate simulations for each treatment. Data are shown for primary schools (top) and secondary schools (bottom) in scenarios with no mitigation (blue bars), weekly or semiweekly testing of students and teachers (left column), alternating student cohorts on a weekly or daily schedule (center column), and vaccination of teachers (right column). Cases are introduced to the school weekly on average, and baseline values are used for all other parameters as outlined in Table 4.B.2.

the school population over the course of a semester. These summary statistics convey outcomes that are of primary interest to school decision makers: the expected number of individuals that may become infected at school and the likelihood of an outbreak that causes a major disruption to the health and operations of the school. However, these statistics abstract away more detailed information about transmission chains that can offer insights about the epidemic dynamics. The jitter figures in the main text illustrate the stochastic distribution of school outbreak sizes, and the bimodality of these distributions provides a clue about the underlying dynamics: most transmission chains trains appear either to contribute few to no secondary cases, or alternatively to develop into a large outbreak that may approach the herd immunity threshold if left unchecked.

This observation is confirmed by Figure 3.C.29 and Figure 3.C.30, which show the distribution of case cluster sizes in primary and secondary schools under different interventions for $R_0 = 1.5$ and $R_0 = 2.25$, respectively. A cluster consists of a primary case that was exposed in the community and any secondary school transmission cases that can be traced back to that primary case. Across both schools and all interventions, a majority of clusters consist of only a single case, which represents an exogenous case introduction that did not generate any secondary cases in the school. In general, smaller clusters are more common than larger clusters. Large clusters (e.g., consisting of 25 individuals or more) are more frequent than intermediate-sized clusters in secondary schools, where school transmission is more likely due to the susceptibility of secondary school students and other factors. When R_0 within the school is high (e.g., $R_0 = 2.25$, Figure 3.C.30), very large clusters become more common in both settings (which reflects the bimodality seen in the jitter

distributions of outcomes for scenarios where $R_0 = 2.25$, e.g., Figure 3.3 in the main text), and singleton clusters decrease in frequency. Testing, cohorting, and vaccination all increase the frequency of terminal primary cases and decrease the frequency of very large clusters, but these interventions do not dramatically impact the distribution of cluster sizes aside from those extremes.

CHAPTER 4

Proactive COVID-19 testing in a partially vaccinated population.

The SARS-CoV-2 virus responsible for COVID-19 is particularly difficult to control because transmission often occurs in individuals who are pre-symptomatic or entirely asymptomatic He et al. (2020b), Kimball et al. (2020), Rivett et al. (2020). Throughout the COVID-19 pandemic, proactive testing has been a valuable tool for mitigating the spread of SARS-CoV-2 in workplaces, academic institutions, and other settings Black et al. (2020), Denny et al. (2020), Larremore et al. (2021), McGee et al. (2021). The aim of proactive testing is to head off transmission from individuals who are not showing symptoms, by detecting infection early and isolating infected persons from the rest of the community.

Widespread COVID-19 vaccination is now underway in many parts of the world, ushering in the prospect of a full return to in-person work. As more individuals become vaccinated, can we afford to reduce or eliminate SARS-CoV-2 proactive testing from our disease control programs? Here we aim to address two aspects of this question. First, what level of vaccination and/or naturally-acquired immunity is required to render proactive testing programs unnecessary Motta et al. (2021)? Second, as we transition from present conditions to that point, what are best practices for tapering off testing efforts?

Methods

To address the questions above, we use two distinct modeling approaches. First, we adapt a simple analytic approximation developed by Bergstrom et al. (2020) to examine how testing and vaccination interact to reduce transmission. We illustrate these interactions as isoclines, or indifference curves, that indicate how increased vaccination coverage can compensate for reduced testing in a population.

Second, we deploy the SEIRS+ modeling framework (McGee 2020) used in our workplace testing (Color Health 2020) and return-to-school (Appendix 3, McGee et al. (2021)) models to consider workplaces or other groups in which vaccination efforts are underway. SEIRS+ is a stochastic, network-based epidemiological simulation model that accounts for the specific details of SARS-CoV-2 transmission and for the structure of social contact networks along which most infections are spread. We use the SEIRS+ model to simulate the dynamics of COVID-19 spread through a group of 1,000 individuals, following a single introduction from the community. We consider how the fraction of the population vaccinated and the extent of pre-existing natural immunity play into COVID-19 transmission dynamics.

The exposure ratio model. To better understand the benefit of proactive SARS-CoV-2 testing in a partially vaccinated workplace or university setting, we adapt the Bergstrom et al. (2020) analytic model to explore how vaccines and testing interact to reduce opportunities for COVID-19 transmission.

The model estimates how much the effective reproduction number R_e—the average number of secondary cases generated by each primary case under current conditions—is reduced by a given level of vaccine uptake and a given cadence of proactive testing. While in practice this depends on the complex biology of the virus and the social dynamics in the host population, we can derive a simple approximation as follows. We envision that during each day of the infectious period, an infected individual is either "at-large" in the community or isolated at home. We assume that at-large individuals transmit at a constant rate corresponding to the effective reproduction number, while individuals isolated at home do not transmit. Testing serves to identify pre-symptomatic, asymptomatic, or paucisymptomatic individuals who would not have otherwise self-isolated. After testing positive, these individuals remove themselves from the population at large and isolate at home, which reduces the number of days they may expose others in the community. Suppose that testing the population every n days reduces the average amount of time that infectious individuals are at large by the *exposure ratio* $\mathcal{Q}(n)$. The effective reproduction number is then reduced by a comparable fraction (Appendix 4.A).

We now add vaccination and immunity from previous infection. Suppose that the basic reproductive number for the disease is R_0, a fraction γ of the population is effectively vaccinated, and a fraction η is immune due to previous infection. If vaccination occurs independent of prior infectious status, the reproductive number with vaccination, testing, and previous immunity is given by

$$R_e = R_0 \, \mathcal{Q}(n) \, (1 - \gamma) \, (1 - \eta). \tag{4.1}$$

In words, the equation states that the effective reproduction number is equal to the product of the basic reproduction number, the fraction of exposure days that remain despite testing, the fraction of the population not yet effectively vaccinated, and the fraction of the population not yet naturally immune. In Appendix 4.A, we sketch out the analytic approach used to develop this equation, and provide a mathematical expression for the exposure ratio $\mathcal{Q}$. Further details and a full derivation are provided in Bergstrom et al. (2020).

The SEIRS+ model. Our open source Python framework SEIRS+ implements stochastic network models of infectious disease transmission (`https://github.com/ryansmcgee/seirsplus`). This model extends the classic SEIRS model in a number of ways: it accounts for pre-symptomatic and asymptomatic disease states, it incorporates a process of quarantine or self-isolation, it accounts for individual heterogeneity in disease parameters, it allows an arbitrary distribution of residence times in each disease state, and it models transmission as occurring along a social contact network. We will briefly describe each of these aspects in turn. The SEIRS+ model has been used by a number of research groups (Karin et al. 2020b, McGee et al. 2021, Patil et al. 2021, Ross et al. 2021), and detailed documentation for can be found on the SEIRS+ github wiki (`https://github.com/ryansmcgee/seirsplus/wiki`).

Classic SEIR models of infectious disease are compartment models with compartments for susceptible (S), exposed (E), infectious (I), and removed (R) individuals. The SEIRS+ platform extends this framework to reflect the biology of SARS-CoV-2. In particular, the infectious class is divided into separate compartments for individuals who are in pre-symptomatic, asymptomatic, and symptomatic disease states. This allows us to model the decisions that some individuals make to self-isolate based on symptoms.

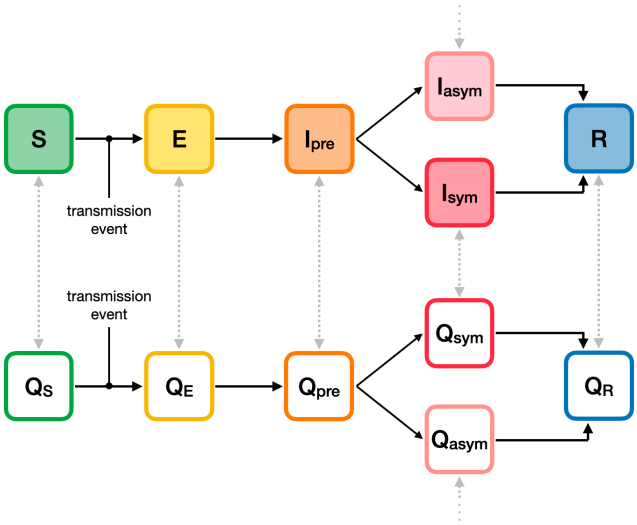


FIGURE 4.1. **The SEIRS+ model extends the structure of a basic SEIR compartment model.** In addition to the traditional S, E, and R states, the I state is split into three separate compartments: presymptomatic I_{pre}, asymptomatic I_{asym}, and symptomatic I_{sym}. In addition, each primary compartment (shaded) has a parallel compartment Q (unshaded) representing individuals at the same disease stage but in quarantine.

Because an important component of COVID-19 control involves isolating individuals in response to symptoms or testing, the SEIRS+ model introduces quarantine compartments that represent individuals in self-isolation (Figure 4.1). An individual may be quarantined in any disease state, and every disease state has a corresponding quarantine compartment. Quarantined individuals follow the same progression through the disease states, but their set of close contacts is defined by a reduced contact network. The transitions into quarantined states occur as the result of a positive test or self-diagnosis of symptoms. Individuals remain in the quarantine set of compartments until the designated isolation period is completed—ten days in this model. Thereafter they are moved into the non-quarantine compartment corresponding to their current disease state.

In the SEIRS+ model, each individual in the model is independently assigned model parameters: for example, infectiousness, symptomaticity if infected, and the duration of time spent in each compartment if infected. This allows us to explicitly represent heterogeneity in disease characteristics (Ball 1985, Hickson and Roberts 2014).

Traditional SEIR models assume that individuals transition among compartments at fixed rates, resulting in exponentially distributed residence times in each compartment. In reality, the distribution of residence times can be very different—and these differences can matter for modeling purposes (Feng et al. 2007, Krylova and Earn 2013). The SEIRS+ model allows residence times drawn from gamma distributions based on empirical data. The dynamics governing disease transmission are proportional to contact rates between susceptible and infectious individuals, accounting for variation in infectiousness, and are described in detail in refs. McGee et al. (2021) and McGee (2020).

Contact networks. In their standard form, epidemic compartment models assume mass-action dynamics of disease transmission, equivalent to assuming that each individual in the population is equally likely to interact

with any other individual. In practice, interactions are highly structured according the physical layout of workplaces as well as by the social contact networks that individuals inhabit (Potter et al. 2015, Weeden and Cornwell 2020), and this can have important consequences for the dynamics of disease spread (Bansal et al. 2007, Keeling 2005). To account for this, the SEIRS+ model specifies a contact network of individuals who are prone to have "close contact" with one another. Those in one's network might include family members, close friends, coworkers, roommates, romantic partners, and so forth. The model assumes that much—though not all—disease transmission occurs among the close contacts represented in the contact network. In the model, SARS-CoV-2 can also be transmitted among "casual" contacts, individuals who make infrequent, brief, and incidental contact such as might occur when passing through the same hallway or shopping in the same store. Here, we assume that 80% of transmission occurs along the network of close contacts, and the remaining 20% is casual in that it occurs at random among all members of the population (Mossong et al. 2008).

The probability that a susceptible individual becomes infected depends on how susceptible they are to infection, how many of their contacts become infected, and how transmissible those infected contacts are. In a network transmission model, some individuals have numerous connections with others, and some have few. People with large numbers of contacts are unlikely to interact as closely with each one of them individually. To account for this, we assume a logarithmic rather than linear scaling of transmission opportunity as a function of network degree (McGee et al. 2021).

Model parameters. In modeling the effects of testing and vaccination, we face multiple sources of uncertainty. Among these, the value of the basic reproduction number R_0 looms largest. Under the traditional definition, R_0 is the mean number of secondary infections generated by an index case in a susceptible population. For our purposes, the relevant value of R_0 is the expected number of transmissions that occur within the institutional setting we are modeling—the number of transmissions that occur at work, for example. This value will depend on the physical structure of the workplace and the interaction patterns therein. Specifically, it will depend on the non-pharmaceutical interventions (NPIs) in place, including masking, ventilation, basic hygiene, and distancing procedures. And it will depend on the strains of SARS-CoV-2 circulating in the community. Because all of these factors vary from workplace to workplace and from week to week, obtaining a precise estimate of workplace R_0 is seldom feasible. As such, the results of this model are more useful for understanding general trends than for making precise quantitative predictions.

The vaccines currently available in the US have demonstrated efficacy of around 65–95% against the wild-type form of SARS-CoV-2 (Baden et al. 2020, Dagan et al. 2021, Haas et al. 2021, Oliver et al. 2021, Polack et al. 2020, Swift et al. 2021, Thompson et al. 2021). The most widely-distributed vaccines—the mRNA-based ones produced by Pfizer and Moderna—fall at the high end of that efficacy range. Thus far, these vaccines show comparable effectiveness against most circulating variants of concern as well (Abu-Raddad et al. 2021, Charmet et al. 2021, Chemaitelly et al. 2021, Lopez Bernal et al. 2021). Available evidence increasingly suggests that these vaccines block SARS-CoV-2 symptoms and transmission at similar rates (Birhane et al. 2021). Here we assume an average effectiveness of 90% against symptoms and transmission alike, intended to represent the mixture of vaccines available to the populations we are modeling. Appendix 4.C shows model results when effectiveness is reduced to 70%.

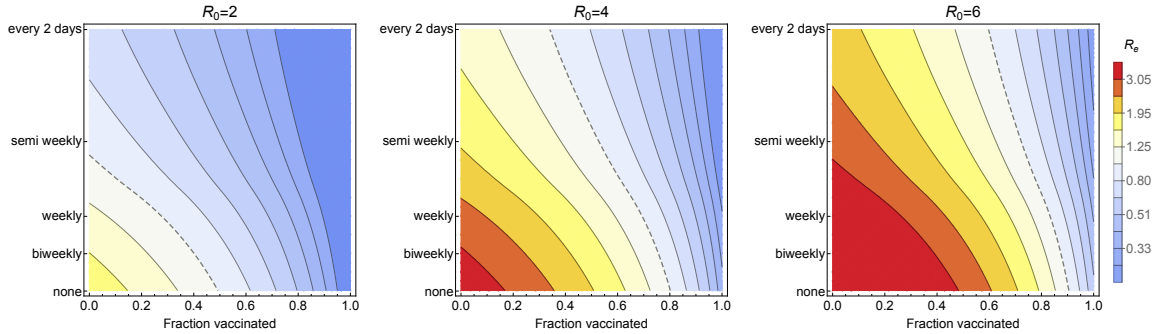


FIGURE 4.2. **Effects of NPIs, vaccination, and testing**. Contour plots of the effective reproduction number R_e show how mitigation efforts depend on NPIs via their effects on R_e, on vaccination coverage (horizontal axis), and on testing cadence (vertical axis). The dashed line in each panel shows the combinations of vaccination and testing that are sufficient to drop R_e to unity for a given R_0 value. (a) When $R_0 = 2.0$, as we might expect for less transmissible variants with active NPIs, the effective reproduction number can be brought below unity with semi-weekly testing, by vaccinating half of the population, or some combination of those interventions. (b) With more transmissible variants and NPIs relaxed so that R_0 rises to 4.0, broad vaccination coverage is an essential component to any control strategy and testing can help get down below the $R_e = 1$ threshold when vaccinate rates are in the range of 50-80%. (c) With highly transmissible variants and minimal NPIs, reducing R_e below one is impossible without very high levels of vaccination. In all three panels, we illustrate a situation in which 10% of the population has been previously infected and vaccines average 90% effectiveness. Additional parameters are summarized in Table 4.B.2.

Results

Analytic approximation. Chance plays an important role in the dynamics of COVID-19 outbreaks. Even if two workplaces have very similar conditions, the introduction of an infected individual may seed a large outbreak in one and lead to no new cases in the other. But as a rule of thumb, an index case is unlikely to seed a sizeable and protracted outbreak when $R_e < 1$. Thus one prerequisite for safely operating a workplace is that mitigation measures therein are sufficient to drop R_e below unity.

To understand how non-pharmaceutical interventions, vaccination adoption, and proactive testing cadence contribute to reducing the likelihood of an outbreak, in Figure 4.2 we examine contour plots of the effective reproduction number R_e for three different R_0 values. The horizontal axis in each panel represents the fraction of the population that have been vaccinated against COVID-19. The vertical axis indicates the cadence of proactive testing across the same population. Along each of the isoclines (solid black lines), the combined effect of proactive testing and vaccination is constant, i.e., each isocline corresponds to a fixed R_e value. Once R_e falls below unity, as indicated by the dashed line, substantial outbreaks are unlikely.

Panel 4.2a illustrates an R_0 value of 2.0, reflecting a situation in which NPIs such as masks and distancing remain in place and in which the predominant SARS-CoV-2 strain is the original non-variant virus. Panel 4.2b illustrates an R_0 value of 4.0, as might be the case in a workplace where more transmissible variants are circulating yet NPIs have been relaxed. Panel 4.2c illustrates an R_0 of 6.0, as we might see where the Delta strain is predominant and NPIs are limited or lacking entirely (Davies et al. 2021, Public Health England 2021).

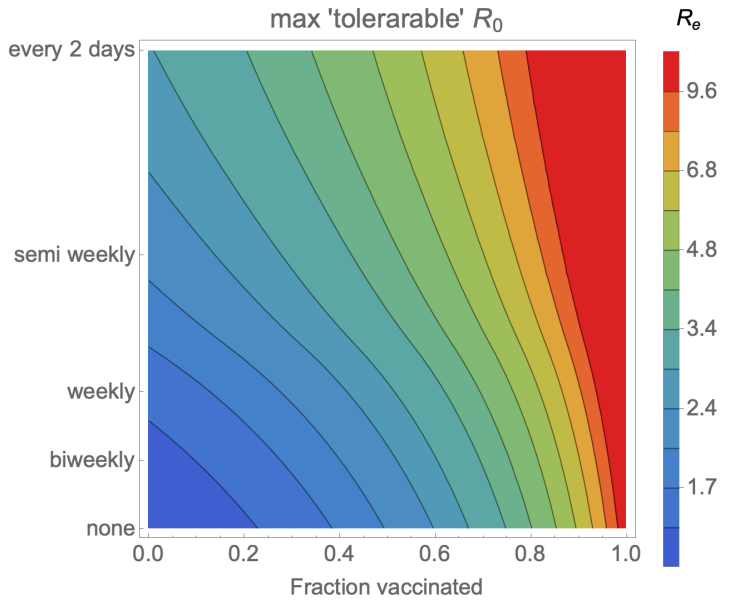


FIGURE 4.3. **Maximum controllable R_0 depends on vaccination and testing rates.** Contour plot illustrates the highest R_0 for which R_e can be brought down below 1.0 by means of the specified combination of vaccination coverage (horizontal axis) and testing cadence (vertical axis). Here 10% of the population has been previously infected and vaccines average 90% effectiveness.

Taken together, these three plots illustrate the synergistic effects of NPIs, vaccination, and testing. The more stringent the NPIs, the further the effective reproduction number R_e is reduced. Frequent testing and broad vaccination coverage also reduce R_e. The higher the initial level of transmission, the greater the amount of testing and vaccination required to mitigate the risk of outbreaks.

When vaccine coverage is insufficient on its own to reduce the effective reproductive number R_e below one, a combination of proactive testing and NPIs can help considerably. For example, we see in Figure 4.2a that semiweekly testing has a comparable impact to vaccinating half the population. As vaccine coverage becomes more extensive, the effect of testing on R_e declines and eventually the benefit of testing becomes marginal.

In each of the three panels in Figure 4.2, R_0 was fixed at a specific value and we looked at how a combination of proactive testing and vaccination reduces R_e. Another way to look at the same mathematical expression is to fix the target level of R_e at 1, the level below which a sustained outbreak is unlikely, and then plot what value of R_0 can be tolerated for a given vaccination level and testing rate (Figure 4.3). This graph allows a number of useful comparisons. For example, a vaccination rate of 60% might be enough to control the original SARS-CoV-2 strains even without supplementing with proactive testing. However, if more highly transmissible strains become established, driving the basic reproduction number upward toward 4 or even higher, vaccinating 60-70% of the population will not be sufficient to control disease on its own. Stronger NPIs and/or frequent, broad-scale proactive testing will be necessary.

Stochastic network-based simulation. The analytical model above relies on a highly simplified picture of disease dynamics. To account for many of the complexities of the real world—superspreading, social contact

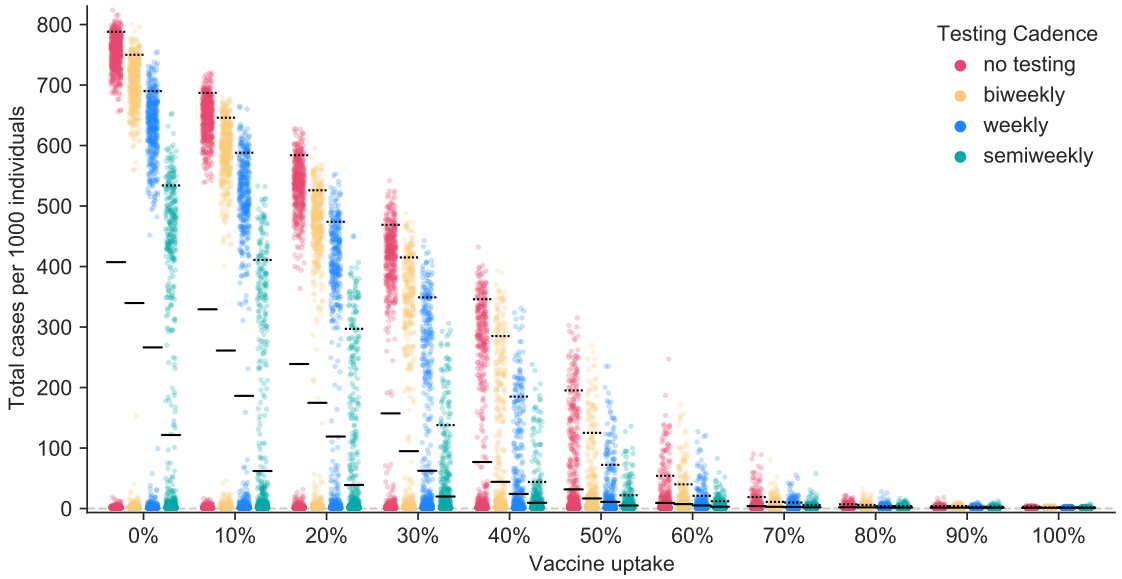


FIGURE 4.4. **Outbreak sizes in the stochastic network simulation model.** Here we illustrate the outcome of 1,000 simulations for each combination of testing cadence and vaccination uptake, when 10% of the population have previously been infected, vaccines are on average 90% effective at preventing infection and transmission, and $R_0 = 4.0$. Each colored dot represents the outcome of a single simulation run. Solid black lines mark the mean outbreak sizes and dashed black lines mark the 95th percentile outbreak sizes for each parameter combination.

networks, variation from person to person in disease progression, and the role of chance in disease outbreaks—we turn to the SEIRS+ simulation model.

Using this model, we consider the consequences of a single introduction into a workplace or other congregate setting of 1,000 individuals. Figure 4.4 illustrates the distribution of outbreak sizes resulting directly from this single introduction, at various testing cadences, as vaccine adoption increases. For each combination of parameters, we run 1,000 replicate simulations and depict the outbreak sizes as jitter plots where each dot represents the outcome of a single simulation run. The mean and 95th percentile outbreak sizes are indicated by the solid and dashed bars, respectively.

At lower levels of vaccine adoption and testing, we see a bimodal distribution of outcomes. In some simulation runs, large outbreaks occur, while in other runs with the same parameter settings, little or no transmission takes place at all. This is the consequence of chance events that contribute to the trajectory of disease spread. This figure also reveals that when less than two thirds of the population has been vaccinated, testing is a powerful tool for reducing both the mean number of cases and the 95th percentile outbreak size.

Figure 4.5 illustrates how the average benefit of testing declines as more of the population becomes vaccinated. The benefit of testing is measured as the reduction in the mean number of individuals infected after a single introduction. In the absence of vaccination and with $R_0 = 4.0$, for example, weekly proactive testing prevents on average about 140 cases after a single introduction into a population of 1,000 people. By the time 70% of the population have been vaccinated, weekly testing prevents only a few cases. In general, we see

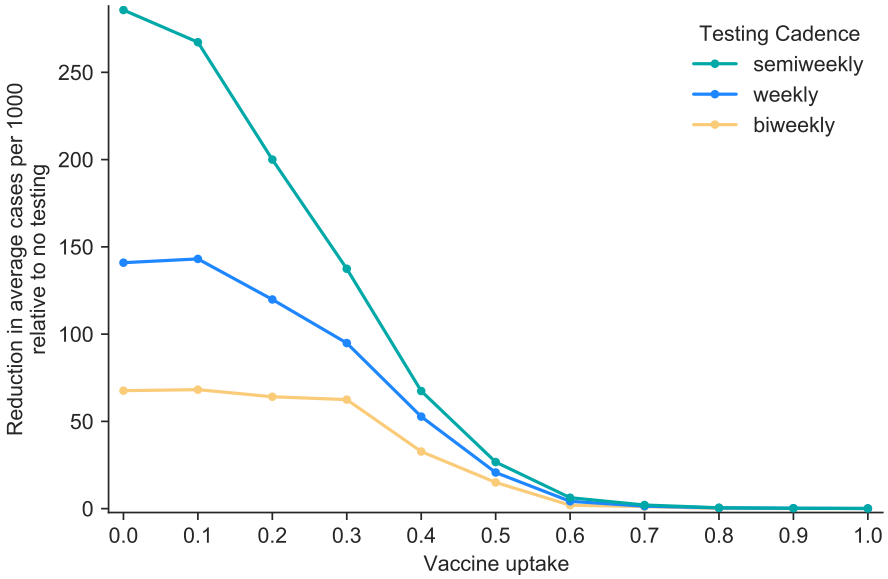


FIGURE 4.5. **Value of testing in the stochastic network simulation model**, as more of the population becomes vaccinated. Here we show how many cases per 1000 individuals are prevented on average when using a given testing cadence relative to not testing, where $R_0 = 4.0$, 10% of the population have previously been infected, and vaccines are on average 90% effective at preventing infection and transmission.

that (1) when vaccination is limited, more frequent testing confers greater benefits, and (2) once vaccination becomes very common, the benefits of testing are substantially diminished.

Because Figures 4.4 and 4.5 illustrate the consequences of a single introduction, a combination of vaccine coverage and testing rate that poses an acceptable risk when community prevalence is low and introductions are rare may nonetheless be intolerably risky when community prevalence is high and frequent introductions are likely to occur.

Figures 4.4 and 4.5 illustrate the results of simulations for a single set of parameters: $R_0 = 4.0$, prior incidence of disease 10%, and an average vaccine effectiveness of 90%. While these may be reasonable approximations of the situation in many locales, the general patterns that we observe are robust to changes in the precise parameter values. To illustrate, we have developed an interactive web application that displays results from the SEIRS+ model across a wide range of parameters (Color Health 2021). This application allows the user to select values of R_0, vaccine efficacy, prior disease incidence, and testing cadence, and presents a jitter plot akin to Figure 4.4 for the chosen values.

Discussion

In this report, we have presented the results from two different modeling approaches: an analytic approximation and a stochastic network-based simulation. These approaches generate concordant results. Both indicate that testing is valuable at lower levels of vaccine adoption, regardless of R_0. When R_0 is high, either due to more transmissible strains or relaxation of non-pharmaceutical interventions, high levels of vaccination

or a combination of vaccination and proactive testing will be necessary to achieve the $R_e \leq 1$ threshold. Once the effective reproduction number R_e drops below 1 without testing, proactive testing offers little additional value and can be suspended if preventing outbreaks is the sole objective (see also Motta et al. (2021)).

This is not to say that proactive testing is useless in even more highly vaccinated populations. One-off transmission events are still possible even when the effective reproductive number $R_e < 1$ and a substantive outbreak is unlikely to occur. Proactive testing can reduce the probability of such transmissions occurring. Thus in addition to surveillance use, proactive testing can be helpful in situations where any transmission event whatsoever is considered unacceptable.

In previous modeling work, we found that testing generally offers a high return-on-investment in an unvaccinated workplace population (https://www.color.com/covid-19-outbreak-model). Here we find that the benefits of testing decline as vaccine adoption increases, but can be sizeable even in partially vaccinated cohorts.

In principle, one could adjust the testing cadence in real time, slowing the cadence as increasing fractions of the population become vaccinated. Given the inevitable uncertainties about the exact value of R_e, and the logistical complexity of adjusting testing cadence in real time, it is likely to be difficult to fine-tune testing cadence by reducing the testing rate gradually as more people are vaccinated. As such, it may be logistically simpler to continue at testing at the original pre-vaccination testing cadence until there is good reason to believe that $R_e < 1$. At that point, the proactive testing program can be halted entirely.

Appendix 4.A. *An analytic approximation for the effects of testing and vaccination*

In the absence of vaccination, we define the exposure ratio $\mathcal{Q}(n)$ as the ratio of the mean exposure days when testing every n days, $\bar{E}(n)$, to the mean exposure days without any testing $E_\emptyset$:

$$\mathcal{Q}(n) = \frac{\bar{E}(n)}{E_\emptyset}. \tag{4.2}$$

We calculate this quantity as follows. Even without testing, some individuals will choose to self-isolate after showing symptoms. If a fraction $1-u$ of infected people develop symptoms and a fraction v of these isolate once symptomatic, the fraction who are symptomatic and choose to isolate is given by $v(1-u)$. The fraction of infected people who are asymptomatic or choose not to self-isolate is $1-v(1-u)$. Suppose that the presymptomatic infectious period is y days and the total infectious period is C days. In the absence of testing, the average number of exposure days is

$$E_\emptyset = \big(1-v(1-u)\big)C + v(1-u)\,y. \tag{4.3}$$

Let $E(t,n)$ be the expected number of infectious days spent at large for a person who is infectious for t days when testing occurs every n days. For an individual who would not otherwise self-quarantine, testing reduces the number of infectious days at large from C to $E(C,n)$. For an individual who would have self-quarantined anyway, testing can still pick up the infection before symptoms appear and thus reduces infectious days at large from y to $E(y,n)$. The mean number of exposure days when testing every t days is then

$$\bar{E}(n) = \big(1-v(1-u)\big)E(C,n) + v(1-u)E(y,n). \tag{4.4}$$

Now all we need is a way of calculating $E(t,n)$. When the testing cadence is slower than the infectious period C minus the reporting delay d for the disease, an individual will be tested and results returned at most once during the course of infection. Bergstrom et al. (2020) show that in this case, i.e. when $n > C-d$,

$$E(C,n) = C - \frac{(1-q)(C-d)^2}{2n}. \tag{4.5}$$

When an individual may be tested twice or more during the course of a single infection, the corresponding equation is somewhat more complicated. Denote the standard floor function by $\lfloor\cdot\rfloor$, let $\bar{x} = \left\lfloor \frac{C-d}{n} \right\rfloor$, and let $r = C - d - n\,\bar{x}$. When the testing cadence is faster or equal to infectious period minus the reporting delay, $n \le C-d$,

$$E(C,\tau) = \left(\frac{n}{2} + d + \frac{n\,q}{1-q}\right) + q^{\bar{x}}\left(\frac{1}{2n}\left(qr^2 - (n-r)^2\right) - \frac{n\,q}{1-q}\right). \tag{4.6}$$

If infectiousness is constant throughout the duration of infection, testing every n days reduces the expected number of individuals infected by an index case by a fraction $\mathcal{Q}(n)$. Thus if the basic reproduction number without testing is R_0, testing every n days would reduce this to $\tilde{R} = \mathcal{Q}(n)\,R_0$

Next we incorporate vaccination and natural immunity. We assume that individuals who have been effectively vaccinated or who have recovered from previous COVID infection will not transmit COVID. If a fraction γ of the population has been effectively vaccinated, a fraction η have recovered from previous infection, and vaccination occurs independently of past infection status, the effective reproductive number with testing, vaccination, and previous infectious is given by

$$R_e = R_0\,\mathcal{Q}(n)\,(1-\gamma)\,(1-\eta). \tag{4.7}$$

Appendix 4.B. *Overview of parameter values*

Here we summarize the parameter values used in the analytic model (Table 4.B.1) and in the SEIRS+ simulations (Tables 4.B.2–4.B.4).

Table 4.B.1. Parameter values for the analytical approximation

Parameter	**Value or values**
Basic reproduction number R_0	2.0, 4.0, and 6.0 as indicated in figure captions
Test sensitivity	90% during the infectious period irrespective of sympaticity. 0% prior to beginning of infectious period.
Delay between testing and quarantine	1 day for those who test positive.
Infectious period	6.2 days (irrespective of symptomaticity)
Presymptomatic period	2.2 days (among those who eventually show symptoms)
Presymptomatic transmissibility	100% relative to symptomatic transmissibility
Asymptomatic transmissibility	100% relative to symptomatic transmissibility
Fraction of asymptomatic cases	30%
Compliance with isolation upon symptoms	30%
Compliance with testing	100%
Vaccine effectiveness	90%
Rate of disease with effective vaccine	0%
Rate of transmission with effective vaccine	0%
Fraction immune due to prior infection	10%

In the analytic model, we assume that test sensitivity is constant at 90% throughout the infectious period. In the SEIRS+ model we assume RNA-based testing with test sensitivities that change over the course of infection, as in McGee et al. (2021). Test sensitivity is 0% during the exposed period and non-infectious period, then climbs to 75% sensitivity for individuals in the first 2 days of their pre-symptomatic period and 80% sensitivity for any pre-symptomatic days beyond that. After the presymptomatic period ends, sensitivities

TABLE 4.B.2. Overview of parameter values for the SEIRS+ model

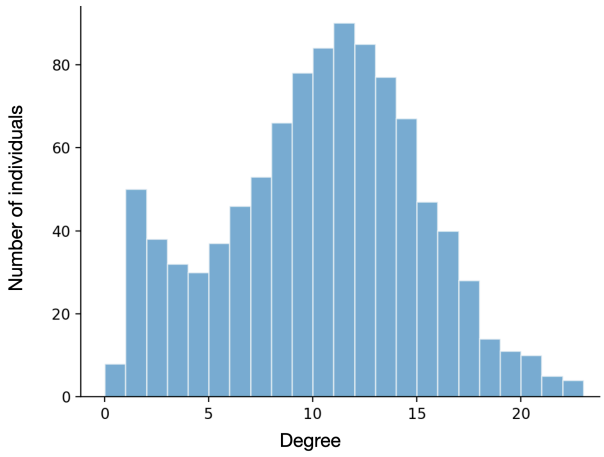
Parameter	**Value or values**
Basic reproduction number R_0 (mean individual reproduction number $R_0^{(i)}$)	4.0
Individual $R_0^{(i)}$ coefficient of variation	2.0 (overdispersed)
Fraction of asymptomatic cases	30%
Compliance with isolation upon symptoms	30%
Compliance with testing	100%
Disease state periods	*See Table 4.B.4*
Network structure	*See Table 4.B.3*
Proportion of global (off-network) transmission	20%
Test sensitivity	Time-varying: see figure 4.B.6
Vaccine effectiveness	90% in main text, 70% in Appendix 4.C.
Rate of disease with effective vaccine	0%
Rate of transmission with effective vaccine	0%
Fraction immune due to prior infection	10%

for symptomatic and asymptomatic individuals alike follow the time course shown in Figure 4.B.6, with values based upon Levine-Tiefenbrun et al. (Levine-Tiefenbrun et al. 2020). We assume there are no false positives.

Appendix 4.C. *Lower vaccine efficacy*

The SEIRS+ simulation results shown in Figures 4.4 and 4.5 of the main text are based on an average vaccine effectiveness of 90%. In some areas, the vaccines used may have lower effectiveness. To provide a sense of how that changes the results, here we show SEIRS+ simulation results for the case in which vaccine effectiveness is 70%. In this situation, testing remains valuable at higher vaccination fractions, but the overall patterns are qualitatively similar.

TABLE 4.B.3. Network property statistics averaged across 1,000 replicate randomly generated contact networks and the degree distribution histogram for a representative random contact network.

Network properties	**Degree distribution**
Degree mean: 9.9 Degree std: 5.0 Degree CV^2: 0.25 Degree assortativity: 0.25 Clustering coeff.: 0.44 Average path length: 3.7	

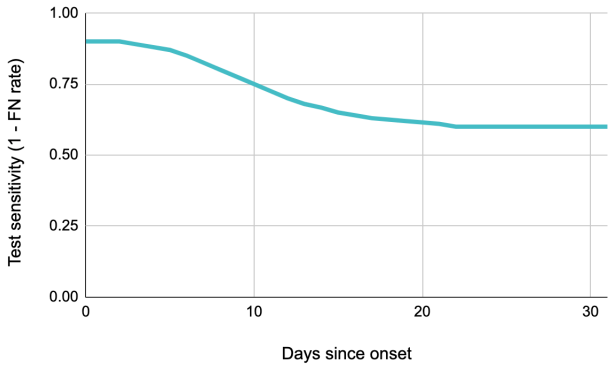
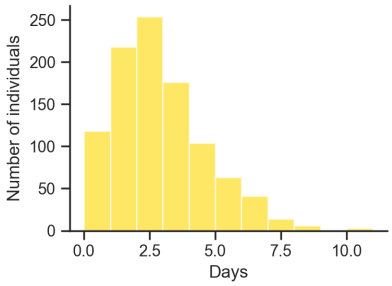
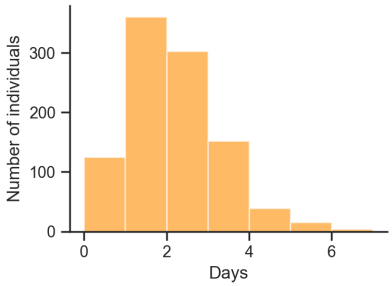
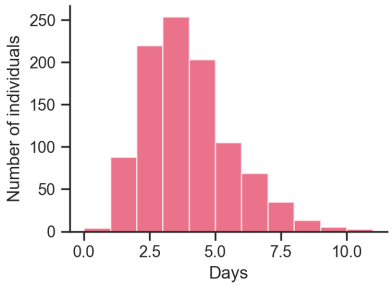
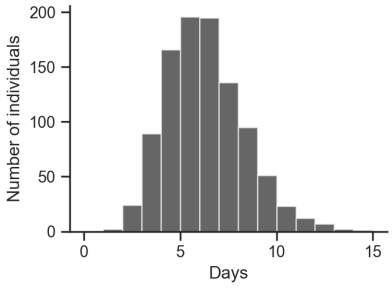


FIGURE 4.B.6. **Test sensitivity in the SEIRS+ model.** The probability of returning a positive test results when testing an infected, symptomatic individual as a function of the number of days since entering that state. The test sensitivity is equivalent to 1 minus the false negative rate.

TABLE 4.B.4. A representative distribution of period values drawn for population of 1,000 individuals is shown for each parameter in the center column below. Statistics across all replicate distributions in our analysis are shown in the rightmost column.

Disease state period	**Distribution**	**Statistics**
Latent period (time in E state)	 `gamma(mean=3.0, CV=0.6)`	mean 3.0 days std 1.8 days 95% CI (0.6, 7.4)
Pre-symptomatic period (time in I_{pre} state)	 `gamma(mean=2.2, CV=0.5)`	mean 2.2 days std 1.1 days 95% CI (0.6, 4.8)
Symptomatic period (time in I_{sym} or I_{asym} state)	 `gamma(mean=4.0, CV=0.4)`	mean 4.0 days std 1.6 days 95% CI (1.5, 7.7)
Total infectious period (total time in I_{pre}, I_{sym}, and I_{asym} states)	 `gamma(mean=2.2, CV=0.5)`	mean 6.2 days std 1.9 days 95% CI (3.0, 10.5)

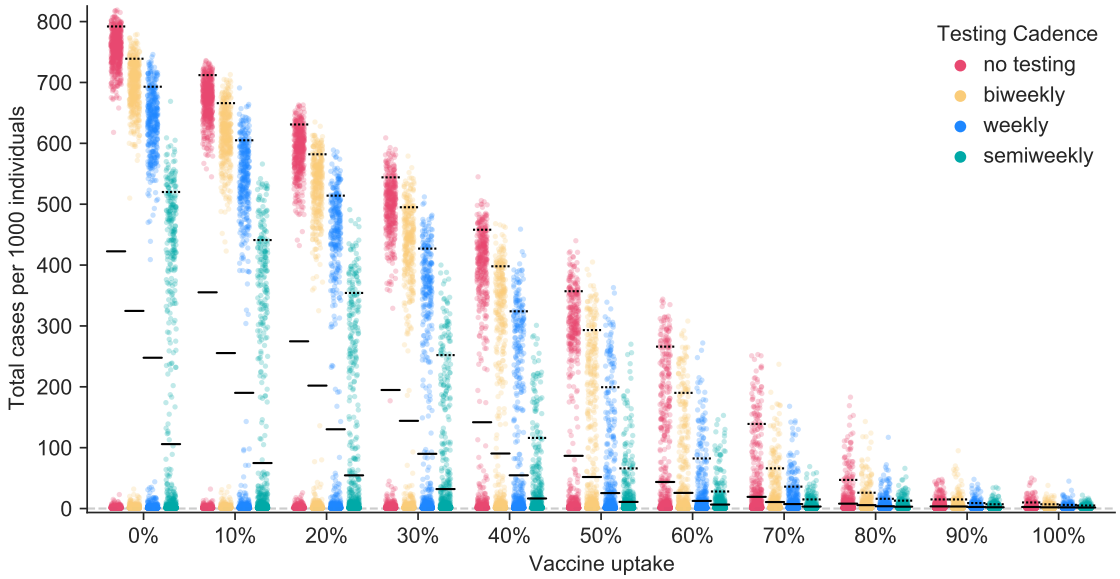


FIGURE 4.C.7. **Outbreak sizes when vaccine effectiveness is 70%** Here we illustrate the outcome of 1000 simulations for each combination of testing cadence and vaccination uptake, when 10% of the population have previously been infected, vaccines are on average 70% effective at preventing infection and transmission, and $R_0 = 4.0$. Solid black lines mark the mean outbreak sizes and dashed black lines mark the 95th percentile outbreak sizes for each parameter combination.

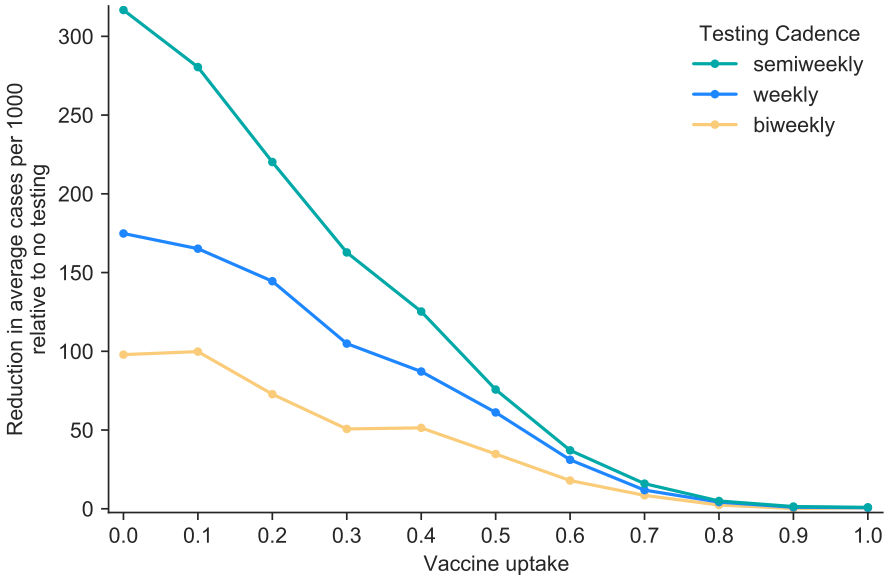


FIGURE 4.C.8. **Value of testing when vaccine effectiveness is 70%**. Here we show how many cases per 1000 individuals are prevented on average when using a given testing cadence relative to not testing, where $R_0 = 4.0$, 10% of the population have previously been infected, and vaccines are on average 70% effective at preventing infection and transmission. As vaccination uptake increases, the benefits of testing here, with 70% effectiveness, decline less rapidly than the benefits with 90% effectiveness as shown in Figure 4.5. Differences between the values for 0% vaccine uptake here and those in Figure 4.5 are merely stochastic.

References

L. J. Abu-Raddad, H. Chemaitelly, and A. A. Butt, "Effectiveness of the BNT162b2 Covid-19 vaccine against the B.1.1.7 and B.1.351 variants," *New England Journal of Medicine*, 2021.

D. Adam, P. Wu, J. Wong, E. Lau, T. Tsang, S. Cauchemez, G. Leung, and B. Cowling, "Clustering and superspreading potential of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections in hong kong," 2020.

C. Adami, "The use of information theory in evolutionary biology," *Annals of the New York Academy of Sciences*, vol. 1256, no. 1, pp. 49–65, 2012.

B. Agiiera y Areas, A. L. Fairhall, and W. Bialek, "What can a single neuron compute?" in *Advances in Neural Information Processing Systems 13: Proceedings of the 2000 Conference*, vol. 13. MIT Press, 2001, p. 75.

A. Aleta, D. Martín-Corral, A. Pastore Y Piontti, M. Ajelli, M. Litvinova, M. Chinazzi, N. E. Dean, M. E. Halloran, I. M. Longini, Jr, S. Merler, A. Pentland, A. Vespignani, E. Moro, and Y. Moreno, "Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19," *Nat Hum Behav*, vol. 4, no. 9, pp. 964–971, Sep. 2020.

B. M. Althouse, E. A. Wenger, J. C. Miller, S. V. Scarpino, A. Allard, L. HÃ©bert-Dufresne, and H. Hu, "Stochasticity and heterogeneity in the transmission dynamics of SARS-CoV-2," *arXiv:2005.13689 [physics, q-bio]*, May 2020, arXiv: 2005.13689. [Online]. Available: http://arxiv.org/abs/2005.13689

S. Arora, E. Hazan, and S. Kale, "The Multiplicative Weights Update Method: A Meta-Algorithm and Applications," *THEORY OF COMPUTING*, vol. 8, p. 44, 2012.

R. Assaker, A.-E. Colas, F. Julien-Marsollier, B. Bruneau, L. Marsac, B. Greff, N. Tri, C. Fait, C. Brasher, and S. Dahmani, "Presenting symptoms of COVID-19 in children: a meta-analysis of published studies," *Br. J. Anaesth.*, vol. 125, no. 3, pp. e330–e332, Sep. 2020.

J. A. Backer, D. Klinkenberg, and J. Wallinga, "Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from wuhan, china, 20-28 january 2020," *Euro Surveill.*, vol. 25, no. 5, Feb. 2020.

L. R. Baden, H. M. El Sahly, B. Essink, K. Kotloff, S. Frey, R. Novak, D. Diemert, S. A. Spector, N. Rouphael, C. B. Creech, J. McGettigan, S. Khetan, N. Segall, J. Solis, A. Brosz, C. Fierro, H. Schwartz, K. Neuzil, L. Corey, P. Gilbert, H. Janes, D. Follmann, M. Marovich, J. Mascola, L. Polakowski, J. Ledgerwood, B. S. Graham, H. Bennett, R. Pajon, C. Knightly, B. Leav, W. Deng, H. Zhou, S. Han, M. Ivarsson, J. Miller, and T. Zaks, "Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine," *N Engl J Med*, Dec. 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7787219/

J. Badham and R. Stocker, "The impact of network clustering and assortativity on epidemic behaviour," *Theoretical Population Biology*, vol. 77, no. 1, pp. 71–75, Feb. 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0040580909001270

F. Ball, "Deterministic and stochastic epidemics with several kinds of susceptibles," *Advances in Applied Probability*, vol. 17, no. 1, pp. 1–22, 1985.

S. Bansal, B. T. Grenfell, and L. A. Meyers, "When individual behaviour matters: homogeneous and network models in epidemiology," *Journal of the Royal Society Interface*, vol. 4, no. 16, pp. 879–891, 2007.

V. C. Barclay, T. Smieszek, J. He, G. Cao, J. J. Rainey, H. Gao, A. Uzicanin, and M. SalathÃ©, "Positive Network Assortativity of Influenza Vaccination at a High School: Implications for Outbreak Risk and Herd Immunity," *PLoS One*, vol. 9, no. 2, Feb. 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3914803/

C. T. Bergstrom and M. Lachmann, "Shannon information and biological fitness," in *Information Theory Workshop, 2004. IEEE*. IEEE, 2004, pp. 50–54.

C. T. Bergstrom and M. Rosvall, "The transmission sense of information," *Biology & Philosophy*, vol. 26, no. 2, pp. 159–176, Mar. 2011.

T. Bergstrom, C. T. Bergstrom, and H. Li, "Frequency and accuracy of proactive testing for COVID-19," Sep. 2020.

A. Bershteyn, H.-Y. Kim, J. B. McGillen, and R. S. Braithwaite, "Which policies most effectively reduce SARS-CoV-2 transmission in schools?" *medRxiv*, 2020.

A. M. Bilinski, J. A. Salomon, J. Giardina, A. Ciaranello, and M. Fitzpatrick, "Passing the test: A model-based analysis of safe school-reopening strategies," *medRxiv*, 2021.

M. Birhane, S. Bressler, G. Chang, T. Clark *et al.*, "COVID-19 vaccine breakthrough infections reported to CDC—United States, January 1–April 30, 2021," *Morbidity and Mortality Weekly Report*, vol. 70, no. 21, p. 792, 2021.

J. R. Black, C. Bailey, J. Przewrocka, K. K. Dijkstra, and C. Swanton, "COVID-19: The case for health-care worker screening to prevent hospital transmission," *The Lancet*, vol. 395, no. 10234, pp. 1418–1420, 2020.

A. Borst and F. E. Theunissen, "Information theory and neural coding," *Nature neuroscience*, vol. 2, no. 11, pp. 947–957, 1999.

C. Bracis, E. Burns, M. Moore, D. Swan, D. B. Reeves, J. T. Schiffer, and D. Dimitrov, "Widespread testing, case isolation and contact tracing may allow safe school reopening with continued moderate physical distancing: A modeling analysis of king county, WA data," *Infect Dis Model*, vol. 6, pp. 24–35, 2021.

L. T. Brandal, T. S. Ofitserova, H. Meijerink, R. Rykkvin, H. M. Lund, O. Hungnes, M. Greve-Isdahl, K. Bragstad, K. Nygård, and B. A. Winje, "Minimal transmission of SARS-CoV-2 from paediatric COVID-19 cases in primary schools, norway, august to november 2020," *Euro Surveill.*, vol. 26, no. 1, Jan. 2021.

A. M. Brues, "The Cost of Evolution Vs. the Cost of Not Evolving," *Evolution*, vol. 18, no. 3, pp. 379–383, 1964, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1558-5646.1964.tb01616.x. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1558-5646.1964.tb01616.x

D. Buitrago-Garcia, D. Egli-Gany, M. J. Counotte, S. Hossmann, H. Imeri, A. M. Ipekci, G. Salanti, and N. Low, "Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis," *PLOS Medicine*, vol. 17, no. 9, p. e1003346, Sep. 2020, publisher: Public Library of Science. [Online]. Available: https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1003346

S. Bunyavanich, A. Do, and A. Vicencio, "Nasal gene expression of Angiotensin-Converting enzyme 2 in children and adults," *JAMA*, vol. 323, no. 23, pp. 2427–2429, Jun. 2020.

J. O. Campbell, "Universal Darwinism As a Process of Bayesian Inference," *Front Syst Neurosci*, vol. 10, Jun. 2016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4894882/

R. Carsetti, C. Quintarelli, I. Quinti, E. Piano Mortari, A. Zumla, G. Ippolito, and F. Locatelli, "The immune system of children: the key to understanding SARS-CoV-2 susceptibility?" *Lancet Child Adolesc Health*, vol. 4, no. 6, pp. 414–416, Jun. 2020.

CDC, "Duration of isolation and precautions for adults with COVID-19," https://www.cdc.gov/coronavirus/2019-ncov/hcp/duration-isolation.html, Dec. 2020, accessed: 2020-12-16.

——, "COVID-19 forecasts: Deaths," https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html, Jan. 2021, accessed: 2021-1-7.

——, "Operating schools during COVID-19," https://www.cdc.gov/coronavirus/2019-ncov/community/schools-childcare/schools.html, Dec. 2020, accessed: 2021-1-7.

——, "Options to reduce quarantine for contacts of persons with SARS-CoV-2 infection using symptom monitoring and diagnostic testing," https://www.cdc.gov/coronavirus/2019-ncov/more/scientific-brief-options-to-reduce-quarantine.html, Dec. 2020, accessed: 2020-12-16.

——, "Healthcare workers," https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html, Dec. 2020, accessed: 2021-1-17.

N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games*, 2006. [Online]. Available: http://www.ii.uni.wroc.pl/~lukstafi/pmwiki/uploads/AGT/Prediction_Learning_and_Games.pdf

T. Charmet, L. Schaeffer, R. Grant, S. Galmiche, O. Chény, C. Von Platen, A. Maurizot, A. Rogoff, F. Omar, C. David *et al.*, "Impact of original, B.1.1.7, and B.1.351/P.1 SARS-CoV-2 lineages on vaccine effectiveness of two doses of COVID-19 mRNA vaccines: Results from a nationwide case-control study in france," *The Lancet Regional Health—Europe*, vol. 8, p. 100171, 2021.

E. Chastain, A. Livnat, C. Papadimitriou, and U. Vazirani, "Algorithms, games, and evolution," *Proceedings of the National Academy of Sciences*, vol. 111, no. 29, pp. 10 620–10 623, Jul. 2014. [Online]. Available: http://www.pnas.org/cgi/doi/10.1073/pnas.1406556111

E. Chastain, "The Multiplicative Weight Updates Method for Evolutionary Biology," Ph.D. dissertation, 2017.

H. Chemaitelly, H. M. Yassine, F. M. Benslimane, H. A. Al Khatib, P. Tang, M. R. Hasan, J. A. Malek, P. Coyle, H. H. Ayoub, Z. Al Kanaani *et al.*, "mRNA-1273 COVID-19 vaccine effectiveness against the B.1.1.7 and B.1.351 variants and severe COVID-19 disease in qatar," *Nature Medicine*, pp. 1–8, 2021.

R. Cheong, A. Rhee, C. J. Wang, I. Nemenman, and A. Levchenko, "Information Transduction Capacity of Noisy Biochemical Signaling Networks," *Science*, vol. 334, no. 6054, pp. 354–358, Oct. 2011, publisher: American Association for the Advancement of Science Section: Report.

Color Health, "Testing & vaccines modeling tool," 2021, online at https://www.color.com/testing-and-vaccines-model.

——, "Return to on-site SARS-CoV-2 testing protocols, v. 2.0," 2020, online at https://www.color.com/covid-19-outbreak-model.

T. M. Cover and J. A. Thomas, *Elements of Information Theory*, ser. Wiley series in telecommunications. USA: Wiley-Interscience, 2006.

J. F. Crow, "Genetic Loads and the Cost of Natural Selection," in *Mathematical Topics in Population Genetics*, ser. Biomathematics, K.-i. Kojima, Ed. Berlin, Heidelberg: Springer, 1970, pp. 128–177. [Online]. Available: https://doi.org/10.1007/978-3-642-46244-3_5

——, "Some possibilities for measuring selection intensities in man," *Human Biology*, vol. 30, no. 1, pp. 1–13, 1958, publisher: Wayne State University Press. [Online]. Available: http://www.jstor.org/stable/41449168

D. Czégel, H. Giaffar, I. Zachar, J. B. Tenenbaum, and E. Szathmáry, "Evolutionary implementation of Bayesian computations," *bioRxiv*, p. 685842, Jun. 2020, publisher: Cold Spring Harbor Laboratory Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.1101/685842v2

N. Dagan, N. Barda, E. Kepten, O. Miron, S. Perchik, M. A. Katz, M. A. HernÃ¡n, M. Lipsitch, B. Reis, and R. D. Balicer, "BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass Vaccination Setting," *New England Journal of Medicine*, vol. 384, no. 15, pp. 1412–1423, Apr. 2021, publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMoa2101765. [Online]. Available: https://doi.org/10.1056/NEJMoa2101765

L. Danon, T. A. House, J. M. Read, and M. J. Keeling, "Social encounter networks: collective properties and disease transmission," *Journal of The Royal Society Interface*, vol. 9, no. 76, pp. 2826–2833, Nov. 2012, publisher: Royal Society. [Online]. Available: https://royalsocietypublishing.org/doi/full/10.1098/rsif.2012.0357

N. G. Davies, S. Abbott, R. C. Barnard, C. I. Jarvis, A. J. Kucharski, J. D. Munday, C. A. Pearson, T. W. Russell, D. C. Tully, A. D. Washburne *et al.*, "Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in england," *Science*, 2021.

T. N. Denny, L. Andrews, M. Bonsignori, K. Cavanaugh, M. B. Datto, A. Deckard, C. T. DeMarco, N. DeNaeyer, C. A. Epling, T. Gurley, S. B. Haase, C. Hallberg, J. Harer, C. L. Kneifel, M. J. Lee, R. Louzao, M. A. Moody, Z. Moore, C. R. Polage, J. Puglin, P. H. Spotts, J. A. Vaughn, and C. R. Wolfe, "Implementation of a pooled surveillance testing program for asymptomatic SARS-CoV-2 infections on a college campus - duke university, durham, north carolina, august 2-october 11, 2020," *MMWR Morb. Mortal. Wkly. Rep.*, vol. 69, no. 46, pp. 1743–1747, Nov. 2020.

M. C. Donaldson-Matasci, C. T. Bergstrom, and M. Lachmann, "The fitness value of information," *Oikos*, vol. 119, no. 2, pp. 219–230, Feb. 2010.

J. O. Dubuis, G. Tkacik, E. F. Wieschaus, T. Gregor, and W. Bialek, "Positional information, in bits," *Proceedings of the National Academy of Sciences*, vol. 110, no. 41, pp. 16 301–16 308, Oct. 2013.

A. Endo, S. Abbott, A. Kucharski, and S. Funk, "Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside china. wellcome open research, 5 (67)," 2020.

J. Fagnan, A. Abnar, R. Rabbany, and O. R. Zaiane, "Modular Networks for Validating Community Detection Algorithms," *arXiv:1801.01229 [physics]*, Jan. 2018, arXiv: 1801.01229. [Online]. Available: http://arxiv.org/abs/1801.01229

L. J. Faherty, B. K. Master, E. D. Steiner, J. H. Kaufman, Z. Predmore, L. Stelitano, J. T. Leschitz, B. Phillips, H. L. Schwartz, and R. Wolfe, "COVID-19 Testing in Kâ12 Schools: Insights from Early Adopters," Mar. 2021, publisher: RAND Corporation. [Online]. Available: https://www.rand.org/pubs/research_reports/RRA1103-1.html

J. Felsenstein, "On the Biological Significance of the Cost of Gene Substitution.pdf," 1971.

Z. Feng, D. Xu, and H. Zhao, "Epidemiological models with non-exponentially distributed disease stages and applications to disease control," *Bull. Math. Biol.*, vol. 69, no. 5, pp. 1511–1536, Jul. 2007.

N. M. Ferguson, *Report 9: Impact of Non-pharmaceutical Interventions (NPIs) to Reduce COVID19 Mortality and Healthcare Demand*. Imperial College London, 2020.

L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-DÃ¶rner, M. Parker, D. Bonsall, and C. Fraser, "Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing," *Science*, vol. 368, no. 6491, May 2020, publisher: American Association for the Advancement of Science Section: Research Article. [Online]. Available: http://science.sciencemag.org/content/368/6491/eabb6936

R. A. Fisher, *The genetical theory of natural selection*, ser. The genetical theory of natural selection. Oxford, England: Clarendon Press, 1930, pages: xiv, 272.

S. A. Frank, "Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory," *Journal of Evolutionary Biology*, vol. 25, no. 12, pp. 2377–2396, Dec. 2012.

Y. Freund and R. E. Schapire, "Adaptive Game Playing Using Multiplicative Weights," *Games and Economic Behavior*, vol. 29, no. 1-2, pp. 79–103, Oct. 1999. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0899825699907388

S. E. Galloway, "Emergence of SARS-CoV-2 b.1.1.7 lineage — united states, december 29, 2020–january 12, 2021," *MMWR Morb. Mortal. Wkly. Rep.*, vol. 70, 2021.

T. Ganyani, C. Kremer, D. Chen, A. Torneri, C. Faes, J. Wallinga, and N. Hens, "Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, march 2020," *Euro Surveill.*, vol. 25, no. 17, Apr. 2020.

J. H. Gillespie, "Natural selection for within-generation variance in offspring number," *Genetics*, vol. 76, no. 3, pp. 601–606, Mar. 1974. [Online]. Available: https://doi.org/10.1093/genetics/76.3.601

e. a. Gillespie DL, "The experience of two independent schools with in-person learning during the covid-19 pandemic," *In press, Journal of School Health*, 2021.

G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Di Filippo, A. Di Matteo, and M. Colaneri, "Modelling the COVID-19 epidemic and implementation of population-wide interventions in italy," *Nat. Med.*, vol. 26, no. 6, pp. 855–860, Jun. 2020.

P. Godfrey-Smith, "Information, Arbitrariness, and Selection: Comments on Maynard Smith," *Philosophy of Science*, vol. 67, no. 2, pp. 202–207, 2000, publisher: [The University of Chicago Press, Philosophy of Science Association].

——, "Information in biology," *The Cambridge companion to the philosophy of biology*, 2007.

——, "On the Theoretical Role of "Genetic Coding"," *Philosophy of Science*, vol. 67, no. 1, pp. 26–44, 2000, publisher: [The University of Chicago Press, Philosophy of Science Association].

E. Goldstein, M. Lipsitch, and M. Cevik, "On the effect of age on the transmission of SARS-CoV-2 in households, schools and the community," *J. Infect. Dis.*, Oct. 2020.

P. E. Griffiths, *The Fearless Vampire Conservator: Philip Kitcher, Genetic Determinism, and the Informational Gene*. Duke University Press, Jan. 2006, pages: 175-198 Publication Title: Genes in Development Section: Genes in Development.

——, "Genetic Information: A Metaphor in Search of a Theory," *Philosophy of Science*, vol. 68, no. 3, pp. 394–412, 2001, publisher: [The University of Chicago Press, Philosophy of Science Association].

W.-J. Guan, Z.-Y. Ni, Y. Hu, W.-H. Liang, C.-Q. Ou, J.-X. He, L. Liu, H. Shan, C.-L. Lei, D. S. C. Hui, B. Du, L.-J. Li, G. Zeng, K.-Y. Yuen, R.-C. Chen, C.-L. Tang, T. Wang, P.-Y. Chen, J. Xiang, S.-Y. Li, J.-L. Wang, Z.-J. Liang, Y.-X. Peng, L. Wei, Y. Liu, Y.-H. Hu, P. Peng, J.-M. Wang, J.-Y. Liu, Z. Chen, G. Li, Z.-J. Zheng, S.-Q. Qiu, J. Luo, C.-J. Ye, S.-Y. Zhu, N.-S. Zhong, and China Medical Treatment Expert Group for Covid-19, "Clinical characteristics of coronavirus disease 2019 in china," *N. Engl. J. Med.*, vol. 382, no. 18, pp. 1708–1720, Apr. 2020.

E. J. Haas, F. J. Angulo, J. M. McLaughlin, E. Anis, S. R. Singer, F. Khan, N. Brooks, M. Smaja, G. Mircus, K. Pan, J. Southern, D. L. Swerdlow, L. Jodar, Y. Levy, and S. Alroy-Preis, "Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data," *The Lancet*, vol. 0, no. 0, May 2021, publisher: Elsevier. [Online]. Available: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(21)00947-8/abstract

J. B. S. Haldane, "The Effect of Variation of Fitness," *The American Naturalist*, vol. 71, no. 735, pp. 337–349, Jul. 1937, publisher: The University of Chicago Press. [Online]. Available: http://www.journals.uchicago.edu/doi/abs/10.1086/280722

J. Haldane, "The Cost of Natural Selection," *Journal of Genetics*, vol. 55, pp. 511–524, 1957.

E. A. Hanushek and L. Woessmann, "The economic impacts of learning losses," 2020.

M. Harper, "Information Geometry and Evolutionary Game Theory," *arXiv:0911.1383 [cs, math, nlin]*, Nov. 2009, arXiv: 0911.1383. [Online]. Available: http://arxiv.org/abs/0911.1383

——, "The Replicator Equation as an Inference Dynamic," *arXiv:0911.1763 [cs, math]*, Nov. 2009, arXiv: 0911.1763.

M. Harper and D. E. A. Fryer, "Stability of Evolutionary Dynamics on Time Scales," *Dyn Games Appl*, vol. 5, no. 3, pp. 318–333, Sep. 2015, arXiv: 1210.5539. [Online]. Available: http://arxiv.org/abs/1210.5539

X. He, E. H. Y. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan, X. Mo, Y. Chen, B. Liao, W. Chen, F. Hu, Q. Zhang, M. Zhong, Y. Wu, L. Zhao, F. Zhang, B. J. Cowling, F. Li, and G. M. Leung, "Temporal dynamics in viral shedding and transmissibility of COVID-19," *Nat. Med.*, vol. 26, no. 5, pp. 672–675, May 2020.

X. He, E. H. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan *et al.*, "Temporal dynamics in viral shedding and transmissibility of COVID-19," *Nature Medicine*, vol. 26, no. 5, pp. 672–675, 2020.

L. Heavey, G. Casey, C. Kelly, D. Kelly, and G. McDarby, "No evidence of secondary transmission of COVID-19 from children attending school in ireland, 2020," *Euro Surveill.*, vol. 25, no. 21, May 2020.

R. Hickson and M. Roberts, "How population heterogeneity in susceptibility and infectivity influences epidemic dynamics," *Journal of Theoretical Biology*, vol. 350, pp. 70–80, 2014.

M. Hilbert, "The More You Know, the More You Can Grow: An Information Theoretic Approach to Growth in the Information Age," *Entropy*, vol. 19, no. 2, p. 82, Feb. 2017, number: 2 Publisher: Multidisciplinary Digital Publishing Institute.

M. Hoffmann, H. Hofmann-Winkler, N. KrÃŒger, A. Kempf, I. Nehlmeier, L. Graichen, A. Sidarovich, A.-S. Moldenhauer, M. S. Winkler, S. Schulz, H.-M. JÃ€ck, M. V. Stankov, G. M. N. Behrens, and S. PÃ¶hlmann, "SARS-CoV-2 variant B.1.617 is resistant to Bamlanivimab and evades antibodies induced by infection and vaccination," *bioRxiv*, p. 2021.05.04.442663, May

2021, publisher: Cold Spring Harbor Laboratory Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.1101/2021.05.04.442663v1

K. E. Holsinger and B. S. Weir, "Genetics in geographically structured populations: defining, estimating and interpreting FST," *Nat Rev Genet*, vol. 10, no. 9, pp. 639–650, Sep. 2009.

J. H. Hurst, S. M. Heston, H. N. Chambers, H. M. Cunningham, M. J. Price, L. Suarez, C. G. Crew, S. Bose, J. N. Aquino, S. T. Carr, S. M. Griffin, S. H. Smith, K. Jenkins, T. S. Pfeiffer, J. Rodriguez, C. T. DeMarco, N. A. De Naeyer, T. C. Gurley, R. Louzao, C. Zhao, C. K. Cunningham, W. J. Steinbach, T. N. Denny, D. J. Lugo, M. A. Moody, S. R. Permar, A. T. Rotta, N. A. Turner, E. B. Walter, C. W. Woods, and M. S. Kelly, "SARS-CoV-2 Infections Among Children in the Biospecimens from Respiratory Virus-Exposed Kids (BRAVE Kids) Study," *Clin Infect Dis*, Nov. 2020.

L. HÃ©bert-Dufresne, B. M. Althouse, S. V. Scarpino, and A. Allard, "Beyond R_0: Heterogeneity in secondary infections and probabilistic epidemic forecasting," *arXiv:2002.04004 [physics, q-bio]*, Apr. 2020, arXiv: 2002.04004. [Online]. Available: http://arxiv.org/abs/2002.04004

S. A. Ismail, V. Saliba, J. L. Bernal, M. E. Ramsay, and S. Ladhani, "SARS-CoV-2 infection and transmission in educational settings: Cross-Sectional analysis of clusters and outbreaks in england," Aug. 2020.

S. A. Ismail, V. Saliba, J. L. Bernal, M. E. Ramsay, and S. N. Ladhani, "SARS-CoV-2 infection and transmission in educational settings: a prospective, cross-sectional analysis of infection clusters and outbreaks in england," 2020.

E. O. i. Kampe, A.-S. Lehfeld, S. Buda, U. Buchholz, and W. Haas, "Surveillance of COVID-19 school outbreaks, germany, march to august 2020," *Eurosurveillance*, vol. 25, no. 38, p. 2001645, Sep. 2020.

O. Karin, Y. M. Bar-On, T. Milo, I. Katzir, A. Mayo, Y. Korem, B. Dudovich, E. Yashiv, A. J. Zehavi, N. Davidovitch, R. Milo, and U. Alon, "Cyclic exit strategies to suppress COVID-19 and allow economic activity," 2020.

O. Karin, Y. M. Bar-On, T. Milo, I. Katzir, A. Mayo, Y. Korem, B. Dudovich, E. Yashiv, A. J. Zehavi, N. Davidovitch *et al.*, "Cyclic strategies to suppress COVID-19 and allow economic activity," 2020.

A. Kaznatcheev, "Computational Complexity as an Ultimate Constraint on Evolution," *Genetics*, vol. 212, no. 1, pp. 245–265, May 2019. [Online]. Available: https://doi.org/10.1534/genetics.119.302000

——, "Evolution is exponentially more powerful with frequency-dependent selection," *bioRxiv*, p. 2020.05.03.075069, May 2020, publisher: Cold Spring Harbor Laboratory Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.1101/2020.05.03.075069v1

——, "Two conceptions of evolutionary games: reductive vs effective," *bioRxiv*, p. 231993, Dec. 2017, publisher: Cold Spring Harbor Laboratory Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.1101/231993v1

M. Keeling, "The implications of network structure for epidemic dynamics," *Theoretical Population Biology*, vol. 67, no. 1, pp. 1–8, 2005.

M. J. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, Sep. 2011.

A. Kimball, K. M. Hatfield, M. Arons, A. James, J. Taylor, K. Spicer, A. C. Bardossy, L. P. Oakley, S. Tanwar, Z. Chisty *et al.*, "Asymptomatic and presymptomatic SARS-CoV-2 infections in residents of a long-term care skilled nursing facility—King County, Washington, March 2020," *Morbidity and Mortality Weekly Report*, vol. 69, no. 13, p. 377, 2020.

M. Kimura, "Evolutionary Rate at the Molecular Level," *Nature*, vol. 217, no. 5129, pp. 624–626, Feb. 1968, number: 5129 Publisher: Nature Publishing Group. [Online]. Available: http://www.nature.com/articles/217624a0

——, "Natural selection as the process of accumulating genetic information in adaptive evolution," 1961.

——, "Natural selection as the process of accumulating genetic information in adaptive evolution," *Genetical Research*, vol. 2, no. 1, pp. 127–140, Feb. 1961. [Online]. Available: https://www.cambridge.org/core/product/identifier/S0016672300000616/type/journal_article

——, *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press, 1983. [Online]. Available: https://www.cambridge.org/core/books/neutral-theory-of-molecular-evolution/0FF60E9F47915B17FFA2620C49400632

——, "Optimum mutation rate and degree of dominance as determined by the principle of minimum genetic load," *Journal of Genetics*, vol. 57, no. 1, pp. 21–34, Jun. 1960. [Online]. Available: http://link.springer.com/10.1007/BF02985336

P. Kitcher, "Battling the undead: How (and how not) to resist genetic determinism," in *Thinking about evolution: Historical, philosophical and political perspectives*, R. Singh, C. Krimbas, J. Beatty, and D. Paul, Eds., 2001, pp. 396–414.

J. Kivinen and M. K. Warmuth, "Additive versus exponentiated gradient updates for linear prediction," in *Proceedings of the twenty-seventh annual ACM symposium on Theory of computing*, ser. STOC '95. New York, NY, USA: Association for Computing Machinery, May 1995, pp. 209–218. [Online]. Available: https://doi.org/10.1145/225058.225121

A. Kolchinsky and D. H. Wolpert, "Semantic information, autonomous agency and non-equilibrium statistical physics," *Interface Focus.*, vol. 8, no. 6, p. 20180041, Dec. 2018.

B. Korber, W. M. Fischer, S. Gnanakaran, H. Yoon, J. Theiler, W. Abfalterer, N. Hengartner, E. E. Giorgi, T. Bhattacharya, B. Foley, K. M. Hastie, M. D. Parker, D. G. Partridge, C. M. Evans, T. M. Freeman, T. I. de Silva, Sheffield COVID-19 Genomics Group, C. McDanal, L. G. Perez, H. Tang, A. Moon-Walker, S. P. Whelan, C. C. LaBranche, E. O. Saphire, and D. C. Montefiori, "Tracking changes in SARS-CoV-2 spike: Evidence that D614G increases infectivity of the COVID-19 virus," *Cell*, vol. 182, no. 4, pp. 812–827.e19, Aug. 2020.

O. Krylova and D. J. D. Earn, "Effects of the infectious period distribution on predicted transitions in childhood disease dynamics," p. 20130098, 2013.

A. J. Kucharski, T. W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, R. M. Eggo, and Centre for Mathematical Modelling of Infectious Diseases COVID-19 working group, "Early dynamics of transmission and control of COVID-19: a mathematical modelling study," *Lancet Infect. Dis.*, vol. 20, no. 5, pp. 553–558, May 2020.

D. B. Larremore, B. Wilder, E. Lester, S. Shehata, J. M. Burke, J. A. Hay, M. Tambe, M. J. Mina, and R. Parker, "Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening," *Sci Adv*, Nov. 2020.

——, "Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening," *Science Advances*, vol. 7, no. 1, p. eabd5393, 2021.

S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler, "The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported

confirmed cases: Estimation and application," *Ann. Intern. Med.*, vol. 172, no. 9, pp. 577–582, May 2020.

R. T. Leeb, S. Price, S. Sliwa, A. Kimball, L. Szucs, E. Caruso, S. Godfred-Cato, and M. Lozier, "COVID-19 trends among School-Aged children - united states, march 1-september 19, 2020," *MMWR Morb. Mortal. Wkly. Rep.*, vol. 69, no. 39, pp. 1410–1415, Oct. 2020.

E. Leidman, "COVID-19 trends among persons aged 0–24 years — united states, march 1–december 12, 2020," *MMWR Morb. Mortal. Wkly. Rep.*, vol. 70, 2021.

K. Leung, M. H. Shum, G. M. Leung, T. T. Lam, and J. T. Wu, "Early transmissibility assessment of the N501Y mutant strains of SARS-CoV-2 in the united kingdom, october to november 2020," *Euro Surveill.*, vol. 26, no. 1, Jan. 2021.

A. Levchenko and I. Nemenman, "Cellular noise and information transmission," *Current Opinion in Biotechnology*, vol. 28, pp. 156–164, Aug. 2014.

M. Levine-Tiefenbrun, I. Yelin, H. Uriel, J. Kuint, L. Schreiber, E. Herzel, R. Katz, A. Ben-Tov, T. Patalon, G. Chodick *et al.*, "Association of covid-19 rt-qpcr test false-negative rate with patient age, sex and time since diagnosis," *medRxiv*, 2020.

M. Levine-Tiefenbrun, I. Yelin, R. Katz, E. Herzel, Z. Golan, L. Schreiber, T. Wolf, V. Nadler, A. Ben-Tov, J. Kuint, S. Gazit, T. Patalon, G. Chodick, and R. Kishony, "Initial report of decreased SARS-CoV-2 viral load after inoculation with the BNT162b2 vaccine," *Nature Medicine*, pp. 1–3, Mar. 2021, publisher: Nature Publishing Group. [Online]. Available: http://www.nature.com/articles/s41591-021-01316-7

Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. M. Leung, E. H. Y. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T. Y. Lam, J. T. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung, and Z. Feng, "Early transmission dynamics in wuhan, china, of novel Coronavirus-Infected pneumonia," *N. Engl. J. Med.*, vol. 382, no. 13, pp. 1199–1207, Mar. 2020.

H. A. Lindsey, J. Gallie, S. Taylor, and B. Kerr, "Evolutionary rescue from extinction is contingent on a lower rate of environmental change," *Nature*, vol. 494, no. 7438, p. 463, Feb. 2013. [Online]. Available: https://www.nature.com/articles/nature11879

Y. Liu, A. A. Gayle, A. Wilder-Smith, and J. Rocklöv, "The reproductive number of COVID-19 is higher compared to SARS coronavirus," *J. Travel Med.*, vol. 27, no. 2, Mar. 2020.

J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, and W. M. Getz, "Superspreading and the effect of individual variation on disease emergence," *Nature*, vol. 438, no. 7066, pp. 355–359, Nov. 2005, number: 7066 Publisher: Nature Publishing Group. [Online]. Available: http://www.nature.com/articles/nature04153

J. Lopez Bernal, N. Andrews, C. Gower, E. Gallagher, R. Simmons, S. Thelwall, J. Stowe, E. Tessier, N. Groves, G. Dabrera *et al.*, "Effectiveness of covid-19 vaccines against the b. 1.617. 2 (delta) variant," *New England Journal of Medicine*, 2021.

K. Macartney, H. E. Quinn, A. J. Pillsbury, A. Koirala, L. Deng, N. Winkler, A. L. Katelaris, M. V. N. O'Sullivan, C. Dalton, N. Wood, and NSW COVID-19 Schools Study Team, "Transmission of SARS-CoV-2 in australian educational settings: a prospective cohort study," *Lancet Child Adolesc Health*, vol. 4, no. 11, pp. 807–816, Nov. 2020.

D. M. MacKay and W. S. McCulloch, "The limiting information capacity of a neuronal link," *The bulletin of mathematical biophysics*, vol. 14, no. 2, pp. 127–135, 1952.

J. Maynard Smith, "The Concept of Information in Biology," *Philosophy of Science*, vol. 67, no. 2, pp. 177–194, Jun. 2000, publisher: The University of Chicago Press.

——, *Evolution and the Theory of Games*. Cambridge University Press, Oct. 1982, google-Books-ID: Nag2IhmPS3gC.

——, ""Haldane's Dilemma" and the Rate of Evolution," *Nature*, vol. 219, 1968. [Online]. Available: http://www.nature.com.offcampus.lib.washington.edu/nature/journal/v219/n5159/pdf/2191114a0.pdf

R. McGee, "SEIRS+ model framework," Aug. 2020, online at https://github.com/ryansmcgee/seirsplus.

R. S. McGee, J. R. Homburger, H. E. Williams, C. T. Bergstrom, and A. Y. Zhou, "Model-driven mitigation measures for reopening schools during the COVID-19 pandemic," *Proceedings of the National Academy of Sciences USA*, 2021, in Press.

R. Mehta, I. Panageas, and G. Piliouras, "Natural Selection As an Inhibitor of Genetic Diversity: Multiplicative Weights Updates Algorithm and a Conjecture of Haploid Genetics [Working Paper Abstract]," in *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science*, ser. ITCS '15. New York, NY, USA: ACM, 2015, pp. 73–73, event-place: Rehovot, Israel. [Online]. Available: http://doi.acm.org/10.1145/2688073.2688118

R. Meir and D. Parkes, "On Sex, Evolution, and the Multiplicative Weights Update Algorithm," Feb. 2015. [Online]. Available: http://arxiv.org/abs/1502.05056v1

S. M. Moghadas, M. C. Fitzpatrick, P. Sah, A. Pandey, A. Shoukat, B. H. Singer, and A. P. Galvani, "The implications of silent transmission for the control of COVID-19 outbreaks," *PNAS*, vol. 117, no. 30, pp. 17 513–17 515, Jul. 2020, publisher: National Academy of Sciences Section: Biological Sciences. [Online]. Available: https://www.pnas.org/content/117/30/17513

P. a. P. Moran, "â□□Haldane's dilemmaâ□□ and the rate of evolution," *Annals of Human Genetics*, vol. 33, no. 3, pp. 245–249, 1970, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1469-1809.1970.tb01649.x. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-1809.1970.tb01649.x

J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, and W. J. Edmunds, "Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases," *PLOS Medicine*, vol. 5, no. 3, p. e74, Mar. 2008, publisher: Public Library of Science. [Online]. Available: https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0050074

F. C. Motta, K. A. McGoff, A. Deckard, C. R. Wolfe, M. A. Moody, K. Cavanaugh, T. N. Denny, J. Harer, and S. B. Haase, "Benefits of surveillance testing and quarantine in a SARS-CoV-2 vaccinated population of students on a university campus," *medRxiv*, 2021.

M. E. J. Newman and J. Park, "Why social networks are different from other types of networks," *Physical Review E*, vol. 68, no. 3, Sep. 2003. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.68.036122

A. Nierenberg and A. Pasick, "Will any more schools reopen in 2020?" *The New York Times*, Nov. 2020.

J. E. Niven and S. B. Laughlin, "Energy limitation as a selective pressure on the evolution of sensory systems," *Journal of Experimental Biology*, vol. 211, no. 11, pp. 1792–1804, Jun. 2008.

P. O'Donald, "â□□Haldane's Dilemmaâ□□ and the Rate of Natural Selection," *Nature*, vol. 221, no. 5183, pp. 815–816, Mar. 1969, number: 5183 Publisher: Nature Publishing Group. [Online]. Available: http://www.nature.com/articles/221815a0

S. E. Oliver, J. W. Gargano, H. Scobie, M. Wallace, S. C. Hadler, J. Leung, A. E. Blain, N. McClung, D. Campos-Outcalt, R. L. Morgan *et al.*, "The advisory committee on immunization practices' interim recommendation for use of Janssen COVID-19 vaccine — United States, February 2021," *Morbidity and Mortality Weekly Report*, vol. 70, no. 9, p. 329, 2021.

D. P. Oran and E. J. Topol, "Prevalence of Asymptomatic SARS-CoV-2 Infection : A Narrative Review," *Ann Intern Med*, vol. 173, no. 5, pp. 362–367, Sep. 2020.

S. E. Palmer, O. Marre, M. J. Berry, and W. Bialek, "Predictive information in a sensory population," *PNAS*, vol. 112, no. 22, pp. 6908–6913, Jun. 2015, publisher: National Academy of Sciences Section: Physical Sciences.

J. Panovska-Griffiths, C. C. Kerr, R. M. Stuart, D. Mistry, D. J. Klein, R. M. Viner, and C. Bonell, "Determining the optimal strategy for reopening schools, the impact of test and trace interventions, and the risk of occurrence of a second COVID-19 epidemic wave in the UK: a modelling study," *Lancet Child Adolesc Health*, vol. 4, no. 11, pp. 817–827, Nov. 2020.

R. Patil, R. Dave, H. Patel, V. M. Shah, D. Chakrabarti, and U. Bhatia, "Assessing the interplay between travel patterns and SARS-CoV-2 outbreak in realistic urban setting," *Applied network science*, vol. 6, no. 1, pp. 1–19, 2021.

C. M. Peak, R. Kahn, Y. H. Grad, L. M. Childs, R. Li, M. Lipsitch, and C. O. Buckee, "Individual quarantine versus active monitoring of contacts for the mitigation of COVID-19: a modelling study," *Lancet Infect. Dis.*, vol. 20, no. 9, pp. 1025–1033, Sep. 2020.

E. Petter, O. Mor, N. Zuckerman, D. Oz-Levi, A. Younger, D. Aran, and Y. Erlich, "Initial real world evidence for lower viral load of individuals who have been vaccinated by BNT162b2," Epidemiology, preprint, Feb. 2021. [Online]. Available: http://medrxiv.org/lookup/doi/10.1101/2021.02.08.21251329

F. P. Polack, S. J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J. L. Perez, G. PÃ©rez Marc, E. D. Moreira, C. Zerbini, R. Bailey, K. A. Swanson, S. Roychoudhury, K. Koury, P. Li, W. V. Kalina, D. Cooper, R. W. Frenck, L. L. Hammitt, Ö. TÃœreci, H. Nell, A. Schaefer, S. Ã□nal, D. B. Tresnan, S. Mather, P. R. Dormitzer, U. Å□ahin, K. U. Jansen, and W. C. Gruber, "Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine," *New England Journal of Medicine*, vol. 383, no. 27, pp. 2603–2615, Dec. 2020, publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMoa2034577. [Online]. Available: https://doi.org/10.1056/NEJMoa2034577

J. Poline, J. Gaschignard, C. Leblanc, F. Madhi, E. Foucaud, E. Nattes, A. Faye, S. Bonacorsi, P. Mariani, E. Varon, M. Smati-Lafarge, M. Caseris, R. Basmaci, N. Lachaume, and N. Ouldali, "Systematic SARS-CoV-2 screening at hospital admission in children:a French prospective multicenter study," *Clin Infect Dis*, Jul. 2020.

G. E. Potter, T. Smieszek, and K. Sailer, "Modeling workplace contact networks: The effects of organizational structure, architecture, and reporting errors on epidemic predictions," *Network Science*, vol. 3, no. 3, pp. 298–325, 2015.

Public Health England, "SARS-CoV-2 variants of concern and variants under investigation in england," *Technical Briefing 12*, 2021.

L. Rajmil, "Role of children in the transmission of the COVID-19 pandemic: a rapid scoping review," *BMJ Paediatr Open*, vol. 4, no. 1, p. e000722, Jun. 2020.

J. M. Read, K. T. Eames, and W. J. Edmunds, "Dynamic social networks and the implications for the spread of infectious disease," *Journal of The Royal Society Interface*, vol. 5, no. 26, pp. 1001–1007, Sep. 2008. [Online]. Available: https://royalsocietypublishing.org/doi/10.1098/rsif.2008.0013

J. M. Read, J. R. E. Bridgen, D. A. T. Cummings, A. Ho, and C. P. Jewell, "Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions," Jan. 2020.

A. Rich, "On the problems of evolution and biochemical information transfer," *undefined*, 1962.

L. Rivett, S. Sridhar, D. Sparkes, M. Routledge, N. K. Jones, S. Forrest, J. Young, J. Pereira-Dias, W. L. Hamilton, M. Ferris *et al.*, "Screening of healthcare workers for SARS-CoV-2 highlights the role of asymptomatic carriage in COVID-19 transmission," *eLife*, vol. 9, p. e58728, 2020.

O. Rivoire, "Informations in Models of Evolutionary Dynamics," *Journal of Statistical Physics*, vol. 162, no. 5, pp. 1324–1352, Mar. 2016.

O. Rivoire and S. Leibler, "The Value of Information for Populations in Varying Environments," *Journal of Statistical Physics*, vol. 142, no. 6, pp. 1124–1166, Apr. 2011.

C. Ross, O. Spector, M. A. Tsadok, Y. Weiss, R. Barnea *et al.*, "BNT162b2 mRNA vaccinations in israel: understanding the impact and improving the vaccination policies by redefining the immunized population," *medRxiv*, 2021.

M. Salathe, M. Kazandjieva, J. W. Lee, P. Levis, M. W. Feldman, and J. H. Jones, "A high-resolution human contact network for infectious disease transmission," *Proceedings of the National Academy of Sciences*, vol. 107, no. 51, pp. 22 020–22 025, Dec. 2010. [Online]. Available: http://www.pnas.org/cgi/doi/10.1073/pnas.1009094108

A. A. Sayampanathan, C. S. Heng, P. H. Pin, J. Pang, T. Y. Leong, and V. J. Lee, "Infectivity of asymptomatic versus symptomatic COVID-19," *Lancet*, vol. 397, no. 10269, pp. 93–94, Jan. 2021.

C. R. Shalizi, "Dynamics of Bayesian updating with dependent data and misspecified models," *Electronic Journal of Statistics*, vol. 3, no. none, pp. 1039–1074, Jan. 2009, publisher: Institute of Mathematical Statistics and Bernoulli Society. [Online]. Available: https://projecteuclid.org/journals/electronic-journal-of-statistics/volume-3/issue-none/Dynamics-of-Bayesian-updating-with-dependent-data-and-misspecified-models/10.1214/09-EJS485.full

C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948, conference Name: The Bell System Technical Journal.

N. Shea, "Consumers Need Information: Supplementing Teleosemantics with an Input Condition," *Philosophy and Phenomenological Research*, vol. 75, no. 2, pp. 404–435, 2007, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1933-1592.2007.00082.x.

——, "Representation in the genome and in other inheritance systems," *Biol Philos*, vol. 22, no. 3, pp. 313–331, Jun. 2007.

R. L. Smith, L. L. Gibson, P. P. Martinez, R. Ke, A. Mirza, M. Conte, N. Gallagher, A. Conte, L. Wang, R. Fredrickson, D. C. Edmonson, M. E. Baughman, K. K. Chiu, H. Choi, T. W. Jensen, K. R. Scardina, S. Bradley, S. L. Gloss, C. Reinhart, J. Yedetore, A. N. Owens, J. Broach, B. Barton, P. Lazar, D. Henness, T. Young, A. Dunnett, M. L. Robinson, H. H. Mostafa, A. Pekosz, Y. C. Manabe, W. J. Heetderks, D. D. McManus, and C. B. Brooke, "Longitudinal assessment of diagnostic test performance over the course of acute SARS-CoV-2 infection," Infectious Diseases (except HIV/AIDS), preprint, Mar. 2021.

[Online]. Available: http://medrxiv.org/lookup/doi/10.1101/2021.03.19.21253964

C. Stein-Zamir, N. Abramson, H. Shoob, E. Libal, M. Bitan, T. Cardash, R. Cayam, and I. Miskin, "A large COVID-19 outbreak in a high school 10 days after schools' reopening, israel, may 2020," *Eurosurveillance*, vol. 25, no. 29, p. 2001352, Jul. 2020.

K. Sterelny, "The "Genetic Program" Program: A Commentary on Maynard Smith on Information in Biology," *Philosophy of Science*, vol. 67, no. 2, pp. 195–201, Jun. 2000, publisher: The University of Chicago Press.

S. P. Strong, R. Koberle, R. R. de Ruyter van Steveninck, and W. Bialek, "Entropy and Information in Neural Spike Trains," *Phys. Rev. Lett.*, vol. 80, no. 1, pp. 197–200, Jan. 1998, publisher: American Physical Society.

J. A. Sved, "Possible Rates of Gene Substitution in Evolution," *The American Naturalist*, vol. 102, no. 925, pp. 283–293, 1968, publisher: [University of Chicago Press, American Society of Naturalists]. [Online]. Available: https://www.jstor.org/stable/2459028

M. D. Swift, L. E. Breeher, A. J. Tande, C. P. Tommaso, C. M. Hainy, H. Chu, M. H. Murad, E. F. Berbari, and A. Virk, "Effectiveness of mRNA COVID-19 vaccines against SARS-CoV-2 infection in a cohort of healthcare personnel," *Clinical Infectious Diseases*, no. ciab361, Apr. 2021. [Online]. Available: https://doi.org/10.1093/cid/ciab361

M. G. Thompson, "Interim Estimates of Vaccine Effectiveness of BNT162b2 and mRNA-1273 COVID-19 Vaccines in Preventing SARS-CoV-2 Infection Among Health Care Personnel, First Responders, and Other Essential and Frontline Workers â Eight U.S. Locations, December 2020âMarch 2021," *MMWR Morb Mortal Wkly Rep*, vol. 70, 2021. [Online]. Available: https://www.cdc.gov/mmwr/volumes/70/wr/mm7013e3.htm

M. G. Thompson, J. L. Burgess, A. L. Naleway, H. L. Tyner, S. K. Yoon, J. Meece, L. E. Olsho, A. J. Caban-Martinez, A. Fowlkes, K. Lutrick *et al.*, "Interim estimates of vaccine effectiveness of BNT162b2 and mRNA-1273 COVID-19 vaccines in preventing SARS-CoV-2 infection among health care personnel, first responders, and other essential and frontline workers — eight US locations, December 2020–March 2021," *Morbidity and Mortality Weekly Report*, vol. 70, no. 13, p. 495, 2021.

M. Tikhonov and W. Bialek, "Complexity in genetic networks: topology vs. strength of interactions," *arXiv:1308.0317 [q-bio]*, Aug. 2013, arXiv: 1308.0317.

L. Tindale, M. Coombe, J. E. Stockdale, E. Garlock, W. Y. V. Lau, M. Saraswat, Y.-H. B. Lee, L. Zhang, D. Chen, J. Wallinga, and C. Colijn, "Transmission interval estimates suggest pre-symptomatic spread of COVID-19," Mar. 2020.

G. Tkacik and W. Bialek, "Information Processing in Living Systems," *Annual Review of Condensed Matter Physics*, vol. 7, no. 1, pp. 89–117, 2016, _eprint: https://doi.org/10.1146/annurev-conmatphys-031214-014803.

UNESCO, "1.3 billion learners are still affected by school or university closures," unesco.org/news/13-billion-learners-are-still-affected-school-university-closures-educational-institutions, Apr. 2020, accessed: 2020-11-9.

L. Van Valen, "Haldane's Dilemma, Evolutionary Rates, and Heterosis," *The American Naturalist*, vol. 97, no. 894, pp. 185–190, May 1963, publisher: The University of Chicago Press. [Online]. Available: http://www.journals.uchicago.edu/doi/10.1086/282267

R. M. Viner, O. T. Mytton, C. Bonell, G. J. Melendez-Torres, J. Ward, L. Hudson, C. Waddington, J. Thomas, S. Russell, F. van der Klis, A. Koirala, S. Ladhani, J. Panovska-Griffiths, N. G. Davies, R. Booy, and R. M. Eggo, "Susceptibility to SARS-CoV-2 infection among children and adolescents compared with adults: A systematic review and meta-analysis," *JAMA Pediatr.*, Sep. 2020.

R. P. Walensky, H. T. Walke, and A. S. Fauci, "SARS-CoV-2 Variants of Concern in the United StatesâChallenges and Opportunities," *JAMA*, vol. 325, no. 11, pp. 1037–1038, Mar. 2021. [Online]. Available: https://doi.org/10.1001/jama.2021.2294

X. Wang, Z. Du, G. Huang, R. F. Pasco, S. J. Fox, A. P. Galvani, M. Pignone, S. C. Johnston, and L. A. Meyers, "Effects of cocooning on coronavirus disease rates after relaxing social distancing," *Emerg. Infect. Dis.*, vol. 26, no. 12, pp. 3066–3068, Dec. 2020.

N. L. Washington, K. Gangavarapu, M. Zeller, A. Bolze, E. T. Cirulli, K. M. S. Barrett, B. B. Larsen, C. Anderson, S. White, T. Cassens, S. Jacobs, G. Levan, J. Nguyen, J. M. Ramirez, C. Rivera-Garcia, E. Sandoval, X. Wang, D. Wong, E. Spencer, R. Robles-Sikisaka, E. Kurzban, L. D. Hughes, X. Deng, C. Wang, V. Servellita, H. Valentine, P. D. Hoff, P. Seaver, S. Sathe, K. Gietzen, B. Sickler, J. Antico, K. Hoon, J. Liu, A. Harding, O. Bakhtar, T. Basler, B. Austin, M. Isaksson, P. G. Febbo, D. Becker, M. Laurent, E. McDonald, G. W. Yeo, R. Knight, L. C. Laurent, E. d. Feo, M. Worobey, C. Chiu, M. A. Suchard, J. T. Lu, W. Lee, and K. G. Andersen, "Genomic epidemiology identifies emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States," *medRxiv*, p. 2021.02.06.21251159, Feb. 2021, publisher: Cold Spring Harbor Laboratory Press. [Online]. Available: https://www.medrxiv.org/content/10.1101/2021.02.06.21251159v1

R. A. Watson and E. Szathmáry, "How Can Evolution Learn?" *Trends in Ecology & Evolution*, vol. 31, no. 2, pp. 147–157, Feb. 2016. [Online]. Available: https://www.cell.com/trends/ecology-evolution/abstract/S0169-5347(15)00293-1

K. A. Weeden and B. Cornwell, "The small-world network of college classes: Implications for epidemic spread on a university campus," *Sociological Science*, vol. 7, pp. 222–241, 2020.

R. Wölfel, V. M. Corman, W. Guggemos, M. Seilmaier, S. Zange, M. A. Müller, D. Niemeyer, T. C. Jones, P. Vollmar, C. Rothe, M. Hoelscher, T. Bleicker, S. Brünink, J. Schneider, R. Ehmann, K. Zwirglmaier, C. Drosten, and C. Wendtner, "Virological assessment of hospitalized patients with COVID-2019," *Nature*, vol. 581, no. 7809, pp. 465–469, May 2020.

S. Wright, "The genetical structure of populations," *Ann Eugen*, vol. 15, no. 4, pp. 323–354, Mar. 1951.

B. E. Young, S. W. X. Ong, S. Kalimuddin, J. G. Low, S. Y. Tan, J. Loh, O.-T. Ng, K. Marimuthu, L. W. Ang, T. M. Mak, S. K. Lau, D. E. Anderson, K. S. Chan, T. Y. Tan, T. Y. Ng, L. Cui, Z. Said, L. Kurupatham, M. I.-C. Chen, M. Chan, S. Vasoo, L.-F. Wang, B. H. Tan, R. T. P. Lin, V. J. M. Lee, Y.-S. Leo, D. C. Lye, and Singapore 2019 Novel Coronavirus Outbreak Research Team, "Epidemiologic features and clinical course of patients infected with SARS-CoV-2 in singapore," *JAMA*, Mar. 2020.

J. Zhang, M. Litvinova, Y. Liang, Y. Wang, W. Wang, S. Zhao, Q. Wu, S. Merler, C. Viboud, A. Vespignani, M. Ajelli, and H. Yu, "Age profile of susceptibility, mixing, and social distancing shape the dynamics of the novel coronavirus disease 2019 outbreak in China," *medRxiv*, p. 2020.03.19.20039107, Mar. 2020, publisher: Cold Spring Harbor Laboratory Press. [Online]. Available: https://www.medrxiv.org/content/10.1101/2020.03.19.20039107v1

K. O. Zimmerman, I. C. Akinboyo, M. A. Brookhart, A. E. Boutzoukas, K. McGann, M. J. Smith, G. Maradiaga Panayotti, S. C. Armstrong, H. Bristow, D. Parker, S. Zadrozny, D. J. Weber, D. K. Benjamin, Jr, and ABC Science Collaborative, "Incidence and secondary transmission of SARS-CoV-2 infections in schools," *Pediatrics*, Jan. 2021.