Dominance Analysis Using Pathway Force Decomposition

Jeremy B. Sato

July 16, 2017

Abstract

This paper proposes a formal and mathematically rigorous framework for defining dominance based on necessary and sufficient conditions. A procedure is outlined for identifying dominant structure which is tested against several forms of the logistic growth model. The procedure is compared with other dominance methods such as PPM, LEEA, Ford's method, and the Loop Impact method. Relationships are identified between each method. The framework also indicates the robustness or fragility of a system. Shifts in dominance were found to occur when systems exist in fragile states in the dominance framework. The framework captures both structural and behavioral aspects of dominance found in other dominance methods. It also establishes a mathematical basis and definition for observed phenomena such as *shadow loop dominance* and *multiple loop dominance*.

1 Introduction

Loop dominance methods have at times produced conflicting or inconsistent results. One of the challenges in model analysis is the lack of a formal and rigorous definition of loop dominance.

This paper proposes a formal definition and criteria for dominance. Using this definition, a procedure is constructed for identifying dominant structure which is then tested against several forms of the logistic growth model. The procedure is then compared with other dominance methods, and relationships are identified between each method. The paper concludes with a discussion on how the proposed dominance framework can be used to characterize the robustness or fragility of a system.

2 Definitions and Implications

Consider the dynamic system described by the following n^{th}-order ordinary differential equation (ODE):

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t))$$

where,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \mathbf{f} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \tag{1}$$

where each f_i is differentiable in $\mathbf{x}$ and in u_i

Towards the goal of identifying elements of system structure which dominate behavior, the following definitions are proposed for *structure*, *behavior*, and *dominance*:

Behavior. *Behavior* is defined as the sign of the second time derivative of a state variable of interest, $x_j(t)$, evaluated at a specific point t_0 along the state's trajectory.

$$\textit{behavior} \text{ of } x_j(t) \text{ at time } t_0 \;=\; \operatorname{sgn} \ddot{x}_j(t_0)$$

Structure. The explanatory elements of system *structure* are the immediate causal pathways from state variables $(x_1, x_2, \ldots, x_n)$ to the first time derivative of the state variable of interest $\dot{x}_j(t) = f_j(\mathbf{x}, u_j)$, as shown in Figure 1. A causal pathway $p_{ijk}(x_i)$ is a scalar function representing a causal process which maps a *cause*, state x_i, to an *effect*, state derivative $\dot{x}_j$. There may exist more than one causal pathway (representing different causal mechanisms) from x_i to $\dot{x}_j$. In such cases, p_{ijk} is the k^{th} pathway from x_i to $\dot{x}_j$.

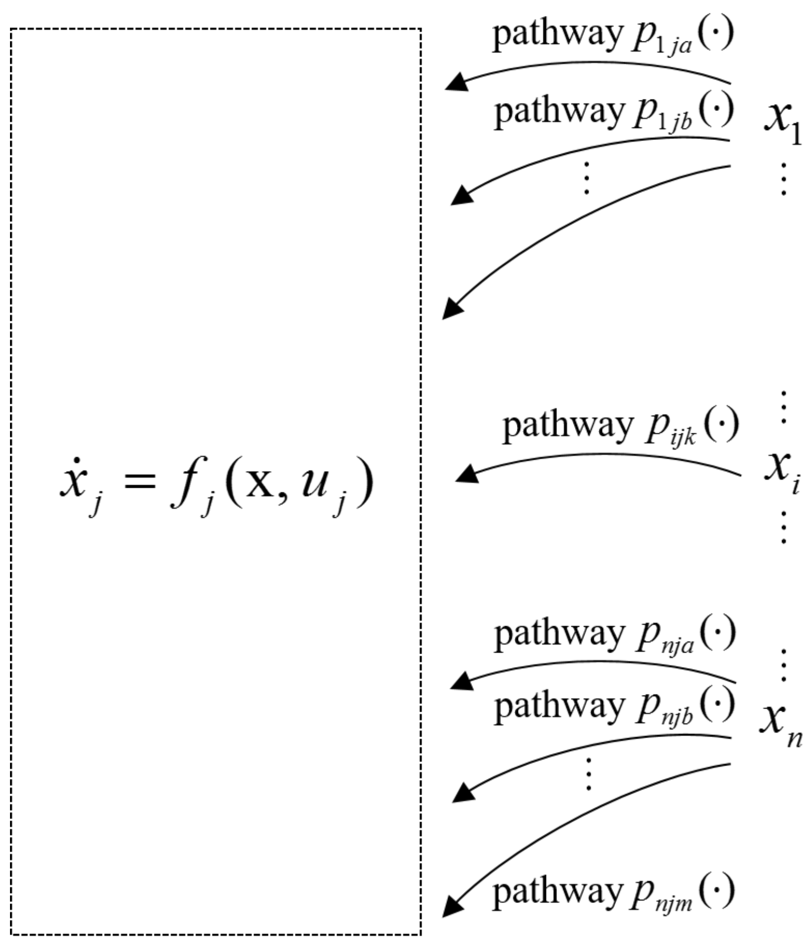

Figure 1: Immediate causal pathways from state variables $(x_1, x_2, \ldots, x_n)$ to the derivative of the state variable of interest $(\dot{x}_j)$.

Dominant. Given a state variable $x_j(t)$ whose behavior is of interest, and point t_0 along its trajectory, causal pathway p_{ijk} is *dominant* if and only if it is both necessary and sufficient for determining $\operatorname{sgn} \ddot{x}_j(t_0)$.

2.1 Decomposing *Behavior* into Pathway Contributions

From Equation 1, the dynamics of $\dot{x}_j$ can be succinctly written as a function of the state variables:

$$\dot{x}_j = f_j(x_1, x_2, \ldots, x_n, u_j) \tag{2}$$

However, in order to distinguish between multiple causal pathways from a common state variable, the dynamics can alternatively be expressed as a function of the pathways depicted in Figure 1, as shown in Equation 3.

$$\dot{x}_j = f_j(p_{1ja}, p_{1jb}, \ldots, p_{ijk}, \ldots, p_{nja}, p_{njb}, \ldots, p_{njm}, u_j) \tag{3}$$

From Equation 3, $\ddot{x}_j$ is derived using the chain rule of differentiation, and expressed as a sum of contributions from each immediate causal pathway p_{ijk}:

$$\ddot{x}_j = \frac{\partial f_j}{\partial p_{1ja}}\,\dot{p}_{1ja} + \frac{\partial f_j}{\partial p_{1jb}}\,\dot{p}_{1jb} + \ldots + \frac{\partial f_j}{\partial p_{ijk}}\,\dot{p}_{ijk} + \ldots + \frac{\partial f_j}{\partial p_{njm}}\,\dot{p}_{njm} + \frac{\partial f_j}{\partial u_j}\,\dot{u}_j \tag{4}$$

In Equation 4, the term

$$\frac{\partial f_j}{\partial p_{ijk}}\,\dot{p}_{ijk}$$

quantifies the contribution of causal pathway p_{ijk} to the second derivative ($\ddot{x}_j$). By analogy to Newton's second law of motion, this term can also be conceptualized as the force exerted by p_{ijk} on the variable x_j, causing an acceleration. The first factor of this term (the partial derivative) represents the gain of p_{ijk}, which is the change in f_j with respect to changes in p_{ijk}. The second factor is the rate of change of p_{ijk}. The second derivative, or acceleration, of the variable of interest ($\ddot{x}_j$) is expressed as the sum of the force contributions from each immediate pathway, and is visually depicted using a free body diagram as shown in Figure 2.

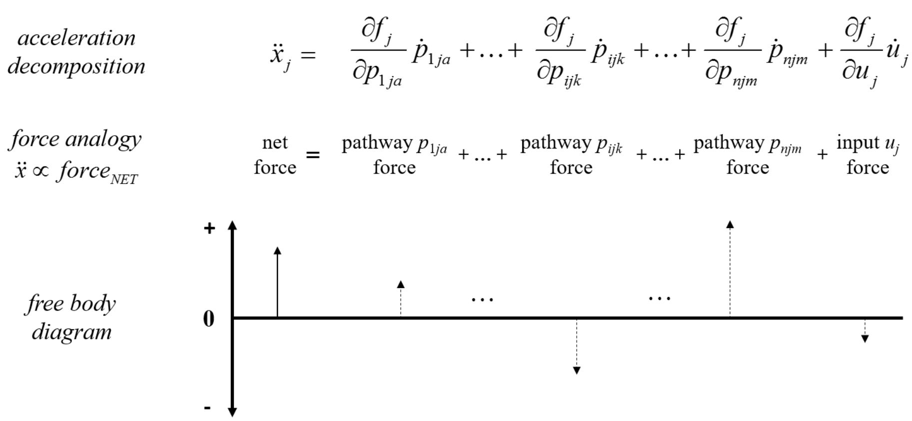

Figure 2: Free body diagram illustrating the contribution of each pathway to the acceleration of the variable of interest.

2.2 Illustration: Pathway Decomposition for the Logistic Growth Model

The logistic growth (or population growth) model, also known as the Verhulst equation as discovered by Pierre-François Verhulst in 1838 (see [11]) is a common model for examining shifting influence

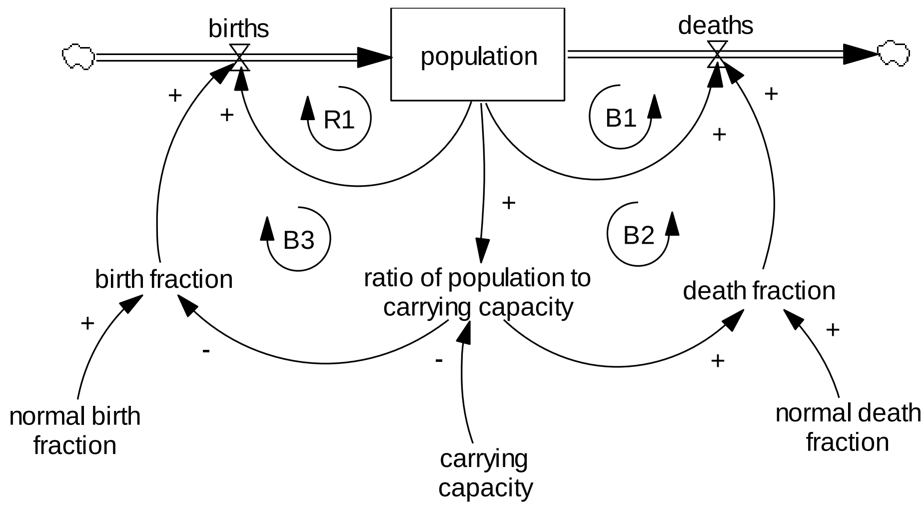

Figure 3: Stock and flow diagram of population growth.

between reinforcing and balancing feedback loops. Figure 3 shows a *stock and flow diagram* of the logistic growth model.

The center box variable, *population*, is the single stock (state variable). The double-lined arrows with valves flowing into and out of population are the flow variables (components of the state derivative), associated with *births* and *deaths*. The clouds represent unconstrained population sources and sinks. $R1$ is the reinforcing feedback loop associated with the birth process, in which births add to the population, further increasing the birth rate. $B1$ is the balancing feedback loop associated with deaths, in which deaths decrease the population, which subsequently slows the death rate. $B2$ and $B3$ are the balancing loops associated with the constraints of a fixed environment with a *carrying capacity* (maximum population size the environment can sustain based on finite resources). The ratio of the current *population* size to its *carrying capacity* affects the $birth\,fraction$ and $death\,fraction$ (also known as the birth and death fractional rates), which represents the fraction of births and deaths per unit of time. $Normal\,birth\,fraction$ and $normal\,death\,fraction$ are the fractional birth and death rates in an unconstrained environment. The dynamics of logistic growth are governed by the following equations:

$$\begin{aligned}
&\frac{d}{dt}population = births\, -\, deaths \\
&\qquad\qquad \dot{P} = b \cdot \left(1 - \frac{P}{C}\right) \cdot P \,-\, d \cdot \left(\frac{P}{C}\right) \cdot P \\
&\quad where \\
&\qquad\qquad P = population \\
&\qquad\qquad b = normal\,birth\,fraction \\
&\qquad\qquad d = normal\,death\,fraction \\
&\qquad\qquad C = carrying\,capacity
\end{aligned} \tag{5}$$

Through a simple change of variables, system 5 can be re-written as a function of the four distinct pathways from P to $\dot{P}$, representing feedback loops $R1$, $B1$, $B2$, and $B3$:

$$\begin{aligned} \dot{P} &= p_{114}(P) \cdot p_{111}(P) + p_{113}(P) \cdot p_{112}(P) \\ &where \\ R1: \ p_{111}(P) &= P \\ B1: \ p_{112}(P) &= -P \\ B2: \ p_{113}(P) &= d \cdot \left(\frac{P}{C}\right) \\ B3: \ p_{114}(P) &= b \cdot \left(1 - \frac{P}{C}\right) \end{aligned} \tag{6}$$

However, logistic equation 5 is also sometimes expressed in the following equivalent form [11, 12]:

$$\begin{aligned} \dot{P} &= \alpha P - \beta P^2 \\ &where \\ \alpha &= b \\ \beta &= \frac{(b+d)}{C} \end{aligned} \tag{7}$$

This alternate form illustrates how for $\alpha \gg \beta$, when P is relatively small, the reinforcing growth of *births* drives the dynamics $\dot{P}$, but as P gets larger and reaches its carrying capacity C, the squared term increases and slows the rate of growth $\dot{P}$. System 7, while mathematically equivalent to 5, can be written as a function of two pathways representing one reinforcing loop and one balancing loop.

$$\begin{aligned} \dot{P} &= p_{111}(P) + p_{112}(P) \\ &where \\ R1: \ p_{111}(P) &= bP \\ B1: \ p_{112}(P) &= -\left(\frac{b+d}{C}\right) \cdot P^2 \end{aligned} \tag{8}$$

This first-order example illustrates how even for simple first-order models, there can be multiple ways to define causal pathways. Systems 6 and 8, while mathematically equivalent, represent two different sets of causal mechanisms. The choice of pathways depends on the causal mechanisms and the correspondence between model variables and real-world processes. The logistic equation has been applied to the fields of biology, chemistry, demography, ecology, economics, sociology, and political science. In each instance, variable and pathway decomposition depends on the phenomena being described. While pathway selection does not change the behavior of the model, as will be shown, it does change how explanations are developed in terms of structure *dominating* behavior. Since different pathway decompositions represent different variable transformations of the system, the chain rule of differentiation allows the modeler to choose the most appropriate decomposition of the second derivative into contributing forces or causal mechanisms.

2.3 Identifying Necessary and Sufficient Pathways

With the second derivative of the variable of interest expressed as a function of causal pathways (Equation 4), dominance is evaluated by examining each pathway and its effect on the sign of the second derivative. Testing for necessary and sufficient conditions requires pathways to be isolated and their effects somehow removed or deactivated, independent from one another. This raises the question, *What does independent removal or deactivation look like mathematically*?

2.3.1 Pathway Removal Versus Deactivation

The practice of isolating and deactivating partial model structures (while leaving the rest of the model intact) and examining the impact on behavior through simulation, has a long history in system dynamics (SD) [10]. For the purpose of identifying dominant structure, there are two primary ways of isolating and deactivating structure. The first is complete removal, and the second is deactivation by holding the partial structure constant [2, 9, 5].

To illustrate the difference between these approaches, consider a second-order system with state variables x and y, in which y is the variable whose behavior is of interest and pathway x to y is to be removed or deactivated, as shown in the upper left corner of Figure 4.

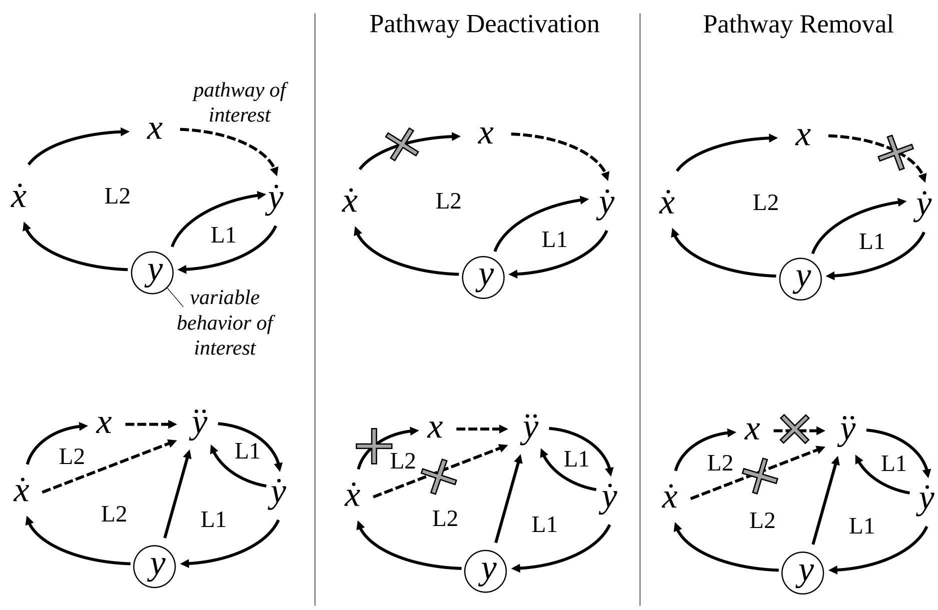

Figure 4: Illustration of pathway deactivation versus pathway removal in a second-order system.

This system has two feedback loops: $L1$ which is a minor feedback loop from y to itself, and $L2$, a major feedback loop going through both y and x. The intent behind deactivating pathway x to y is to test the influence of feedback loop $L2$ on $\ddot{y}$. The bottom left diagram in Figure 4 shows an equivalent representation of the causal structure using the decomposition of $\ddot{y}$, highlighting the role of both the states and their derivatives. The middle column of the figure represents the deactivation approach. The pathway is deactivated by holding the value of the pathway constant, which requires holding x constant (accomplished by setting $\dot{x}$ to zero). The result (lower middle figure) is that the impact of the dynamics of the pathway are eliminated (hence $L2$ is eliminated), while the pathway itself remains constant and still appears in the equation for $\ddot{y}$, as a gain applied to feedback loop $L1$. The right column represents the approach of fully removing the pathway from x to y. Under this approach, the pathway is completely removed from the equation of $\dot{y}$, and hence neither x nor $\dot{x}$ affect $\ddot{y}$.

For this paper, the deactivation approach (middle column) is used to test for necessary and sufficient pathways. The primary reason is because the full removal method appears to test two effects simultaneously, which is undesirable. It tests the effect of both the existence of a pathway as well as its dynamics. This is evident in the lower right diagram of Figure 4 in which both the effects of x and $\dot{x}$ are eliminated, whereas in the deactivation method (lower middle diagram), only the effect of $\dot{x}$ is eliminated. In other words, the deactivation approach tests the effect of the *dynamics* of the pathway (generated by feedback loop $L2$), but not its *existence*. This is especially preferable if pathway x to y is an integral aspect of the theory for y (for example, is required to make the equation $\dot{y}$ logically and dimensionally consistent). Through deactivation, one may desire to know if x can be replaced by a parameter or exogenous variable, or if it being a state variable within a feedback loop is required in order for y to exhibit certain behaviors.

Furthermore, to test the full removal of a pathway goes beyond the question of dominant structure, and to one of model simplification. To test the full removal of pathway x to y requires the formulation of an alternate, dimensionally consistent theory about the causal mechanisms affecting $\dot{y}$. There is no general approach for removing a variable from an equation - the method of removal depends on the nature of the equation and requires thoughtful consideration from the modeler. In some cases, when pathways are nonlinearly coupled, it may not be possible to remove one pathway without affecting other pathways, which violates the principle that the influence of different elements of structure should be isolated and tested independently [10]. In some cases, it may not be possible to reformulate a coherent rate equation without the causal pathway of interest[1]. In contrast, the theory (equation) of $\dot{y}$ remains unaltered and dimensionally consistent if the pathway is deactivated and held constant, but not removed. This also results in a stronger test in which only the dynamics of the pathway are being isolated and tested, as opposed to both dynamics and existence.

There are two important additional observations about the deactivation of pathways, made evident in Figure 4. First, with respect to model *behavior*, is that deactivation only impacts $\ddot{y}$, maintaining the smoothness of trajectories, whereas full removal impacts both $\ddot{y}$ and $\dot{y}$ at the same time, resulting in non-smooth trajectories at the time of removal. This also illustrates how, when conducting pathway deactivation, the second derivative of y alone is sufficient for determining dominance and shifts in dominance, providing further support for the proposed definition of *behavior*.

Second, with respect to model *structure*, suppose there exists an additional feedback loop, $L3$, through state variable z, between y and $\dot{x}$, as shown in Figure 5. Also, suppose that after deactivating and testing the pathway from x to y, it is shown to be necessary for determining the sign of $\ddot{y}$. A natural question might be which feedback loop, $L2$ or $L3$, is the most influential loop influencing y at that time, since both contain the necessary pathway from x to y. Figure 5 shows that, with respect to $\ddot{y}$, this question is unanswerable. At the time of testing, the influence of loops $L2$ and $L3$ cannot be distinguished from each other. Their effect is mediated through $\dot{x}$, and thus one additional simulation time step is required to observe their impact on $\ddot{y}$ [2]. A similar observation was also made by Hayward and Boswell in their discussion of the limitations of the Loop Impact method [4]. It would be more appropriate to ask about the relative influence between loops $L2$ and $L3$ with respect to $\ddot{x}$, instead of $\ddot{y}$. Because the systems are state-determined, and states accumulate

[1]For example, in the Lotka-Volterra predator prey model, the causal pathways of the major balancing feedback loop between predators and prey cannot be removed without also removing other minor feedback loops as well (due to their nonlinear coupling), in order to maintain a dimensionally consistent and coherent set of equations.

[2]Deactivating a pathway n state variables removed from the variable whose behavior is of interest requires n simulation time steps before an effect is observed. This fact makes some loop deactivation methods inadequate for the proposed definition of *dominance*.

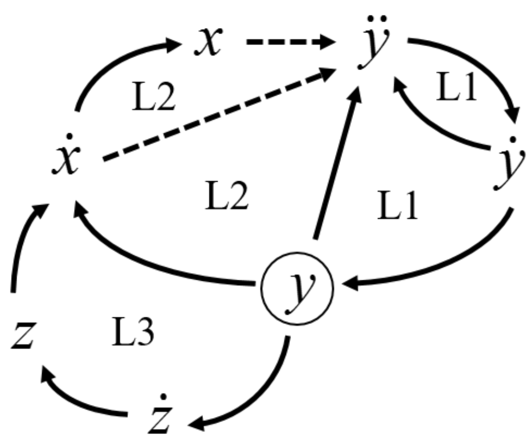

Figure 5: Inability to distinguish between loops containing a common pathway.

(integrate) the effects of feedback loops from history to the present, the influence of loops sharing the same immediate pathway to the variable of interest cannot be distinguished from each other. Their influence, rather, unfolds as the simulation progresses over time and their impacts observed on the pathway they have in common. This lends further support to considering pathways, rather than feedback loops, as the explanatory element of system *structure*.

2.3.2 Tests for Necessary and Sufficient Pathways

The following table summarizes the tests for necessary, sufficient, and contributory pathways.

Type	Description	Necessary?	Sufficient?	Test
1	necessary pathway	yes	no	$\operatorname{sgn} \ddot{x} \rightarrow$ pathway no pathway $\rightarrow -\operatorname{sgn} \ddot{x}$
2	sufficient pathway	no	yes	pathway $\rightarrow \operatorname{sgn} \ddot{x}$ $-\operatorname{sgn} \ddot{x} \rightarrow$ no pathway
3	necessary and sufficient pathway	yes	yes	pathway $\leftrightarrow \operatorname{sgn} \ddot{x}$ no pathway $\leftrightarrow -\operatorname{sgn} \ddot{x}$
4	contributory pathway	no	no	sgn pathway $= \operatorname{sgn} \ddot{x}$
5	none of the above	no	no	sgn pathway $= -\operatorname{sgn} \ddot{x}$

Table 1: Five types of causal pathways.

Equation 4 expresses *behavior* $\ddot{x}_j$ as a sum of force contributions (F_{ijk}) from each pathway

$$\ddot{x}_j = F_{1ja} + F_{1jb} + \ldots + F_{ijk} + \ldots + F_{njm} + F_{u_j}$$

where:

$$F_{ijk} = \frac{\partial f_j}{\partial p_{ijk}} \dot{p}_{ijk}$$

Pathway p_{ijk} is *necessary* if its deactivation changes the sign of $\ddot{x}_j$. Deactivating pathway p_{ijk} is accomplished by setting $\dot{p}_{ijk} = 0$ in Equation 4 at the time of deactivation, which zeros out the force contribution F_{ijk}. Thus, p_{ijk} is a necessary pathway if the following inequality holds:

$$\operatorname{sgn} \ddot{x}_j \neq \operatorname{sgn} (\ddot{x}_j - F_{ijk}) \tag{9}$$

Pathway p_{ijk} is *sufficient* if it alone guarantees the sign of the $\ddot{x}_j$. The first condition for this to hold is that the sign of F_{ijk} must be the same as the sign of $\ddot{x}_j$:

$$\operatorname{sgn} \ddot{x}_j = \operatorname{sgn} F_{ijk}$$

Secondly, sufficiency of p_{ijk} requires that $\operatorname{sgn} \ddot{x}_j$ be unaffected by any combination of other pathways. Equivalently, disproving the sufficiency of p_{ijk} requires finding some combination of other active pathways which results in a sign change of $\ddot{x}_j$. Because pathway contributions to $\ddot{x}_j$ are purely additive, this is equivalent to the absolute value of F_{ijk} being greater than the absolute value of the sum of all opposing pathway contributions (i.e., pathways whose contributions are opposite in sign of F_{ijk}). Therefore, p_{ijk} is a sufficient pathway if the following inequality holds:

$$|F_{ijk}| > \left| \sum_{\left\{\operatorname{sgn} F_{ljm} \neq \operatorname{sgn} F_{ijk}\right\}} F_{ljm} \right| \tag{10}$$

Lastly, if pathway p_{ijk} is neither sufficient nor necessary, but contributes in the same direction as the observed behavior,

$$\operatorname{sgn} \ddot{x}_j = \operatorname{sgn} F_{ijk}$$

then it is called a *contributory* pathway.

2.4 Illustration: Dominance Framework Applied to a Linear Model

The above criteria for necessary, sufficient, and contributory pathways creates a *dominance framework* which is applied to a first-order linear model containing multiple pathways, in order to illustrate the different possible combinations of pathway types. Consider the following first-order system (Figure 6 and Equation 11) with state variable x, governed by three linear inflows and a single outflow.

$$\begin{gathered} \dot{x} = \alpha_1 x + \alpha_2 x + \alpha_3 x - \beta x \\ where, \\ \alpha_1,\, \alpha_2,\, \alpha_3,\, \beta \, \geq 0 \end{gathered} \tag{11}$$

In this model, the four pathways could be aggregated into a single pathway with fractional rate $(\alpha_1 + \alpha_2 + \alpha_3 - \beta)$, however, assume each pathway represents a unique causal process and the desire is to understand the relative influence of each pathway on the behavior of x. The pathways are identified as:

$$\begin{aligned} \text{path}\,1: &\ p_{111}(x) = \alpha_1 x \\ \text{path}\,2: &\ p_{112}(x) = \alpha_2 x \\ \text{path}\,3: &\ p_{113}(x) = \alpha_3 x \\ \text{path}\,4: &\ p_{114}(x) = -\beta x \end{aligned}$$

The second derivative of x is expressed as the sum of each pathway contribution:

$$\ddot{x} = \alpha_1 \dot{x} + \alpha_2 \dot{x} + \alpha_3 \dot{x} - \beta \dot{x} \tag{12}$$

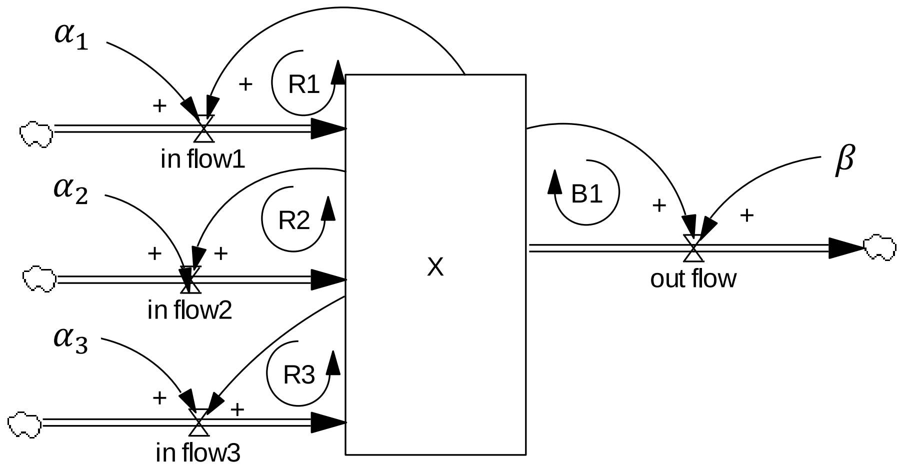

Figure 6: Stock and flow diagram of first-order linear model with four causal pathways.

The relative contributions of each pathway are solely determined by the parameter values which are constant. Therefore, there can be no shifts in dominance over time. However, postulating different sets of values for the parameters illustrates the different possible outcomes with respect to necessary, sufficient, and contributory pathways. Without loss of generality, assume at the time of interest t_0, $\dot{x}(t_0) = 1$. Equation (12) reduces to:

$$\ddot{x}(t_0) = \alpha_1 + \alpha_2 + \alpha_3 - \beta \tag{13}$$

Six sets of parameter values in Table 2 illustrate the possible outcomes of necessary, sufficient, and contributory pathways, corresponding to different points on the dominance framework in Figure 7.

	$\boldsymbol{\alpha_1}$	$\boldsymbol{\alpha_2}$	$\boldsymbol{\alpha_3}$	$\boldsymbol{\beta}$
case 1	3	3	3	5
case 2	6	3	3	5
case 3	6	6	1	5
case 4	4	2	2	5
case 5	6	1	1	5
case 6	3	3	1	5

Table 2: Six sets of parameter values in the first-order linear model.

Using a visual representation similar to a free body diagram, the results of each case are displayed in Figure 8, in which $\ddot{x}(t_0)$ is represented by the first vertical bar, labeled *sum*, and the subsequent vertical bars are the individual contributions of each pathway.

Figure 9 illustrates how the six cases map onto the dominance framework.

Case 1: No necessary or sufficient pathways. Multiple pathways (1, 2, and 3), each exerting a force smaller than opposing pathway 4 (and thus insufficient), collectively exert a force

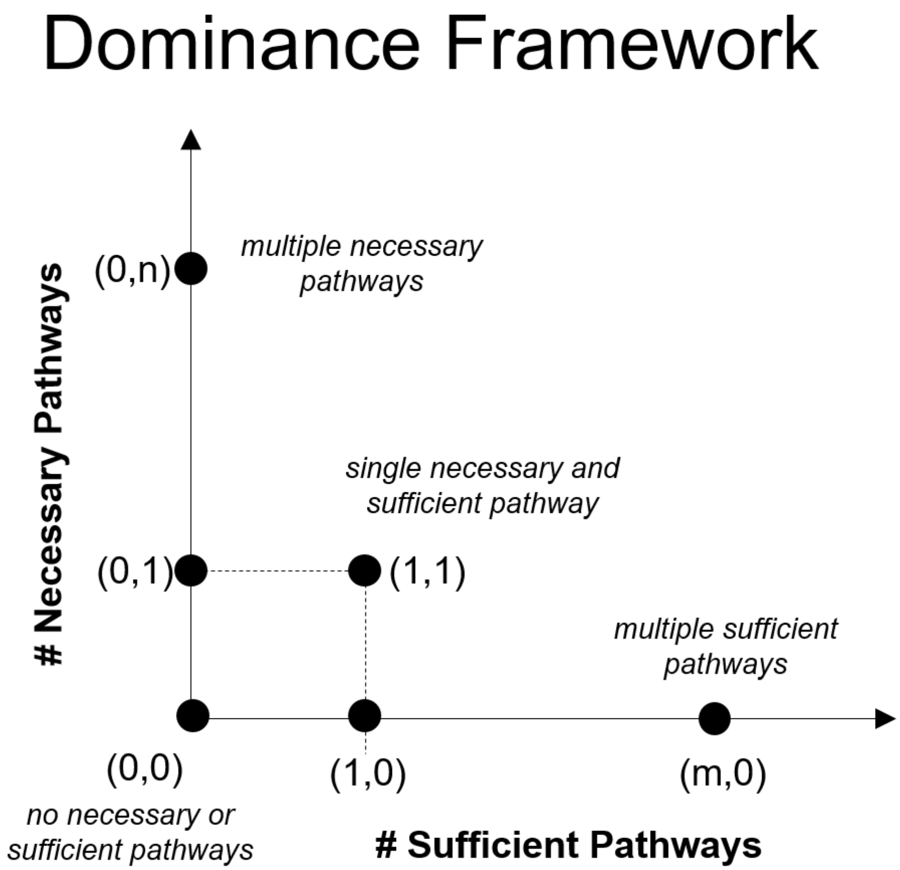

Figure 7: Dominance framework depicting the number of necessary and sufficient pathways that determine the behavior of a variable of interest.

greater than opposing pathway 4 with margin such that no pathway is necessary. Pathways 1, 2, and 3 are therefore considered *contributory* pathways. An analogy is a game of *tug of war* in which on one side of the rope is a team of many weaker individuals who collectively greatly over-power a single opposing strong person. **claim:** A minimum of four pathways are required for a system to be in this condition. **proof:** Suppose there are only three pathways, one opposing the other two. The one opposing pathway cannot exert greater strength than the other two together, otherwise it would be necessary and sufficient. Of the remaining two pathways, if one is not necessary, then the other must be sufficient for determining the sign, which contradicts the claim that no pathways are necessary or sufficient.

Case 2: One sufficient and no necessary pathways. Pathway 1 is sufficient (has magnitude greater than that of opposing pathway 4), but is not necessary, in that pathways 2 and 3, together, also exert a greater force than opposing pathway 4. As in case 1, pathways 2 and 3 are neither necessary nor sufficient. Using the same *tug of war* analogy in case 1, this case occurs when one of the weaker individuals on the winning team is replaced by a strong person. **claim:** A minimum of four pathways are required for a system to be in this condition. **proof:** Suppose there are only three pathways, one opposing the other two. The one opposing pathway cannot exert greater strength than the other two together, otherwise it would be necessary and sufficient. Of the remaining two pathways, only one of them is not sufficient. The other then is not only sufficient, but must also be necessary, otherwise its removal would cause the sign to change, which contradicts the claim that

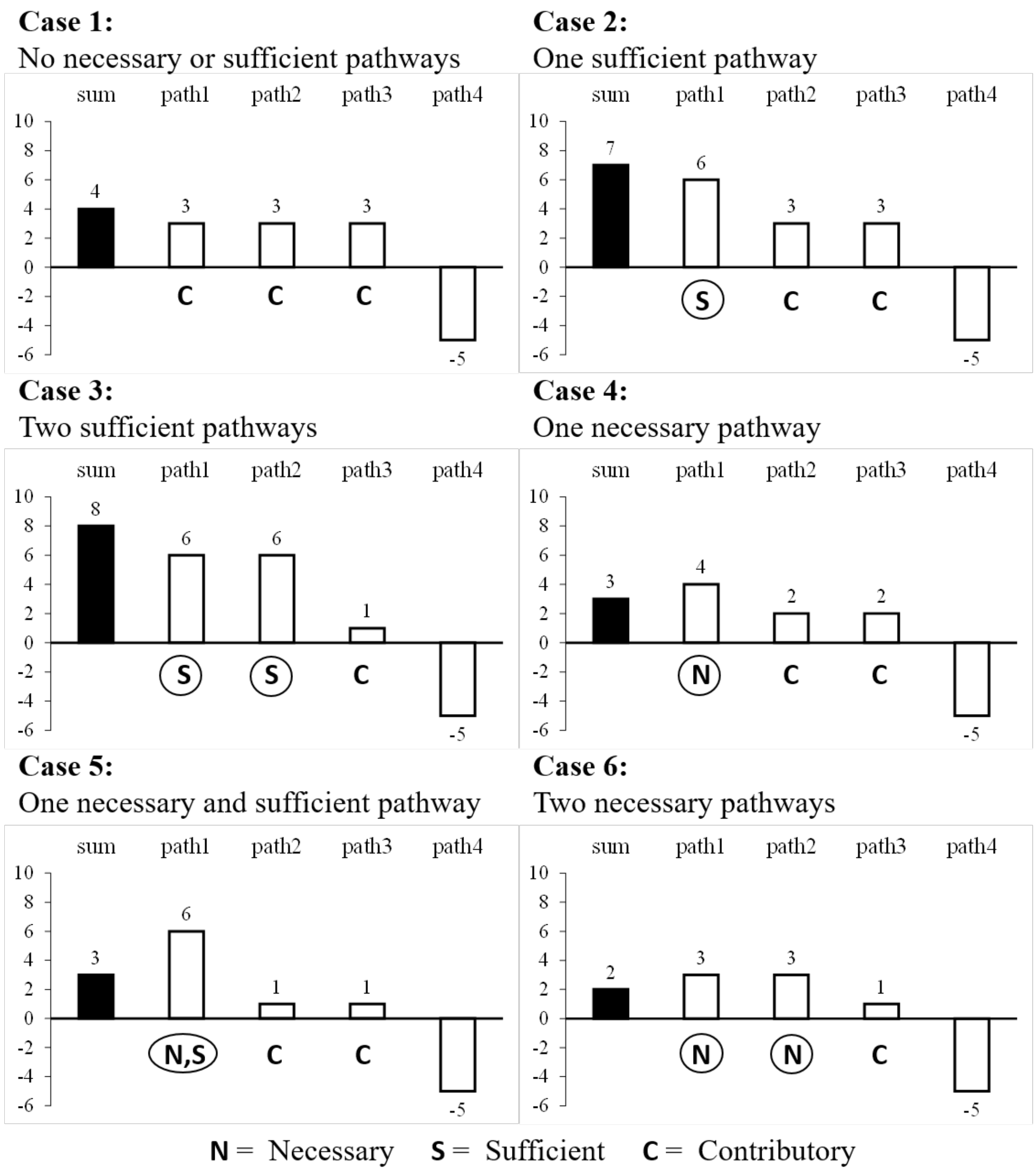

Figure 8: Six cases: combinations of necessary, sufficient, and contributory pathways.

none are necessary.

Case 3: Two sufficient and no necessary pathways. Pathways 1 and 2 are both sufficient (their force magnitudes are each greater than the opposing force from pathway 4), and therefore neither alone are necessary since their individual deactivation does not change the sign. Pathway 3, as in the previous cases, is neither sufficient nor necessary, but contributes in the same direction as the observed behavior. **claim:** A minimum of two pathways are required for a system to be in this condition. **proof:** By definition, a minimum of two pathways are required for there to exist two sufficient pathways. To illustrate how only two pathways are required, consider the current example of Case 3 but with pathways 3 and 4 completely removed. Pathways 1 and 2 would remain individually sufficient for producing the sign, but each alone unnecessary since they are both contributing in the same direction with no opposing force.

Case 4: One necessary and no sufficient pathway. Pathway 1 is necessary (deactivating it

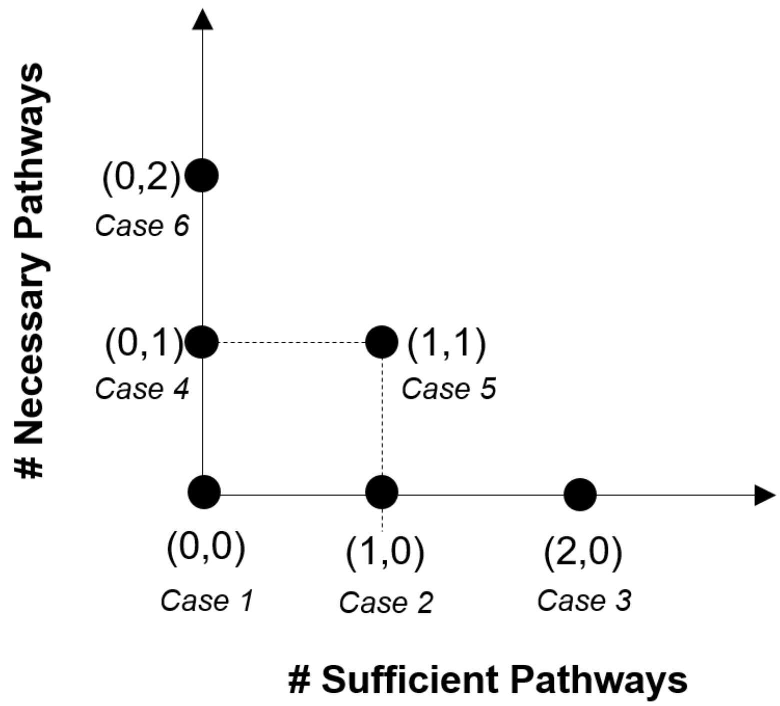

Figure 9: Six cases mapped onto the dominance framework.

results in a change in sign), but it alone is not sufficient (its force magnitude is less than that of the opposing pathway 4). Pathways 2 and 3, as in Cases 1 and 2, are contributory, but neither necessary nor sufficient. In this case, consider the analogy of a sports team with a single all-star player whose talent is required for the team to win, but who alone does not make up a team and therefore is insufficient. The remaining team members contribute to the team's success but are neither sufficient nor necessary because they are easily replaced. **claim:** A minimum of four pathways are required for a system to be in this condition. **proof:** Suppose there are only three pathways. At least one pathway must exert a force in an opposite direction from the other two, otherwise all pathways would be sufficient. Consider one pathway exerting an opposite force from the two others. Of the remaining two, if one is necessary but not sufficient, then the other must also be necessary, which contradicts the assumption that only one is necessary.

Case 5: One necessary and sufficient pathway. Pathway 1 is both necessary and sufficient. If pathway 1 is deactivated, $\ddot{x}(t_0)$ changes signs, so it is necessary. The absolute value of pathway 1's contribution is greater than that of all opposing paths, so pathway 1 is also sufficient. This is the only case in which a pathway meets the established criteria for *dominance*. There can be at most one necessary and sufficient pathway at a point in time, and that if a system has a single necessary pathway and a single sufficient pathway, they must be one and the same. **claim:** A minimum of one pathway is required for a system to be in this condition. **proof.** Consider this example but with pathways 2, 3, and 4 removed. Pathway 1 would remain necessary and sufficient for determining the behavior since its removal would change the sign to zero, and since there are no opposing pathways.

Case 6: Two necessary and no sufficient pathways. Pathways 1 and 2 are both necessary in that their individual deactivations result in a change in sign, but they are also individually insufficient in that their force magnitudes are less than the opposing force magnitude from pathway 4. To use another sports analogy, consider a team with no back-up players and in which every member of the team is necessary to play their position. **claim:** A minimum of three pathways are required for a system to be in this condition. **proof.** Consider a system with only two pathways. As demonstrated earlier, if both contribute in the same direction then both are sufficient and neither necessary. If they contribute in opposite directions, then the one with larger magnitude will be necessary and sufficient. Therefore, this case is not possible with only two pathways. For a three pathway system, consider the current example but with pathway 3 removed. Pathways 1 and 2 would still both be necessary but not individually sufficient.

3 Pathway Force Decomposition Procedure

The above process for identifying necessary, sufficient, and contributory pathways, derived from the proposed definitions of *structure*, *behavior*, and *dominance*, is now summarized in the following steps, and is referred to as the pathway force decomposition (PFD) procedure.

1. Select state variable whose behavior is of interest.

2. Specify initial conditions and time horizon of interest, over which to perform dominance analysis.

3. Specify immediate pathways to the state variable of interest, as functions of other state variables. Express the derivative of the state variable of interest as a function of the pathways.

4. Using the chain rule, express the second derivative of the behavior of interest as a sum of contributions from each pathway, by computing each partial derivative.

5. For each time step within the time interval of interest, calculate the force contributions of each pathway and test each pathway for necessary, sufficient and contributory conditions.

4 Application to Logistic Growth Model

The PFD procedure is applied to the logistic growth model, which exhibits transient S-shape behavior. This model is chosen due to its association with general growth processes from which the term *feedback loop dominance* originally emerged. The logistic growth model is also the canonical example of *shifts in loop dominance* and is therefore an important test for a formal definition of dominance. Three forms of the logistic model are analyzed and results compared with current methods.

4.1 Logistic Model Form 1: Reinforcing Loop and Constraining Loop

Richardson analyzed the two-loop logistic growth model (Figure 10) in which $R1$ is reinforcing growth, and $B1$ is the constraining effect from carrying capacity C on the fractional growth rate α [11] (also, see [12, p. 296]).

Step 1: Behavior of interest. Since population P is the only state variable, it is the variable of interest.

Step 2: Initial conditions and time horizon. The population begins with a single member, $P(0) = 1$. The model is analyzed from $t = 0$ to $t = 100$.

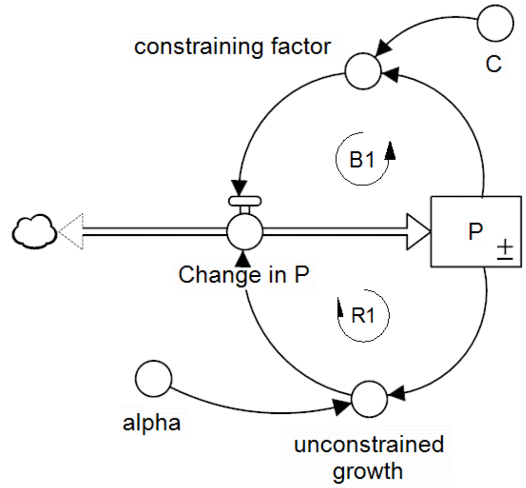

Figure 10: Stock and flow diagram of the logistic growth model (form 1).

Step 3: Express $\dot{P}$ as a function of its immediate pathways.

$$\begin{aligned} \dot{P} &= p_{111} \cdot p_{112} \\ p_{111} &= \alpha P \\ p_{112} &= 1 - \frac{P}{C} \end{aligned} \tag{14}$$

Pathway p_{111} is the pathway from P through auxiliary variable *unconstrained growth*, and represents the reinforcing growth feedback loop $R1$. Pathway p_{112} is the pathway from P through auxiliary variable *constraining factor*, and represents the balancing feedback loop $B1$ from carrying capacity C which constrains growth.

Step 4. Express $\ddot{P}$ as a function of pathway force contributions.

$$\begin{aligned} \ddot{P} &= F_{111} + F_{112} \\ F_{111} &= \frac{\partial \dot{P}}{\partial p_{111}} \dot{p}_{111} = \left(1 - \frac{P}{C}\right) \cdot \alpha \dot{P} \\ F_{112} &= \frac{\partial \dot{P}}{\partial p_{112}} \dot{p}_{112} = \alpha P \cdot \left(\frac{-\dot{P}}{C}\right) \end{aligned} \tag{15}$$

Step 5. For each time step, calculate force contributions and identify necessary, sufficient, and contributory pathways. Figure 11 shows the results of the simulation and analysis. The top graph is the behavior over time of state variable P. The middle graph plots $\ddot{P}$ (*Net Force*) along with the force decomposition of Path 1 (p_{111}) and Path 2 (p_{112}). The bottom graph shows which paths are necessary and which are sufficient for determining the sign of $\ddot{P}$.

4.1.1 Results

P exhibits the expected S-shape growth: divergent exponential growth followed by convergent goal-seeking growth. The inflection point ($\ddot{P} = 0$) occurs around time $t = 46$, when P reaches half its carrying capacity (50). Path 1 (associated with reinforcing feedback) always exerts a positive force, while Path 2 (associated with the balancing feedback loop) always exerts a negative force. From

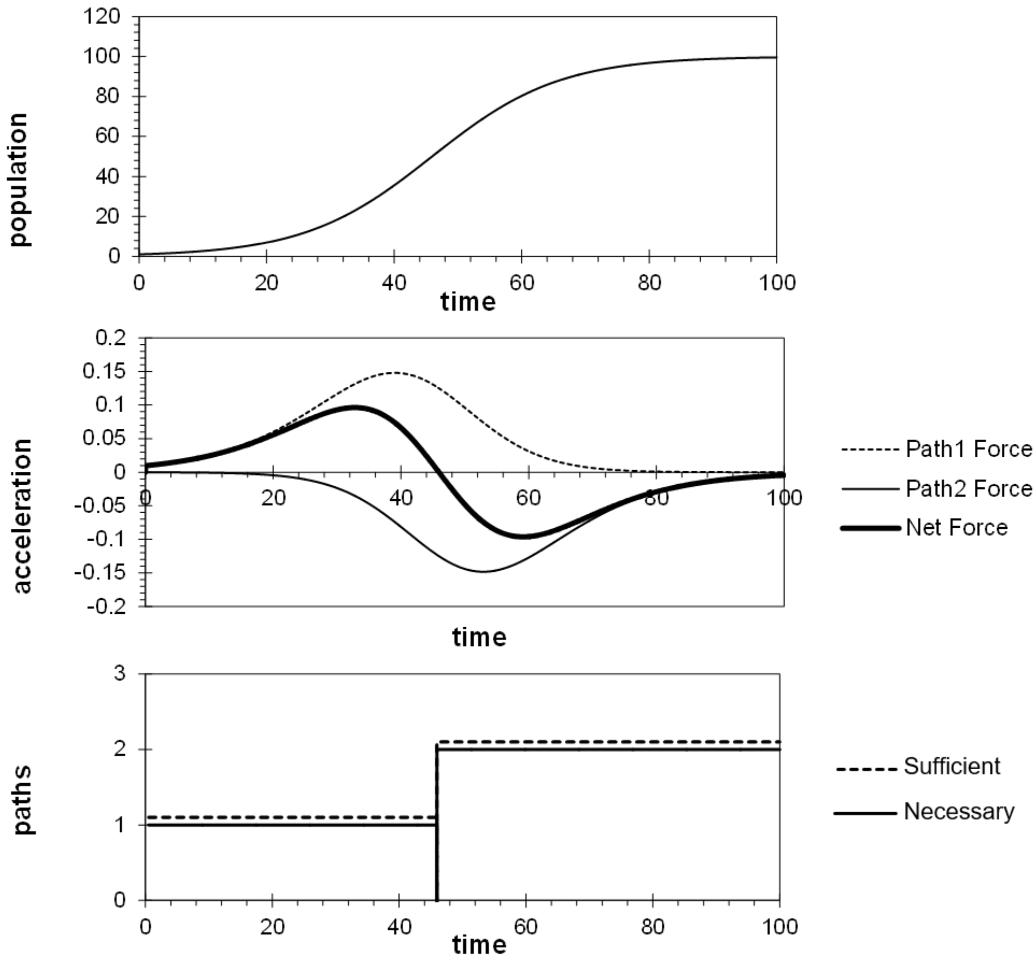

Figure 11: Simulation results for logistic growth model (form 1).

$t = 0$ to the inflection point $t = 46$, the force exerted by Path 1 is greater in magnitude than the force exerted by Path 2, and thus is both necessary and sufficient for creating positive acceleration (and by definition is *dominant*). After $t = 46$, Path 2 exerts a greater force than Path 1 and is necessary and sufficient for creating negative acceleration (deceleration), and therefore *dominance shifts* to Path 2. As time continues, both forces decrease as P approaches equilibrium.

4.1.2 Excursion

Consider the case in which the population begins above its carrying capacity (Figure 12). P exhibits exponential decay as it approaches its carrying capacity, C. The net force is always positive, slowing the decline. Path 2 (representing the balancing loop associated with the carrying capacity) exhibits the greatest force, however both Paths 1 and 2 exhibit positive forces and so both are sufficient (neither are necessary) for generating the observed behavior. Therefore, neither pathway is dominant. The result that Path 2 is sufficient is intuitive since it is associated with the balancing feedback loop and has a negative closed-loop gain. To understand, however, how Path 1 (associated with the reinforcing growth loop) is sufficient for generating this behavior, observe that if $P > C$, Path 2 is negative and therefore, when deactivated, switches the polarity of the feedback loop associated with Path 1. This illustrates how nonlinear coupling of pathways can change the polarity of feedback loops and lead to shifts, not just in loop dominance, but the polarity of feedback loops

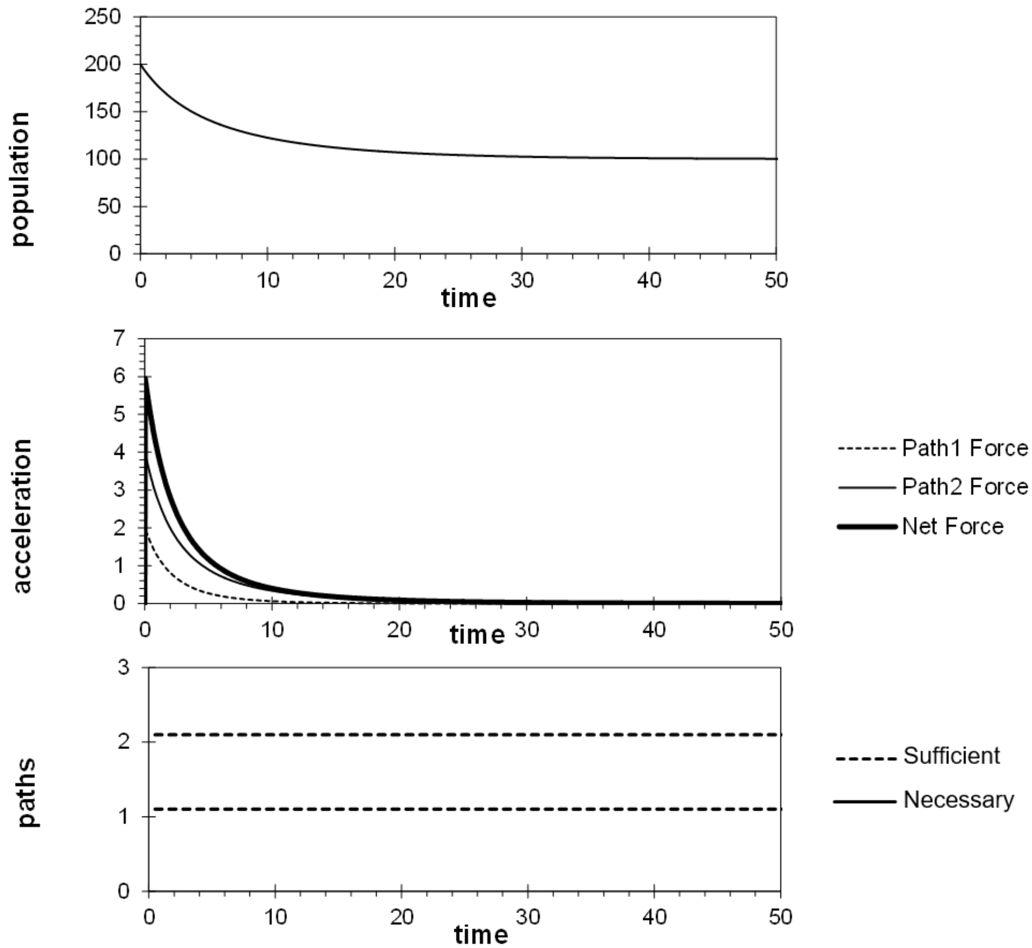

Figure 12: Simulation results for logistic growth model (form 1) ($P > C$).

in different regions of the state space.

4.1.3 Discussion and Comparison with Other Methods

The results from applying the PFD procedure agree with the conclusions from Richardson [11] and Sterman [12] who use the concept of *dominant polarity* to determine loop dominance for first-order two-loop systems, where dominant polarity is positive if population is less than half the carrying capacity (and thus the reinforcing loop dominates), and dominant polarity is negative if population is greater than one half (and thus the balancing loop dominates). The results also agree with numerous examples from SD instructional literature which informally introduce the concept of *shifts in dominance* based on the logistics model [1, 7]. However, while applying the concept of *dominant polarity* to identify loop dominance is straightforward when $P < C$, for $P > C$, Richardson and Sterman are silent on the specific roles of the two feedback loops. Other literature on logistic growth is also silent on loop dominance when $P > C$. The PFD procedure provides a formal and rigorous answer to this question.

Mojtahedzadeh performs dominance analysis on the logistic growth model, in the form of the Susceptible-Infectious (SI) epidemic model, using the Pathway Participation Method (PPM) [8]. The SI model uses the following interpretation of variables which make-up the two pathways in system 14:

$$\begin{aligned}
population\,(P) &= infectious\,(I)\\
\alpha &= contact\,rate\,(c) \times infectivity\,(i)\\
carrying\,capacity\,(C) &= population\,size\,(N)\\
\text{Contagion Reinforcing Loop}: p_{111} &= c\,i\,I\\
\text{Depletion Balancing Loop}: p_{112} &= \left(1 - \frac{I}{N}\right)
\end{aligned}$$

PPM, which uses the total pathway participation metric (TPPM) to determine dominance, concludes that the contagion loop (p_{111}) is dominant in the first phase of S-shape growth, followed by the dominance of the depletion loop (p_{112}) in the second phase. For this two-loop model, because PPM identifies the loop which contributes most to ($\ddot{I}/\dot{I}$) as the dominant loop, and each loop always has opposite contributions, the one with largest magnitude measure is also guaranteed to be both necessary and sufficient, and so the results agree with the PFD procedure. The case in which population starts above the carrying capacity does not apply to the SI model, since the infectious population can never be greater than the total population, so this case was not analyzed by PPM.

Similarly, Kampmann and Oliva [6] apply PPM to a simple diffusion model of $adopters\,(A)$ and $potential\,adopters\,(N-A)$ having the same mathematical and causal structure as the SI model, in which:

$$\begin{aligned}
population\,(P) &= adopters\,(A)\\
\alpha &= contact\,rate\,(c) \times adoption\,fraction\,(i)\\
carrying\,capacity\,(C) &= population\,size\,(N)\\
\text{Word of mouth Reinforcing Loop}: p_{111} &= c\,i\,A\\
\text{Market Saturation Balancing Loop}: p_{112} &= \left(1 - \frac{A}{N}\right)
\end{aligned}$$

The PPM results agree with the PFD results and identify the $word\,of\,mouth$ reinforcing loop as dominant in the first phase, and the $market\,saturation$ balancing loop as dominant in the second phase. Kampmann and Oliva also apply loop eigenvalue elasticity analysis (LEEA) to the same diffusion model, using influence metrics based on the derivative of eigenvalues of the linearized system with respect to loop parameters. LEEA also identifies the $word\,of\,mouth$ loop as dominant in the first phase, followed by the dominance of the $market\,saturation$ loop, agreeing with both PPM and the PFD procedure.

Taylor and colleagues [13] test their experimental statistical screening procedure on the same diffusion model, which identifies highly influential elements of structure by measuring the correlation between structural and behavioral changes. $Initial\,adopters$ is identified as most influential in the beginning because it defines the initial state. Shortly after, $adoption\,fraction$ and $contact\,rate$ both have the highest correlation coefficients, which agrees with the PFD procedure in that both contribute to the gain of the causal pathway associated with the reinforcing loop. In the last phase, $initial\,potential\,adopters$ has the highest correlation coefficient which also agrees with the PFD procedure, in that it restricts the system through the balancing $market\,saturation$ loop.

PFD results are now compared to Ford's behavioral method of loop deactivation. First, consider the case in which population P starts below the carrying capacity C. Loops $B1$ and $R1$ (Figure 10) are each individually deactivated once during the divergent growth phase ($t = 10$) and once again during the convergent growth phase ($t = 50$). The resulting simulated responses are shown in Figure 13.

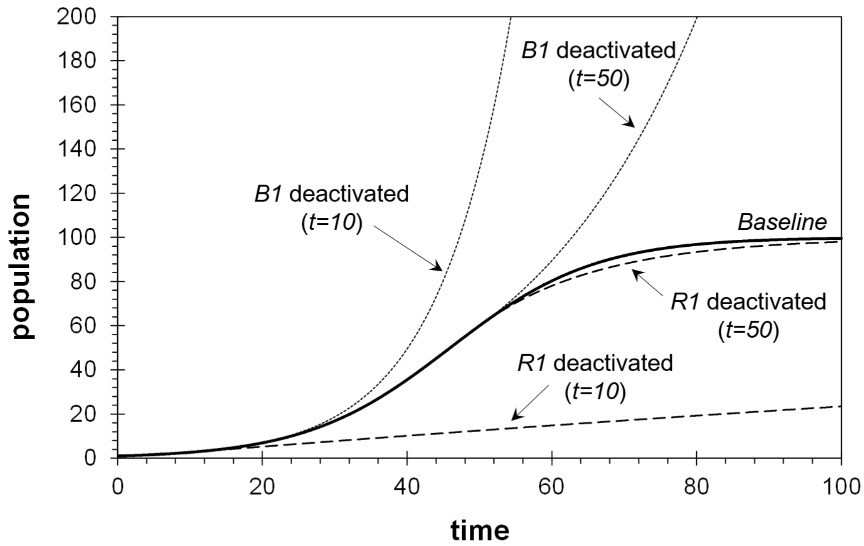

Figure 13: Ford's behavioral test for dominance applied to the logistic growth model (form 1).

During the divergent growth phase, deactivating $B1$ does not change the atomic behavior pattern (exponential), and thus, according to Ford's method, is not dominant. Deactivating $R1$ does change the atomic behavior pattern from exponential to logarithmic, and thus $R1$ is dominant. In the second phase, deactivating $B1$ changes the atomic behavior pattern from logarithmic to exponential, and thus $B1$ is dominant. Deactivating $R1$ does not change the atomic behavior pattern, and is not dominant. Because Ford's criteria for dominance employs a counterfactual test, it identifies necessary pathways. In this model, since the dominant pathways are both necessary and sufficient, and there are no pathways which are necessary but not sufficient, Ford's test identifies the same dominant structure as the PFD procedure.

Now consider the case in which $P > C$. The model exhibits only one behavior pattern (logarithmic) and loops $B1$ and $R1$ are individually deactivated at time ($t = 10$), as shown in Figure 14. Deactivating $R1$ results in no noticeable change in behavior, and thus $R1$ is not dominant. Deactivating $B1$ causes a noticeable change but does not change the atomic behavior pattern, and thus $B1$ is also not dominant. However, deactivating $R1$ and $B1$ simultaneously changes the atomic behavior pattern from logarithmic to linear, indicating what Ford calls a *shadow dominance* condition. However, because their are only two loops, deactivating both loops always results in a change in atomic behavior pattern unless the model is in equilibrium, therefore, Ford's test for shadow dominance may be degenerate in this case. None-the-less, Ford's test agrees with the PFD procedure in not identifying any dominant loops, since Ford's test identifies necessary pathways, and in this case, both loops are found by PFD to be sufficient and not necessary. This also indicates a potential relationship between *shadow dominance* and cases in which there exists one or more sufficient pathways and no necessary pathways. In examining Ford's method, it appears that a pair of shadow loops are identified when either loop is sufficient for changing the behavior, in the absence of the other, which indeed indicates a sufficient condition. He describes this as a different type of dominance. He also admits that his algorithm is more difficult as the number of shadow feedback structures increase (i.e. the number of sufficient pathways increase).

One observation is that if the two sufficient loops are considered as a set, then together they are both sufficient and necessary for creating the observed behavior, and thus can be considered as a dominant set, meeting the established criteria of dominance. This expands the definition of

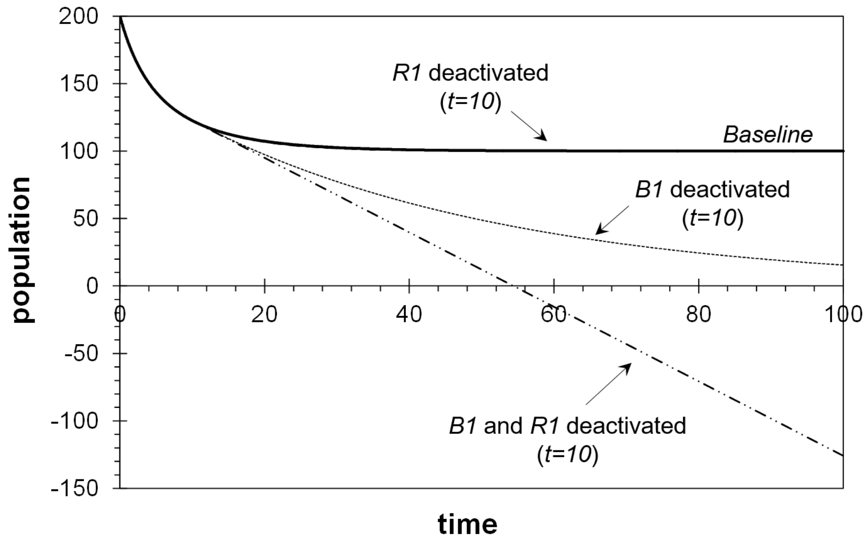

Figure 14: Ford's behavioral test for dominance applied to the logistic growth model (form 1) ($P > C$).

structure to include not just single pathways, but sets of pathways. Under this expanded definition, $B1$ and $B2$ would be identified by the PFD procedure as a dominant pair of pathways. This directly corresponds to Ford's results which identifies $B1$ and $B2$ as a pair of shadow feedback structures which are together dominant.

4.2 Logistic Model Form 2: A Second Balancing Loop (Deaths)

A three-loop version of logistic growth is analyzed and compared against the Loop Impact method [4]. This version (Figure 15), adds a second balancing feedback loop $B2$ representing population decline through deaths based on a constant fractional death rate b. Accordingly, a third pathway, p_{113}, representing balancing loop $B2$, is added to the decomposition of P:

$$\begin{aligned} \dot{P} &= p_{111} \cdot p_{112} + p_{113} \\ p_{111} &= \alpha P \\ p_{112} &= 1 - \frac{P}{C} \\ p_{113} &= -b\,P \end{aligned} \tag{16}$$

$$\begin{aligned} \ddot{P} &= F_{111} + F_{112} + F_{113} \\ F_{111} &= \left(1 - \frac{P}{C}\right) \cdot \alpha \dot{P} \\ F_{112} &= \alpha P \cdot \left(\frac{-\dot{P}}{C}\right) \\ F_{113} &= -b\,\dot{P} \end{aligned} \tag{17}$$

The simulation begins from an initial population of 1. Figure 16 shows the results of the simulation and force decomposition analysis.

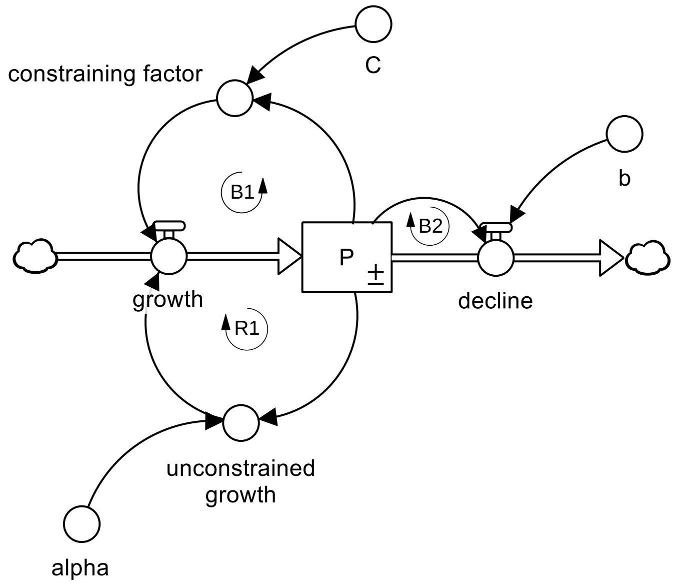

Figure 15: Stock and flow diagram of the logistic growth model (form 2).

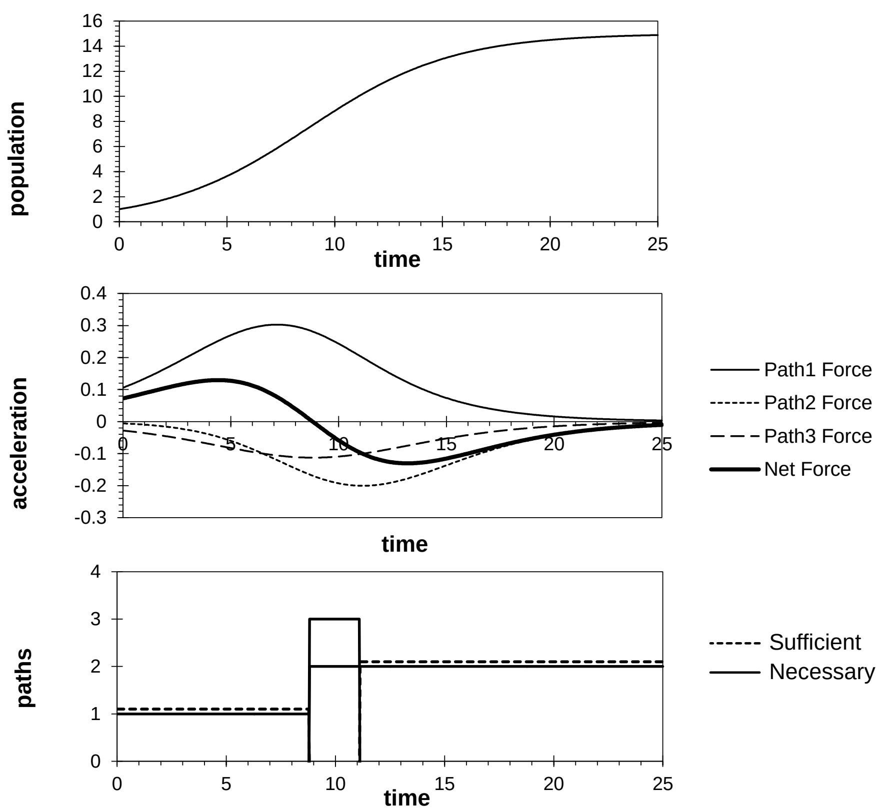

Figure 16: Simulation results for logistic growth model (form 2).

4.2.1 Results

As before, the inflection point occurs when $P = C/2$. Path 1 ($R1$) always exerts positive force, while Path 2 ($B1$) and Path 3 ($B2$) always exert negative force. There are three distinct phases of dominance. Before the inflection point (phase 1), Path 1 is necessary and sufficient (dominant).

After the inflection point, there is a brief period in which Paths 2 and 3 are both necessary, and neither sufficient, to produce deceleration. Then, Path 2 becomes necessary and sufficient and dominates for the remainder of the trajectory.

Phases 1 and 3 agree with the Loop Impact method, identifying $R1$ and $B1$ as dominant, respectively. In phase 2, the Loop Impact method identifies $B1$ and $B2$ as dominant *together*, or as a set, and concludes that in this phase there is no single dominant pathway or loop. The PFD method also concludes that in phase 2, there is no single dominant pathway (i.e. no pathway is both necessary and sufficient). However, if Path 1 and Path 2 are considered as a set, then as a set they are necessary and sufficient in phase 2, and meet the criteria for dominance. Thus, if the definition of dominance is expanded to include as elements of structure *sets of pathways*, then the results are the same as in the Loop Impact method.

The Loop Impact method, by construction, finds the minimum combination of loops of like polarity (i.e forces contributing in same direction) whose combined impact is greater than the sum of all loops of opposite polarity [4]. This criteria, by construction, identifies either a single sufficient loop[3], or a minimum set of loops which are sufficient. Thus, it is expected that the PFD procedure, which identifies dominant structures as those which are both necessary and sufficient, would agree with the Loop Impact method, however the Loop Impact method also classifies loops as dominant which are sufficient and not necessary, and thus identifies dominance more frequently than the PFD procedure.

Next, Ford's behavioral method is applied to this model. In Ford's procedure, locations of deactivation are based on changes in behavior patterns. Therefore, in this form of the model, deactivating in two locations is insufficient for identifying the three phases of dominance discussed earlier, which is a shortcoming of Ford's method [4]. The dominance transition from phase 2 to phase 3 is not associated with a behavior mode change. Regardless, Ford's deactivation method is applied to each loop in each of the three phases ($t = 5$, $t = 10$, $t = 15$) to see what insights are discovered. Figure 17 shows the results of Ford's procedure.

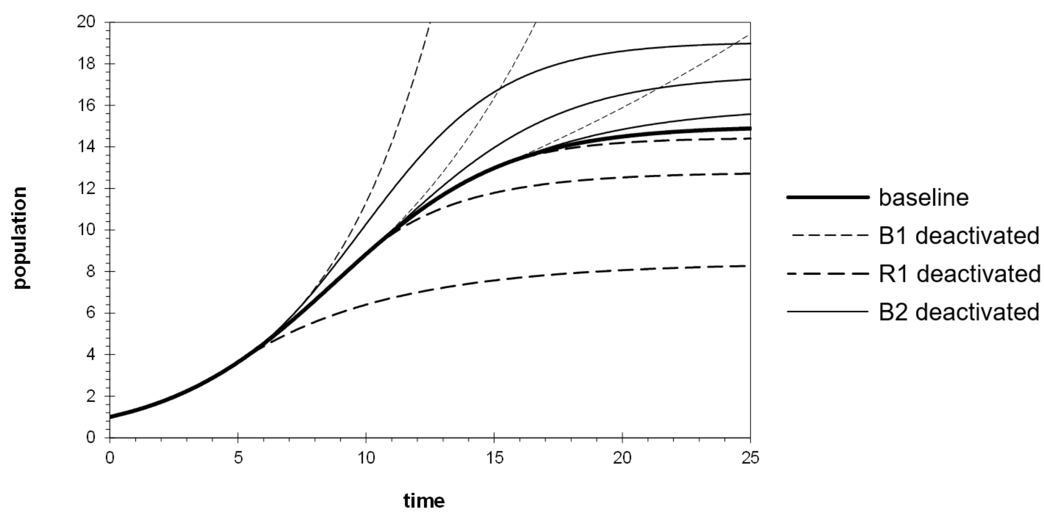

Figure 17: Ford's behavioral test for dominance applied to the logistic growth model (form 2).

[3]It is not clear how the Loop Impact method works when there are multiple single sufficient loops.

[4]Extensions to Ford's method include automated testing at every point along the trajectory and not just at behavior mode transition points, and thus address this shortfall in Ford's method [9].

In phase 1, deactivation of $R1$ is the only case which causes an immediate behavior mode change (exponential to logarithmic), and thus $R1$ is dominant. In phase 2, deactivating $B1$ results in a behavior mode change, as does deactivating $B2$, so both are found to be individually dominant. Ford refers to this as *simultaneous multiple loop dominance*. In phase 3, deactivation of $B1$ is the only case resulting in a behavior mode change, and thus $B1$ is dominant.

Ford's behavioral method agrees with PFD and the Loop Impact method in phases 1 and 3. In phase 2, it differs from both PFD and Loop Impact in that it identifies $B1$ and $B2$ as individually dominant. This is because, as observed earlier, Ford's criteria for dominance is a necessary condition (not sufficient), thus the PFD procedure explains why Ford's method produces a different result. Any time loops are necessary (regardless of whether or not they are also sufficient), Ford's method will identify them as dominant, which leads to the potential identification of multiple dominant loops.

4.3 Logistic Model Form 3: Alternate Two Loop Version

Figure 18 shows an alternative two-pathway representation of the logistic model (compare with form 1, Figure 10). Whereas in form 1, pathways represent the reinforcing growth loop and the

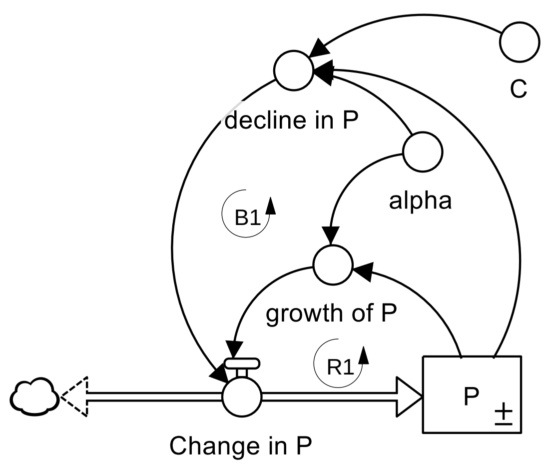

Figure 18: Stock and flow diagram of logistic growth (form 3).

loop constraining the fractional growth rate, in this form, pathways represent the reinforcing growth loop and the balancing death loop:

$$\begin{aligned} \dot{P} &= p_{111} + p_{112} \\ p_{111} &= \alpha\, P \\ p_{112} &= \left(\frac{-\alpha}{C}\right) \cdot P^2 \end{aligned} \tag{18}$$

$$\begin{aligned} \ddot{P} &= F_{111} + F_{112} \\ F_{111} &= \alpha\, \dot{P} \\ F_{112} &= \frac{-2\,\alpha\, P\, \dot{P}}{C} \end{aligned} \tag{19}$$

Note that Equations 18 and 19 are equivalent to Equations 14 and 15, respectively, but have different pathway decompositions. Both forms are often found in literature on sigmoid growth, as observed earlier. For $P < C$, the simulation and analysis results are shown in Figure 19.

4.3.1 Results

The results are the same as in form 1 ($P < C$), in which Path 1 ($R1$) is necessary and sufficient (dominant) in the first phase of exponential growth, and Path 2 ($B1$) is necessary and sufficient (dominant) for logarithmic growth in the second phase.

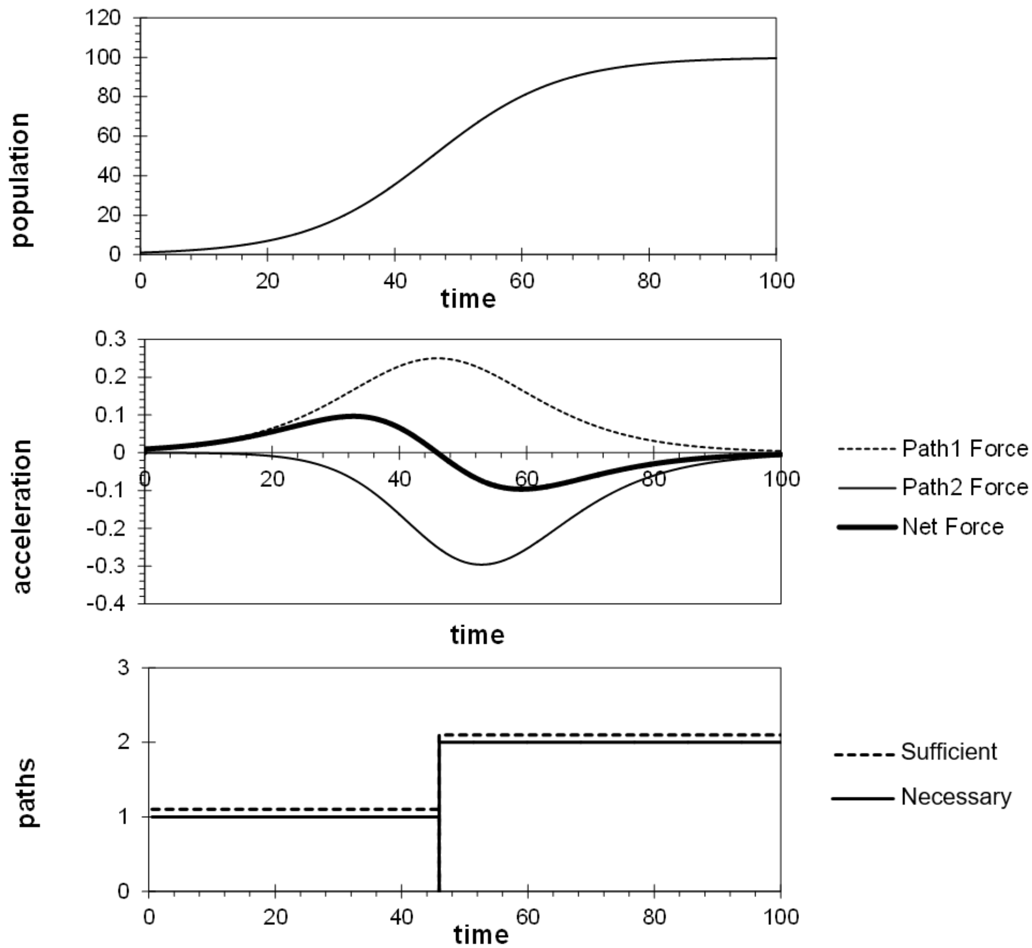

Figure 19: Simulation results of logistic growth (form 3).

Next, the results for condition ($P > C$) are shown in Figure 20.

These results differ from form 1 (compare with Figure 12). Whereas in form 1, both paths were sufficient for producing the exponential decay behavior, in form 3, the balancing loop (path 2) is necessary and sufficient for producing the behavior while the reinforcing loop (path 1) has an opposing force. In form 1, it was noted that the balancing loop changed the polarity of the reinforcing loop to behave like a balancing loop, so both were sufficient for generating the behavior. In form 3, the balancing and reinforcing loops are added together instead of multiplied, and therefore always represent positive and negative forces strictly associated with births and deaths. The necessary and sufficient pathways for each case are summarized in Tables 3 and 4.

In summary, while the pathway decomposition does not change the dynamic behavior of the

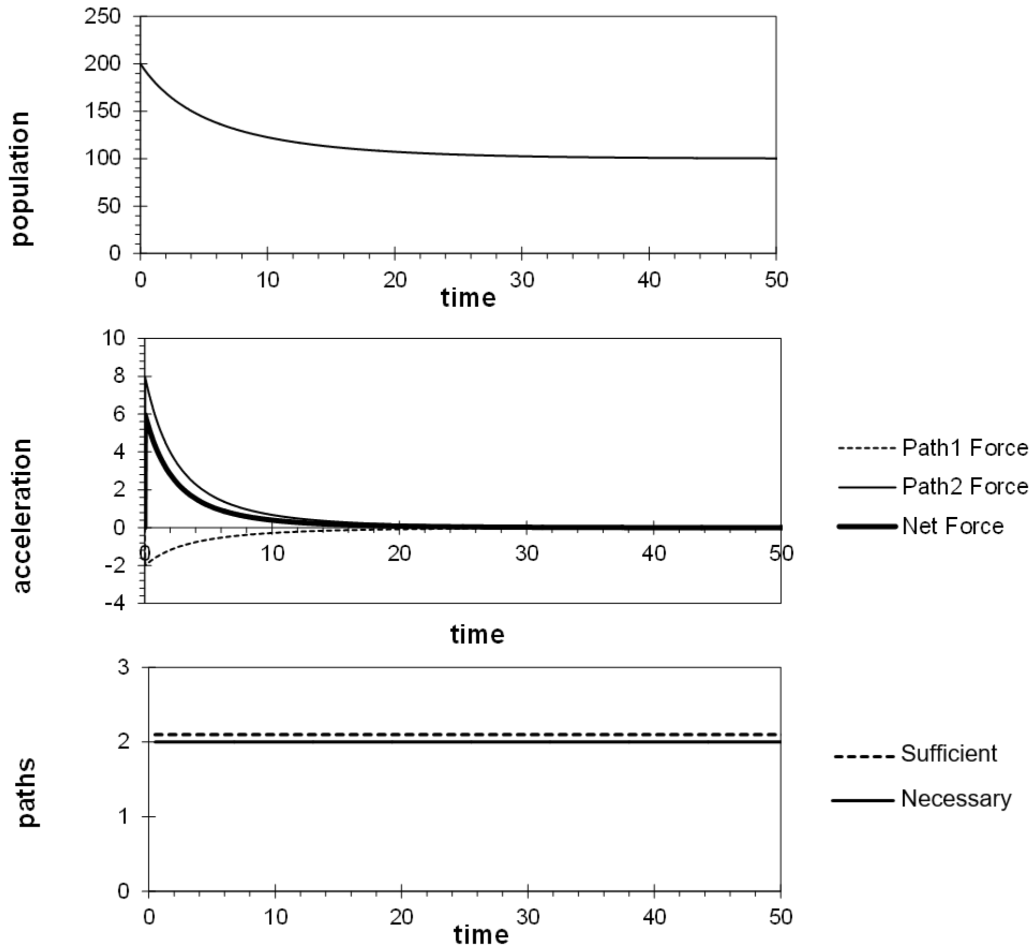

Figure 20: Simulation results of logistic growth (form 3) ($P > C$).

Loop	$\mathbf{P < \frac{C}{2}}$	$\mathbf{\frac{C}{2} < P < C}$	$\mathbf{P > C}$
R1	N & S		S
B1		N & S	S

Table 3: Dominance of loops in logistic equation form 1.

Loop	$\mathbf{P < \frac{C}{2}}$	$\mathbf{\frac{C}{2} < P < C}$	$\mathbf{P > C}$
R1	N & S		
B1		N & S	N & S

Table 4: Dominance of loops in logistic equation form 3.

system, this example illustrates how pathway definition and decomposition affects the explanation for how structure determines behavior, and the nature of dominant structure.

5 Conclusions

5.1 Procedure for Identifying Dominant Structure

Using the proposed definitions for *behavior*, *structure*, and *dominance*, behavior is formally expressed as a sum of contributions from individual elements of structure (pathways). There may be more than one way to decompose behavior, depending on how causal mechanisms are defined. The dominance criteria requires deactivation of pathways to identify necessary and sufficient conditions and it was shown that examining the second derivative is sufficient for detecting dominance and shifts in dominance.

Furthermore, it was shown that immediate pathways (as opposed to feedback loops) are adequate for identifying necessary and sufficient structure. Concise mathematical tests were proposed for detecting necessary and sufficient pathways. A procedure, pathway force decomposition (PFD), was developed for identifying necessary and sufficient pathways and was applied to several forms of the logistic model. The PFD procedure has qualities similar to exploratory/behavioral methods of dominance analysis in that the criteria for dominance is anchored in a behavioral criteria (sign change of the second derivative).

The PFD procedure also has qualities similar to formal/structural methods in which the criteria for dominance is developed from the structure (equations) of the model. Thus, the PFD procedure may be considered both a behavioral and a formal/structural dominance method. One reason for this is that the criteria for dominance has behavioral/objective qualities in that dominance requires an objective behavior change, as well as structure-relative qualities in that the metric of force contribution can be directly compared for different pathways. Therefore, the PFD procedure captures both the behavioral-objective dimension of dominance as well as the structural-relative dimension.

When tested against the logistic growth model, the PFD procedure showed that the loops commonly associated as dominant satisfied both necessary and sufficient conditions, providing further support for the proposed definition of dominance. The procedure also illustrated how pathway choice affects structure-behavior explanations.

5.2 Relationship Between Dominance Methods

All methods of dominance analysis have been applied to the logistic growth model and produce consistent results, but for different reasons.

Ford's behavioral test employs a counterfactual criteria and identifies necessary structure as dominant. Conversely, the Loop Impact method, by construction, identifies sufficient elements of structure as dominant. Where the structure is both sufficient and necessary, Ford's behavioral method, the Loop Impact method, and PFD identify the same dominant structure. Where structure is necessary but not sufficient, or sufficient but not necessary, the methods will not identify the same dominant structure. PFD applies Ford's behavioral criterion analytically when it identifies necessary structure, and thus can also be viewed as an automated, analytical version of Ford's method. PFD also improves upon Ford's method in that it directly identifies multiple shadow feedback structures analytically. PFD is also easily applied at each time step along a trajectory and does not require manual deactivation of feedback structure. This is important because, as observed, not all shifts in dominance are associated with changes in behavior modes. Additionally, manual deactivation and testing of models can be a tedious exercise and clutter up the model with switch variables. In Ford's method, deactivation points must be manually identified, and multiple combinations of loops deactivated simultaneously in the case of shadow structure. By using a completely analytical procedure, these challenges are avoided.

Likewise, the Loop Impact method has many similarities to PFD, but uses the Impact metric which is the acceleration contribution divided by the first derivative of the variable of interest. Thus, it uses the same metric as the PPM method.

PPM results agree with PFD and the behavioral methods when the dominant loop also happens to have the largest TPPM contribution, as in the case of the logistics model. In fact, when there only exists a single loop contributing in the direction of the observed behavior, PPM will always agree with the behavioral method and the PFD. TPPM is simply the second derivative divided by the first derivative, and thus is very similar to the proposed definition of behavior. However, because PPM relies on a relative metric for determining dominance, it will always identify a single dominant loop. Whenever there is a change of loop with the largest TPPM, a shift in loop dominance will be identified, thus it is expected that PPM identifies dominance and shifts in dominance more frequently than behavioral-based methods such as Ford's procedure and the PFD procedure. The PFD procedure improves upon the PPM method in that it precisely determines each pathway's contribution to the observed behavior, as opposed to a normalized proxy measure of behavior. The PFD procedure also allows for the possibility that multiple loops or no loops dominate, and it is well-defined when the derivative of the variable of interest is equal to zero.

For similar reasons as for PPM, LEEA (also based on a normalized influence metric), agrees with the results of PFD and the behavioral methods in the case of the logistic model since there is only one reinforcing and one balancing loop. Like PPM, LEEA will also always identify a single dominant loop and detects shifts in dominance more frequently than in behavioral methods. One way to interpret the results of PPM and LEEA is that they identify highly influential structures which are contributory and *potentially* dominant, and are good candidates for subsequent testing for dominance [3].

Ford's behavioral method identifies cases of *shadow loop dominance*, which others have also described as *shared dominance* [9]. It also describes situations of *multiple loops dominating simultaneously*. The phenomenon of *shadow loop dominance* is associated with multiple sufficient and unnecessary pathways, and the phenomenon of *multiple loop dominance* is associated with multiple necessary and insufficient pathways.

5.3 Implications of Dominance Framework for Policy Design

5.3.1 Expanded Definition of Dominance

Ford's method considers pairs of shadow feedback structures (i.e. sets of sufficient pathways/loops) as dominant together. Similarly, the Loop Impact method considers sets of pathways/loops which are collectively sufficient to be a dominant set. One observation is that multiple sufficient pathways may form a set that is both sufficient and necessary. Likewise, multiple necessary pathways may form a set that is both necessary and sufficient.

Expanding the definition of dominance to consider not just single pathways which are necessary and sufficient, but sets of pathways which together are necessary and sufficient, permits a broader use of the term.

Therefore, the following modified definition is proposed:

> Given a state variable $x_j(t)$ whose behavior is of interest, and point t_0 along its trajectory, a set of causal pathways are *dominant* if and only if the set is both necessary and sufficient for determining $\operatorname{sgn} \ddot{x}_j(t_0)$.

5.3.2 System Robustness and Fragility

Ford's method and the Loop Impact method above illustrate three possible compositions of dominant sets:

1. A dominant set contains a single necessary and sufficient pathway.
2. A dominant set contains some combination of necessary and contributory pathways, but no sufficient pathways.
3. A dominant set contains some combination of sufficient and contributory pathways, but no necessary pathways.

The robustness or fragility of a variable's behavior at a given time depends on the number of necessary and sufficient pathways to that variable at that time. Consider Figure 21, which expands upon the dominance framework introduced earlier.

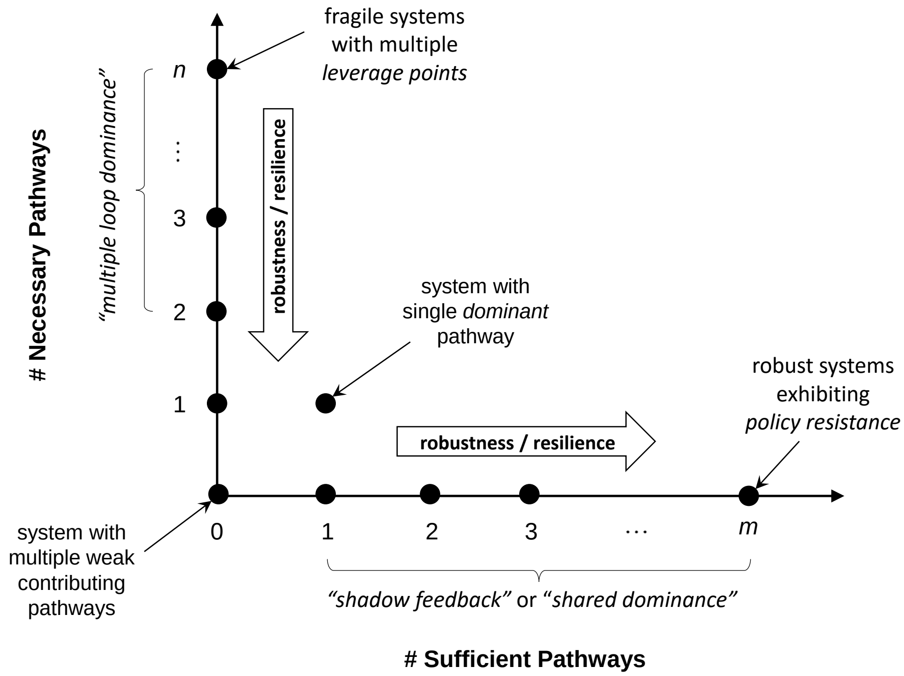

Figure 21: System robustness depending on the number of necessary and sufficient pathways.

Figure 21 shows the typical case of *dominance* associated with a single necessary and sufficient pathway. It also shows that *shadow feedback* or *shared dominance* occurs when there exists one or more sufficient and no necessary pathways. Similarly, *multiple loop dominance* or *simultaneous dominance* occurs when there exists two or more necessary and no sufficient pathways. A system with no sufficient or necessary pathways contains multiple contributing pathways. The robustness or resilience of a system increases as the number of necessary pathways decrease and as the number of sufficient pathways increase. Here, the terms *robustness* or *resilience* are used to describe the likelihood of system behavior change when there are changes in causal pathways.

5.3.3 Design Implications for Robust Systems

General principles for making a system more or less robust can be inferred from the dominance framework. First, robustness can be increased by decreasing the necessary pathways. This is accomplished by adding new supporting causal pathways, increasing the force of the weaker contributing pathways, or by reducing the force of the opposing pathways. Second, robustness can be increased by increasing the sufficient pathways. This is accomplished by increasing the number of individual pathways which are stronger than the opposing pathways, either by increasing the force of supporting pathways, adding new strong supporting pathways, or reducing the strength of opposing pathways.

This also raises new questions of how to identify the best way to intervene in a system that is operating in one of these six states. Also, can these insights be used to develop design heuristics for sustainable policies and interventions? Can they be used to better understand policy resistance observed in real systems?

6 Summary

A formal definition of *dominance* is proposed and tested against a simple model alongside other methods. In the process, some discrepancies between current methods were addressed, and ambiguities associated with shadow loop dominance, multiple loop dominance, and pathway/loop representation were also addressed. Additionally, based on the tests, the definition of dominance was expanded to include not just single pathways, but sets of pathways. Based on the proposed definition of dominance and the PFD procedure, new insights were offered regarding the relationship between dominance and system robustness, policy resistance, and leverage points.

Additional analyses and comparison between PFD and other loop dominance methods on larger and more complex models is warranted.

References

[1] Terri Duhon. (D-4286) A system dynamics approach to rabbits (S-shaped growth). 1992.

[2] David N. Ford. A behavioral approach to feedback loop dominance analysis. *System Dynamics Review*, 15(1):3–36, March 1999.

[3] Nathan B Forrester. *A dynamic synthesis of basic macroeconomic theory: implications for stabilization policy analysis*. PhD thesis, Massachusetts Institute of Technology, 1982.

[4] John Hayward and Graeme P. Boswell. Model behaviour and the concept of loop impact: A practical method. *System Dynamics Review*, 30(1-2):29–57, January 2014.

[5] J. Huang, E. Howley, and J. Duggan. An extension of loop deactivation in the behavioral method. In *Proceedings of the 27th International Conference of the System Dynamics Society*, Seoul, South Korea, 2010.

[6] Christian Erik Kampmann and Rogelio Oliva. Analytical methods for structural dominance analysis in system dynamics: An assessment of the current state of affairs. In *Proceedings of the 25th International Conference of the System Dynamics Society, Boston, USA*, 2007.

[7] Leslie A Martin. (D-4476-2) Exploring S-shaped growth. 1996.

[8] Mohammad Mojtahedzadeh. Consistency in explaining model behavior based on its feedback structure. *System Dynamics Review*, 27(4):358–373, October 2011.

[9] H.G. Phaff. Generalized loop deactivation method: An extension of fords behavioural approach to loop dominance analysis and its applications. In *Proceedings of the 26th International Conference of the System Dynamics Society*, Athens, Greece, 2008.

[10] Hazhir Rahmandad. Connecting strategy and system dynamics: an example and lessons learned. *System Dynamics Review*, 31(3):149–172, 2015.

[11] George P. Richardson. Loop polarity, loop dominance, and the concept of dominant polarity (1984). *System Dynamics Review*, 11(1):67–88, March 1995.

[12] John D Sterman. *Business Dynamics: Systems Thinking and Modeling for a Complex World.* Irwin McGraw-Hill, New York, 2000.

[13] Timothy R. B. Taylor, David N. Ford, and Andrew Ford. Improving model understanding using statistical screening. *System Dynamics Review*, 26(1):73–87, January 2010.