Estimating System Dynamics Models Using Indirect Inference

Abstract

System Dynamics research has not reached its potential to impact many social science fields, partly because it is difficult to estimate SD models using common datasets which include few data points over time for many units under analysis. Here, we introduce indirect inference, a simulation-based estimation method which can be applied to common data structures and is applicable to SD models which often include intractable likelihood functions. In this method, the parameters of the model are estimated in a way that simulated data and empirical data produce similar statistics. We also present a case study in the context of depression research where we apply the method, estimate the unknown parameters and their confidence intervals, and assess the model's fit to an empirical dataset. The overall results suggest that indirect inference can extend the application of SD models to new application areas and leverage common panel datasets to provide unique insights.

Keywords: indirect inference; simulation-based estimation methods; system dynamics; Depression

Background

Most system dynamics (SD) models use a single case study and apply traditional estimation methods (e.g., mean squared error, mean absolute percentage error, etc.) to time series data for that case to specify unknown parameter values. However, more flexible methods of estimation are needed in both theoretical and practical applications to leverage data structures beyond single case time series. With increasing availability of datasets on various research subjects, from individual level to firm and country level phenomena, formal model calibration has become a requisite step in producing credible model-based analyses that is trusted by various academic audiences. However, there are three major challenges in estimating SD models. First, SD models are often complex and nonlinear and the likelihood functions are intractable. Thus many conventional statistical methods do not directly apply. Second, due to the structure of many datasets, even heuristic calibration methods common in the SD practice that minimize the differences between empirical time series and simulated counterparts may not apply. For example, many "panel" datasets include data at only a few points in time, but for many units under analysis (e.g., many individuals, organizations, or countries), complicating the matching of the simulations to data using traditional methods that require many data points over time for each unit. For the same reason, other methods such as kalman filtering (Kalman, 1960) or extended kalman filtering (G. L. Smith, Schmidt, McGee, Aeronautics, & Administration, 1962) which adjust state variables based on measured system behaviors cannot be used effectively when very few data points are available over time. Third, in many applications, randomness which is exogenous to the model boundaries has a significant role in the behavior of the system; therefore, noise should be considered explicitly in the estimation of the model. These complications call for the introduction of more rigorous simulation-based estimation methods to the SD literature.

The simulation-based estimation methods were introduced with the increasing computational power of computers that made it possible to run many numerical simulations of large datasets in short periods of time. The basic idea behind these methods is to match properties of the simulated data to those of the empirical data. These methods include method of simulated moments (Duffie & Singleton, 1993; Jalali, Rahmandad, & Ghoddusi, 2013; Mcfadden, 1989; Pakes & Pollard, 1989), efficient method of moments (Durlauf & Blume, 2008), and indirect inference (Gourieroux, Monfort, & Renault, 1993; Gouriéroux, Phillips, & Yu, 2010; A. A. Smith, 1993) to name a few. These methods are mostly useful for models with intractable likelihood function such as nonlinear dynamic models and models with missing or incomplete data.

In this article, we provide an introduction to one of the most flexible methods available in this space, and present how it can be applied in SD modeling. First, we introduce the indirect inference method and explain the steps to estimate unknown parameters of a model. We then present an SD model that relates depression, rumination and stressful life events to demonstrate the estimation of this model using indirect inference. Finally, we conclude and discuss under what conditions SD studies can benefit from the indirect inference.

Indirect Inference Method

General properties and historical background. The main idea behind the indirect inference method is to match properties of empirical and simulated data in order to estimate the unknown parameters of the model of interest. This method was developed to overcome the challenges of estimating parameters of complex models for which the likelihood function is intractable. In indirect inference method, the simulated data is generated by simulating the model of interest, then an "auxiliary model", typically consisting of simple regression (s), is selected and parameters of the auxiliary model are estimated by using both the empirical and the simulated data. The difference between these two sets of parameters of the auxiliary model is minimized to estimate the parameters of the model of interest.

The indirect inference method has several advantages. First, there are few limitations to the types of models to which it can be applied. The only requirement is that the model of interest can be simulated for different values of its parameters. Second, although this method is a simulation-based technique, it can be relatively inexpensive to compute when the auxiliary model uses a maximum likelihood estimator, and thus the auxiliary model parameters have small variance and could be matched reliability with few simulations (Gourieroux et al. 2010). Third, the indirect inference method inherits the beneficial properties of the estimation method used for the auxiliary model (Gourieroux et al. 2010). For instance, if the maximum likelihood is used to find the parameters of the auxiliary model, the estimated parameters resulting from indirect inference would also have small variance. Forth, it can be used for both estimating and validating a model. The validation step allows the modeler to decide if the model's outputs are indistinguishable from empirical data, or notable differences exist after estimation which could inform further model refinement. In this article, we discuss one such validation test as well. We also investigate the method's validity using a separate approach where indirect inference is applied to a synthetic dataset generated by simulation of the calibrated model, and method's ability to recover correct parameters (from a structurally precise model) is evaluated.

The method of simulated moments (MSM) proposed by Mcfadden (1989) is one of the first rigorous simulation-based estimation methods, which is the workhorse of modern econometrics, and motivates the idea of indirect inference. In this method, parameters of a model are estimated by minimizing the difference between selected moments (e.g., mean and variance) of empirical data and corresponding moments of model-generated simulated data. There are only a few studies that have implemented MSM to calibrate SD models. Rahmandad and Sabounchi (2011) calibrate a dynamic model of obesity at both individual and population levels by using MSM and Jalali et al. (2013) discuss the application of MSM to SD models. The indirect inference method, proposed independently by Gourieroux et al. (1993) and A. A. Smith (1993), is very similar to MSM in matching some functions of empirical data against the same function calculated on simulated data. However, it is more general because rather than only the statistical moments, a wider set of functions of the empirical and simulated data can be matched to estimate the unknown parameters. These functions are created using auxiliary models. The auxiliary model is typically a separate estimation, but does not need to capture the true data generating process. The auxiliary model only serves as a lens through which we view the empirical data and calculate functions which we will then match against their simulated counterparts. The parameters of the model are set in a way that both empirical and simulated data produce very similar images as they pass through this lens.

There are other methods that follow a similar logic. Structural equation modeling (SEM) is based on matching the observed covariance matrix and model-generated covariance matrix (Anderson & Gerbing, 1988). In actor-based network models, the statistical properties (such as degree distribution, centrality, and clustering) of empirical networks are compared with those of the simulated networks to estimate the parameters of a model (Snijders, 2001). Overall, indirect inference and its derivatives are among the most flexible econometric methods for estimating complex models using various data structures. Given its moderate computational costs, the method could be applied easily to models of modest size and when a handful of model parameters are to be estimated. However, estimating a large number of parameters (e.g. in the hundreds) could be much more challenging because the underlying optimization problem is nonconvex. There is currently no study in the literature that applies the indirect inference method to SD models. In the next section, we introduce the method formally and present a step-by-step guide for applying it.

The Description of the Method. Consider a general dynamic model with stock (state) variables z, dynamics of which are described as $\frac{dz}{dt} = f(\theta_1, z, u, \varepsilon_1)$ and a set of exogenous variables, *u*, and observable variables, x, which are a function of z:

$$x = g(\theta_2, z, u, \varepsilon_2) \tag{1}$$

Here function *f* describes the dynamics of the system and function *g* the measurement process and structure of both of these functions are assumed to be known. A vector of random errors with a known distribution[1] ($\varepsilon = \cup(\varepsilon_2, \varepsilon_1)$) adds uncertainty to the dynamics and measurements. Finally, a set of parameters, $\theta = \cup(\theta_2, \theta_1)$, is unknown and the goal of the estimation process is to find these parameters. Note that the model and measurement functions may apply to a single case or multiple units of the phenomenon of interest. For example a panel dataset includes measures on dynamics of several parallel units (e.g., people, firms, or countries) over time. Figure 1 summarizes the steps to estimate the model parameters (θ) by using the indirect inference method.

First, suitable statistics of empirical data, *x*, are generated. Suitable statistics include coefficients of an auxiliary model (e.g., a regression that estimates some elements of x based on other elements or lagged values) or they can be any statistics of a dataset such as mean and standard deviation (Wood, 2010). We call them empirical-auxiliary statistics (S_{EmpAux}). Second, the corresponding simulated statistics, simulated-auxiliary statistics (S_{SimAux}), are calculated and estimated. For a given value of θ, the model of interest (SD model) is simulated H times by using H different streams of noise over time, ε_t ($=\varepsilon_1^h, \ldots, \varepsilon_T^h$), h=1,…H. As a result, H number of *x* are generated. Then, S_{SimAux} is estimated for each *x*. Third, the average of these estimators are found ($\frac{1}{H}\sum_{h=1}^{H} S_{SimAux}{}^{h}$) and θ is changed to minimize the difference between the empirical-auxiliary statistics and the average of simulated-auxiliary statistics (Gourieroux et al., 1993). Each step is explained in more details in the following sections.

---

[1] The distribution of the ε does not need to be known. ε can be a function of a white noise with a known distribution and an unknown parameter of the model of interest (θ) (Gourieroux et al. 1993). Moreover, if there is uncertainty in the initial conditions of stock variables, that uncertainty could be incorporated into the ε.

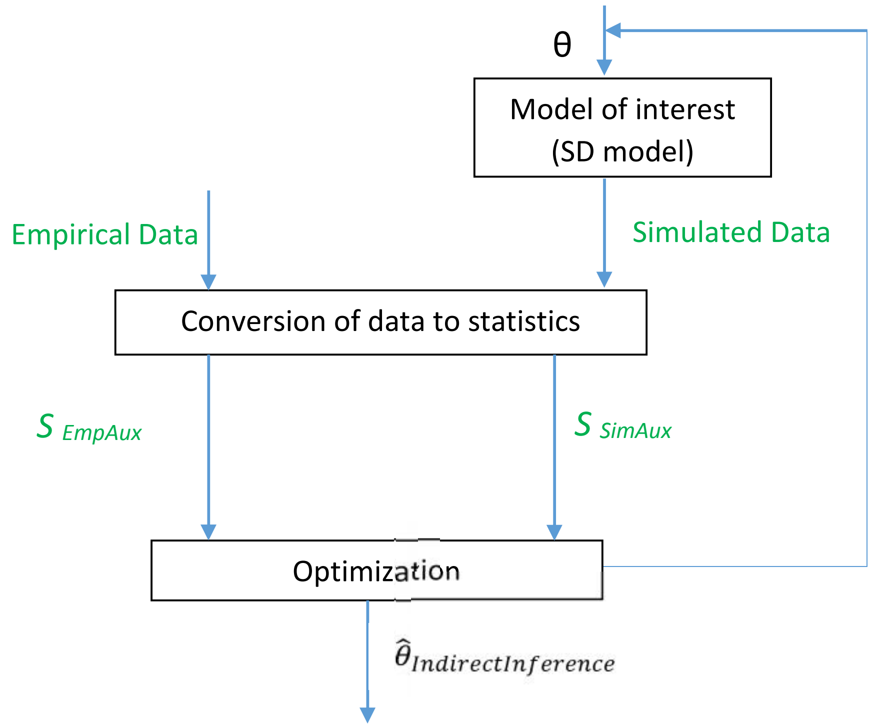

Figure 1. Required steps for estimating parameters of a dynamic model

1. **Define and estimate a set of empirical-auxiliary statistics.** The first step is to select a set of empirical-auxiliary statistics, which when matched in simulation, allow us to find the model parameters. There is substantial flexibility in terms of defining these statistics. Common empirical statistics include mean, standard deviation, autocorrelation, and correlation matrixes of observed variables. These statistics are typically calculated across different units of analysis (for cross sectional and panel data, e.g. mean weight in a group of subjects), but could also be calculated over time for a single case, if we can assume the observations are coming from a stationary system (i.e., a system in stochastic equilibrium). Besides simple statistics, more complex auxiliary models could be defined that relate some of the observed variables to the other ones, or to the lagged values of the same variable. The coefficients of these models (i.e., regression coefficients) could be then appropriate statistics to include in our empirical-auxiliary statistics vector. Note that

auxiliary models do not need to be accurate (i.e., the density function may not accurately describe the conditional distribution of *x* for the element of *x* being estimated (Durlauf & Blume, 2008)). It is an approximate model which, unlike the model of interest, can be easily estimated with limited computational costs (e.g., using a simple linear regression). However, the estimation would be more efficient if the auxiliary model were defined as precisely as possible, i.e. the auxiliary model is a good approximation of some aspects of *f* and *g* functions which are reflected in the estimated relationship (Guvenen & Smith, 2010). A more precise model reduces the variance of estimated regression coefficients (elements of S_{EmpAux}) and thus enables a reliable estimation with a smaller number of simulations, H.

A couple of examples help to illustrate the idea of auxiliary models. Consider a dynamic, stochastic, general equilibrium (DSGE) model of the macroeconomy which describes the trend of a macroeconomic variable such as consumption (x_t). This model could be quite complex and hard to estimate directly. An auxiliary model for estimating the more complex DSGE can be a vector autoregression for the variable of interest: $x_t = \beta x_{t-1} + \varepsilon_t$, (A. A. Smith, 1993). Another example is a two-level logistic model, $x_{ik} = logit^{-1}(p_{ik}) + e_{ik}$, in which $logit(p_{ik}) = \theta_0 + \theta_1 z_{ik} + u_k$ and x_{ik} is the i^{th} observation in the k^{th} group. This model has intractable likelihood function and conventional estimation methods cannot estimate it. An auxiliary model for implementing indirect inference can be $x_{ik}^* = \beta_0 + \beta 1 z_{ik} + u_k + e_{ik}^*$, (Mealli & Rampichini, 1999).

A good empirical-auxiliary statistic has four key characteristics. First, it should be relatively stable, that is, its value should not be very sensitive to the process and measurement noise streams (ε). The empirical value of a noise-sensitive statistic is not reliable and as such does not have much information to guide the identification of model parameters. Second, good statistics are sensitive to changes in at least one of the parameters in θ. In the extreme, if a statistic does not change with changes in any of the model parameters, there is no way to backtrack the value of any parameter based on the information in that statistic. Both of these conditions could be partially tested using simulations. One can simulate the model in the range of parameters being considered and measure the sensitivity of the simulated statistics with respect to model parameters ($\frac{\partial S_{SimAux}}{\partial \theta}$) and their sensitivity to different noise streams. Third, empirical statistics should be inexpensive to calculate, or otherwise the multiple iterations needed to solve the optimization problem may become infeasible. Therefore simple linear regressions are preferred over regression models that require non-convex optimizations. Forth, the number of statistics should be equal or more than the number of the parameters that need to be estimated. In other words, $q \geq p$ where p and q are the number of elements in the vector θ and the statistics vector S_{EmpAux}, respectively. After choosing appropriate statistics, including the auxiliary model(s), the empirical-auxiliary statistics (S_{EmpAux}) are estimated or calculated using the empirical data *x*.

2. **Generate the simulated data using the SD model.** First, H independent drawings of ε_t ($\varepsilon_1^h, \ldots, \varepsilon_T^h$) are generated. These streams of random numbers are generated only once and kept constant in the rest of the process. Then for a given θ, the SD model is simulated H times (H replications using the independent drawings above). This process creates the simulated data which contains H paths ($x_1^h, \ldots, x_T^h$) where h=1,…,H. The number of observations in each path should be equal to the number of observations in the empirical

data. It should also be noted that the same ε_t $(= \varepsilon_1^h, \ldots, \varepsilon_T^h)$ are used for each value of θ simulated during optimization (i.e., we use the same noise seed values for each iteration of the optimization).

3. **Estimate the simulated-auxiliary statistics using the auxiliary model and simulated data.** For each of the H paths, the simulated-auxiliary statistics are estimated in the same fashion they were calculated for the empirical-auxiliary statistics. The only difference is that instead of using empirical data, simulated data are used to estimate those statistics. The key point is to generate the same statistics using the empirical and simulated data (they are both $< q \times 1 >$ vectors). After finding the simulated statistics for each path, the average of these H simulated-auxiliary parameters is found as:

$$\left(\frac{1}{H}\sum_{h=1}^{H} S_{SimAux}{}^{h}\right) \quad (2)$$

Typical values of H could range between a handful and hundreds, depending on the variance of the simulated auxiliary statistics. If that variance is high, a larger H is recommended to reduce error that is due to simulation of statistics. However, note that computational costs scale linearly with H and incremental value of increasing H is limited, because for the empirical statistics we only have a single path available, and thus the total sampling error approximately scales with (1+1/H).

4. **Minimize the difference between the auxiliary-empirical statistics and the auxiliary-simulated statistics.** The unknown parameters (θ) are estimated by minimizing the weighted differences between the empirical-auxiliary statistics (S_{EmpAux}) and the average of the simulated-auxiliary statistics ($\frac{1}{H}\sum_{h=1}^{H} S_{SimAux}{}^{h}$). In other words, the parameters of the model of interest are estimated as:

$$\hat{\theta} = argminMin\left(S_{EmpAux} - \frac{1}{H}\sum_{h=1}^{H} S_{SimAux}{}^{h}\right)' W \left(S_{EmpAux} - \frac{1}{H}\sum_{h=1}^{H} S_{SimAux}{}^{h}\right)$$

(3)

Weighting matrix (W) can be any positive definite matrix in theory, but good choices of W are critical for getting reliable parameter estimates. Therefore, the calibration of the indirect inference is usually performed in a two-step procedure with two different values of W. In the first step, W can be chosen to be a diagonal matrix in which the diagonal element i of the matrix is the inverse of square of the i^{th} empirical statistic ($1/S_{EmpAux\_i}{}^2$) and the non-diagonal elements are zero (let us call this matrix W_1). W_1 ensures that some statistics do not dominate the optimization if their magnitude is much larger than the others. However, W_1 is not theoretically optimal in the sense that it does not provide the lowest standard deviation for the estimated parameters. After performing optimization using the initial W1 and getting estimates of the model parameters ($\widehat{\theta_1}$), we switch to another W, the inverse of the variance-covariance matrix of the simulated statistics (using $\widehat{\theta_1}$ to estimate this matrix) and repeat the estimation process. It is important to note that in calculating the variance-covariance matrix a large number of simulations, using distinct noise streams, will be needed (in thousands). However this step is done only once and not repeated during optimizations, so computational costs are not a concern. The intuition behind using the inverse of variance-covariance matrix is that those statistics which have large variance (i.e., they are sensitive to the choice of random noise) should get lower weights. Although in many applications stopping after the second estimation gives reliable results, W can be re-estimated (based on $\hat{\theta}$ achieved in the second step) to estimate a new set of parameters. This process can be iterated through until the estimated parameters converge across successive iterations.

The initial assumed values for θ can impact the speed of convergence in the optimization process or trap the optimization in a local optimum. If the coefficients of the auxiliary model and the unknown parameters in the main model are similar in their meaning, the initial values for model parameters could be chosen to equal the corresponding empirical-auxiliary statistics. If they are not similar, qualitative information on the appropriate range of those parameters or rough initial estimates using a relevant estimation method can help initialize the model from a more promising point in the parameter space. Even with good initial points however, the optimization may get stuck in a local optimum, so the optimization algorithm should include multiple start points to increase the chances of finding the global optimum.

5. **Model assessment test.** When q>p, the optimal value of the objective function can be used to test how well the model has been specified. The following statistic (ξ_T) is distributed asymptotically as a chi-square with $q-p$ degrees of freedom. The null hypothesis is that the model of interest (our SD model) is not different from the true data generating process. If the test statistic is larger than the threshold for chi-square distribution with the desired precision, then we reject the null hypothesis, inferring that the model's structure could be improved further.

$$\xi_T = \frac{H}{1+H} Min\left(S_{\text{EmpAux}_t} - \frac{1}{H}\sum_{h=1}^{H} S_{\text{SimAux}_t^h}\right)' W \left(S_{\text{EmpAux}_t} - \frac{1}{H}\sum_{h=1}^{H} S_{\text{SimAux}_t^h}\right) \qquad (4)$$

An Applied Example

Here we demonstrate the use of indirect inference for estimating an SD model using a panel dataset, one of the more common scenarios in which this method can prove beneficial. Major depressive disorder (MDD) is a disabling illness that causes feeling of sadness and loss of interest. Different mechanisms including genetics, cognitive, environmental, and biological factors contribute to the disorder. To keep the applied example simple, we only focus on the impact of rumination (a cognitive factor) and stressful life events (an environmental factor) on MDD. In the following sections, we explain the data, develop our model, and show how the unknown parameters can be estimated using the indirect inference method.

Data. The dataset includes 1,065 adolescents from two middle schools in Connecticut (Michl, McLaughlin, Shepherd, & Nolen-Hoeksema, 2013). The tendency to ruminate was assessed at three points in time (T_1, T_2, *and* T_3) while the questionnaires related to MDD (The Children's Depression Inventory) and stressful life events (The Life Events Scale for Children) were completed only at T_1 and T_3. The time between the first and second assessments and the second and third assessments are four and three months, respectively. Table 1 summarizes the variables and the time of data collection.

The Life Events Scale for Children (Coddington, 1972) contains 25 instances of stressful life events and asks participants to indicate whether they have experienced any of the events in the past 6 months (e.g., "Your parents got divorced" and "You got suspended from school"). The Children's Response Style Questionnaire (Abela, Brozina, & Haigh, 2002) is composed of 25 items which measures the extent to which participants ruminate when they experience sad feelings. The Children's Depression Inventory is composed of 27 items and measures depressive symptoms in children and adolescents (Kovacs, 1992). After dropping missing values, the sample includes 661 observations. Table 1 summarizes the means of variables and the time of data collection. Fifty three percent of our subjects are female. As we have only a few data points over time, it is not feasible to estimate the unknown parameters with the conventional time series estimation methods in the SD literature.

Table 1: Means and standard deviations of variables

Variable	Time 1, T_1	Time 2, T_2	Time 3, T_3
MDD symptoms	9.48 (6.28)	-	9.78 (7.64)
Rumination	11.59 (7.52)	10.85 (7.62)	9.95 (7.95)
Stressful life events	4.96 (3.32)	-	4.20 (3.70)

Standard deviations are in parentheses.

The Depression-Rumination Model. Figure 2 depicts the proposed model, an individual-level model of MDD, which is based on the response style theory. This model is the result of multiple iterations of model building, estimation, and model refinement; however, due to limited space we only report the final model and its structure. The response style theory defines rumination as repetitive thinking about the causes and consequences of a stressor without focusing on coping strategies or engaging in problem solving (Nolen-Hoeksema, 2004). Engaging in rumination increases the duration of recalling a stressor (i.e., *memory time constant*) and thus increases the accumulation of *stressor memory*, causing an even higher level of *rumination* (loop R1). Current *rumination* is formulated as a stock adjusting with a time constant towards *indicated rumination* which is a linear function of *stressor memory* (Michl et al., 2013), current *MDD symptoms* (Nolen-Hoeksema, Stice, Wade, & Bohon, 2007), and *gender* (Nolen-Hoeksema, Larson, & Grayson, 1999). Only those stressors that are perceived negatively cause rumination and are tracked in our formulation; *stressor memory* is multiplied by a fraction to calculate the stressors contributing to rumination. In addition, higher level of *rumination* predicts more *MDD symptoms* and longer duration of depression (Nolen-Hoeksema, 1991) (loop R2). The *MDD symptoms* is the smooth of *indicated MDD* which is a function of *rumination*. Moreover, the random events outside the model boundaries affect rumination and MDD. The randomness recognizes that *indicated MDD* and *indicated rumination* are not deterministic and they vary by factors outside the model boundary based on a probability distribution; however, there is some autocorrelation in how those chance events unfold. Therefore, normally distributed pink noises are added to the *indicated rumination* and *indicated MDD*, $RumNoise \sim N(0, \sigma_r^2)$ and $DepNoise \sim N(0, \sigma_d^2)$ with correlation time ρ, respectively. All equations are presented in the online Appendix.

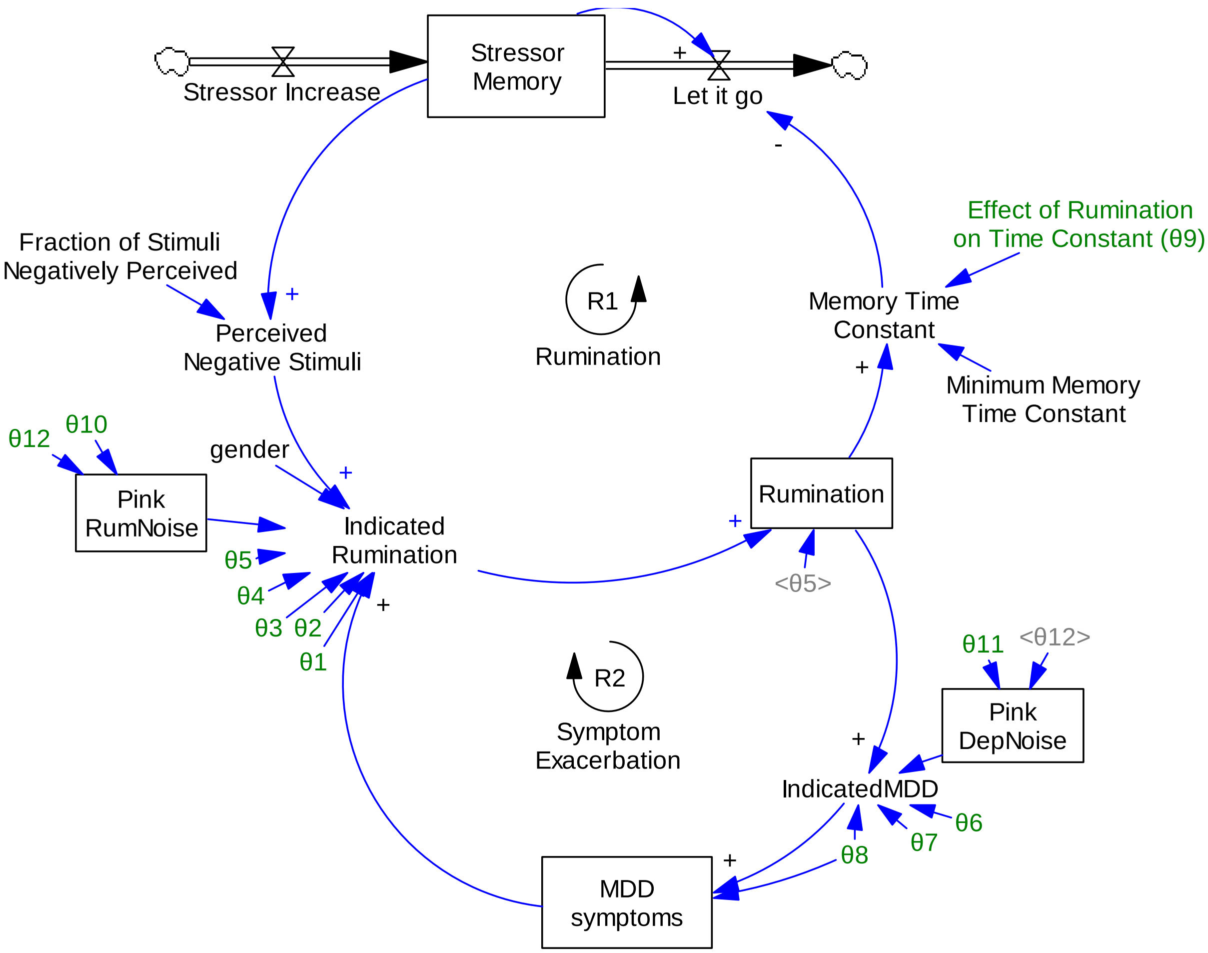

Figure 2: The MDD-Rumination model. Green parameters are the unknown parameters to be estimated.

The Stressful Life Events variable does not completely represent the stock of *stressor memory* in the model. An individual may ruminate about a stressor that happened more than or less than 6 months ago. Stressor memories are not stressful events but they are memories of those events that may cause stress when a person thinks about them. As painful experiences are forgotten, he or she reaches a point when remembering those events would not cause stress. However, our data does not capture this concept; we only observe the stressful events that occurred in the prior six months. To address this limitation as much as possible, we estimate the *stressor memory* from the Stressful Life Events questionnaire by assuming that both stocks (*Stressor Memory* and Stressful Life Events) have the same inflow and the initial value of the *Stressor Memory* is proportional to the Stressful Life Events at Time 1 (the ratio is *Memory Time Constant* divided by six months). The online appendix shows how we estimate the inflow of these stocks. The remaining twelve unknown parameters (p=12) to estimate are listed in Table 2.

Table 2: Unknown parameters in the model

Unknown Parameters (θ)	**Unit**
Rumination Constant (θ_1)	RumScore
Depression Effect Coefficient (θ_2)	RumScore/DepScore
Gender Coefficient (θ_3)	RumScore
Perceived Stress Coefficient (θ_4)	RumScore/Disruption
Rumination Coefficient (θ_5)	Dmnl*
Depression Constant (θ_6)	DepScore
Rumination Effect Coefficient (θ_7)	DepScore/RumScore
Depression Coefficient (θ_8)	Dmnl
Effect of Rumination on Time Constant (θ_9)	1/RumScore
RumNoise Standard Deviation (θ_{10})	Dmnl
DepNoise Standard Deviation (θ_{11})	Dmnl
Correlation Time (θ_{12})	Month

*Dimensionless

Steps to estimate the parameters of the MDD-rumination model

1. **Define and estimate a set of empirical-auxiliary statistics.** Our auxiliary models include three regressions. The first regression, presented in equation (5), relates to rumination. Different measurements of the same concept are distinguished by subscripts. As mentioned in the previous section, rumination is influenced by gender, stressful life events, and depression. Besides these variables, we also included the previous values of rumination. The second regression presented in equation (6) captures the impact of rumination on depression and the regressors are rumination and previous values of MDD. The previous values of rumination and depression were included in the first and second

regressions, respectively, because the predictive power of models improved by adding them. In addition, incorporating previous values account for the inertia observed in those variables and encodes information about some of the time constants in our SD model. The third regression is an approximation of the change in stressor memory per month, presented in equation (7). The change in stressor memory was divided by seven months (the time between the two measurements) to get the stressor memory change per month.

$$Rum_3 = b_0 + b_1\, MDD_3 + b_2\, gender + b_3\, PerNegStim + b_4\, Rum_2 + b_5\, Rum_1 \quad (5)$$

$$MDD_3 = a_0 + a_1\, Rum_3 + a_2\, MDD_1 \quad (6)$$

$$(StressorMemory_3 - StressorMemory_1)/7 = c_0 - c_1 \frac{StresscorMemory_1}{\tau Rum_1} \quad (7)$$

To estimate the auxiliary-empirical statistics, we ran the three regressions (5-7). In addition, we included the mean of MDD at T_3 and rumination at T_2 and T_3 as statistics. The resulting empirical statistics (q=14) are listed in Table 3. Because q>(p=12), we have enough degrees of freedom to also test the model's specification quality after estimation.

Table 3: The value of empirical-auxiliary statistics

Regression	Statistic	Empirical-auxiliary Statistic
Equation (5)	b_0	-0.4663
	b_1	0.2313
	b_2	1.2021
	b_3	0.1316
	b_4	0.4548
	b_5	0.1749
Equation (6)	a_0	2.0012
	a_1	0.2526
	a_2	0.5559
Equation (7)	c_0	-0.0201
	c_1	-0.1222
Mean	Mean_MDD at T_3	9.7852
	Mean_Rum at T_2	10.8487
	Mean_Rum at T_3	9.9546

2. **Generate the simulated data using the SD model.** For generating a simulated data path, we first set the value of stocks to the corresponding empirical values (e.g.,

$MDDsymptoms_0$=MDD at T_1). Then, we generate H=200 paths by adding random noise to the indicated rumination and indicated depression for each individual (the resulting noise matrix has 200 columns and 661 rows, with two noise values for each cell corresponding to the MDD and Rumination noise). We repeat this procedure every time step as we simulate each individuals over the 7 months of simulation.

3. **Estimate the simulated-auxiliary statistics using the auxiliary model and simulated data.** After generating the simulated data, for a given θ, the simulated-auxiliary statistics are estimated similar to the empirical-auxiliary statistics for each path. In this case, we run three regressions parallel to those in equations 5-7, and include the other statistics to create a ${S_{SimAux}}^h$. Then, the average of these H simulated-auxiliary statistics is found as $(\frac{1}{H}\sum_{h=1}^{H} {S_{SimAux}}^h)$.

4. **Optimization.** A good estimate for the initial value of parameters can be found by running regressions on equations of indicated rumination and indicated MDD. The initial value for standard deviations of RumNoise and DepNoise are the residuals of these two regressions. The initial values of other parameters, effect of rumination on time constant (θ_9), and correlation time (θ_{12}), are arbitrarily selected. The initial values are summarized in the first column of Table 4. The unknown parameters (θ) are estimated by using fmincon solver in MATLAB. The same set noise matrices are used in each iteration of the optimization. The estimated parameters are shown in Table 4. All materials for estimating the parameters of the model are provided in the online appendix.

Results

Table 4 shows the estimated parameters of the SD model, including estimated parameters found in the first round of optimization, and estimated parameters after 10 rounds of optimizations. In the first round of optimization the weighting matrix is W_1, defined above. In the next rounds of optimizations, the weighting matrix is the inverse of the variance-covariance matrix of the statistics based on parameters estimated in the previous round of optimization. We run 2000 simulations to estimate the weight matrices. The parameters have fully converged after 7 rounds of optimization.

Table 4: Estimated parameters in the first and 10th rounds of optimization

Unknown Parameters	**Initial Value (θ_0)**	**First Round of Optimization**	**10th Round of Optimization**
Rumination Constant (θ_1)	0.3320	-0.5064	-1.2504 [-3.1920,0.6911]
Depression Effect Coefficient (θ_2)	0.2490	0.1187	0.4236 [-0.1661,1.0132]
Gender Coefficient (θ_3)	1.3540	0.7883	2.5152 [0.5518,4.4787]
Perceived Stress Coefficient (θ_4)	0.1240	0.0824	0.2518 [0.0227,0.4809]
Rumination Coefficient (θ_5)	0.5470	0.6202	0.1639 [-0.8064,1.1342]
Depression Constant (θ_6)	2.0010	0.3207	0.3730 [0.2968,0.4491]
Rumination Effect Coefficient (θ_7)	0.2520	0.0530	0.0699 [0.0638,0.0759]
Depression Coefficient (θ_8)	0.5560	0.9102	0.8894 [0.8822,0.8967]
Effect of Rumination on Time Constant (θ_9)	1.0000	2.1865	1.4741 [1.3735,1.5747]
RumNoise Standard Deviation (θ_{10}= [illegible]	5.8000	2.8678	7.8735 [-0.1391,15.8861]
DepNoise Standard Deviation (θ_{11} = [illegible]	6.0500	0.0016	0.0002 [-0.0307,0.0311]
Correlation Time (θ_{12})	1.0000	0.4266	1.6008 [0.0456,3.1559]

95% confidence interval are presented in parentheses.

Figure 3 compares the results of the first round of optimization and the final optimization. The blue circles represent the simulated-auxiliary statistics and the red bars depict the 95% confidence interval of empirical-auxiliary statistics (where such confidence intervals are available from auxiliary model estimations). The estimated parameters from the first round of optimization generate a few simulated-auxiliary statistics that are far away from the 95% confidence interval of the empirical-auxiliary statistics (Figure 3-A). After 10 rounds of optimization, almost all of the simulated-auxiliary statistics are within the 95% confidence intervals of the empirical-auxiliary statistics (Figure 3-B). Table 5 presents the values of the simulated and empirical auxiliary statistics shown in Figure 3.

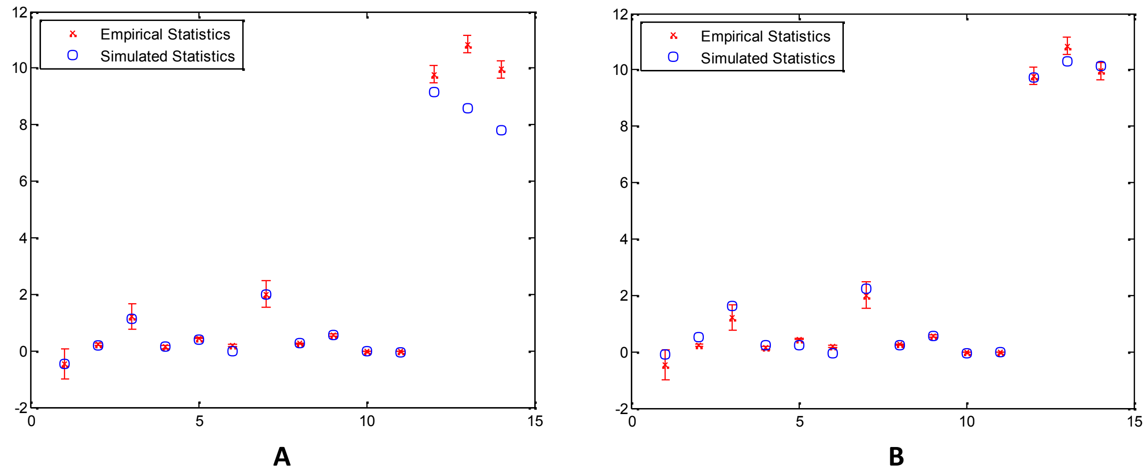

Figure 3: Empirical-auxiliary statistics and simulated-auxiliary statistics generated using the estimated parmeters from the first (**A**) and the 10th rounds of optimization (**B**)

Table 5: The values of empirical-auxiliary statistics and simulated-auxiliary statistics generated using the estimated parameters from the first and higher rounds of optimization

Regression	**Statistics**	**Empirical-Auxiliary Statistics**	**Simulated-Auxiliary Statistics (First round of optimization)**	**Simulated-Auxiliary Statistics (10th round of optimization)**
Equation (5)	b_0	-0.4663 [-1.4883, 0.5557]*	-0.48496	-0.09359
	b_1	0.2313 [0.1680, 0.2947]	0.196941	0.511783
	b_2	1.2021 [0.3046, 2.0996]	1.128222	1.609224
	b_3	0.1316 [0.0064, 0.2569]	0.134688	0.221293
	b_4	0.4548 [0.3819, 0.5276]	0.38188	0.225195
	b_5	0.1749 [0.1028, 0.2470]	-0.01602	-0.07331
Equation (6)	a_0	2.0012 [1.0910, 2.9113]	2.004988	2.234688
	a_1	0.2526 [0.1894, 0.3157]	0.263558	0.234232
	a_2	0.5559 [0.4760, 0.6358]	0.538416	0.5394
Equation (7)	c_0	-0.0201 [-0.06850, 0.0282]	-0.02015	-0.06472
	c_1	-0.1222 [-0.1588, -0.0857]	-0.04511	-0.00205
Mean	Mean_MDD at T_3	9.7852 [9.2013, 10.3690]	9.169462	9.722089
	Mean_Rum at T_2	10.8487 [10.2665, 11.4309]	8.563383	10.31732
	Mean_Rum at T_3	9.9546 [9.3475, 10.5617]	7.8181	10.13648

*95% confidence interval are presented in brackets.

Model specification and refinement

As it was explained in the fifth step of the method, when q>p, a test can be used to assess how well a model has been specified. Using equation 4, the ξ_T found to be 220. The 99% cut-off value for a Chi-square distribution with 2 degrees of freedom (14-12=2) is 9.2. Although almost all simulated-auxiliary statistics are within the confidence interval of the empirical-auxiliary statistics, our ξ_T is still higher than the cut-off value ($\xi_T = 220 > \chi_2^2 = 9.2$), hence, the model can be further refined.

Internal validity of this method can be tested using simulations. We thus check whether the parameters estimated by applying the indirect inference to a synthetic dataset, generated by simulating the calibrated model with coefficients reported in Table 4, are similar to true values (used for creating the synthetic data). The main idea behind this test is that in this case the data generating process is perfectly modeled and true parameter values are known, so any errors in parameter estimates can be attributed to the estimation method. The parameters in the third column of Table 4 are used to simulate the model and generate a synthetic dataset. Then all steps explained in the description of the model are applied to the synthetic data to find the indirect inference estimates. As it is shown in Table 6, the true parameters that are used to generate the synthetic data (first column) are within the 95% confidence interval of the estimated parameters using the synthetic data (second column). ξ_T is 7.48 which is lower than the cut off value.

Table 6: Estimated parameters using empirical data and synthetic data

Unknown Parameters	**Parameters Used to Generate Synthetic Data**	**Estimated Parameters Using Synthetic Data**
Rumination Constant (θ_1)	-1.2504	-0.0915 [-3.83, 3.65]*
Depression Effect Coefficient (θ_2)	0.4236	0.3111 [-0.03, 0.66]
Gender Coefficient (θ_3)	2.5152	2.8423 [-1.58, 7.27]
Perceived Stress Coefficient (θ_4)	0.2518	0.2411 [-0.19, 0.67]
Rumination Coefficient (θ_5)	0.1639	0.1722 [-1.19, 1.54]
Depression Constant (θ_6)	0.3730	-0.4226 [-7.00, 6.15]
Rumination Effect Coefficient (θ_7)	0.0699	0.0948 [0.05, 0.14]
Depression Coefficient (θ_8)	0.8894	0.8923 [0.77, 1.01]
Effect of Rumination on Time Constant (θ_9)	1.4741	1.4920 [1.39, 1.60]
RumNoise Standard Deviation (θ_{10}= [illegible]	7.8735	7.1009 [0.90, 13.30]
DepNoise Standard Deviation (θ_{11} = [illegible]	0.0002	17.9914 [-77.01, 113.00]
Correlation Time (θ_{12})	1.6008	2.7767 [-1.22, 6.77]

*95% confidence interval are presented in brackets.

Conclusion

This article provides a step-by-step introduction to the indirect inference method for estimating unknown parameters of dynamic models. In this method, the unknown parameters of the model of interest are estimated by matching the properties of empirical data and simulated data. In many applications, there are few empirical data points available over time; as a result, it is not feasible to use the conventional estimation methods such as the least squared error. In addition, unlike traditional methods, the indirect inference does not require the calculation of likelihood function, which may well be intractable for complex models. The indirect inference method extends the MSM by removing the requirement that the matching statistics be a set of valid moments. They can be parameters of an auxiliary model which is not an accurate description of the data generating process but it can be estimated easily by conventional estimation methods. When the dynamic model is too complicated with intractable likelihood function, when there are very few empirical data points over time, or when the number of available valid moments are smaller than the number of parameters of the model, indirect inference might be one of the few feasible options to estimate the unknown parameters of an SD model.

Results could make the contribution of SD to other fields more salient. For example, previous models of MDD have not incorporated the feedback mechanisms we discussed in our model. Our modeling and estimation results suggest that these feedbacks are indeed important and may be central to understanding MDD. Without the SD perspective, the previous literature ignored these feedback mechanisms, or at least did not quantify them. On the other hand, in the absence of indirect inference, traditional calibration methods in SD literature would not allow for using the common data structures available in this field (e.g. with 2-3 data points per person) to estimate a feedback-rich model. Many empirical datasets in psychology, medicine, organization studies, economics, and sociology share a similar structure. The results thus show the potential synergies between SD and indirect inference which could be explored well beyond MDD research.

Many advances in statistics have enabled researchers to estimate increasingly complex and realistic models with diverse types of data over the past three decades. We believe for SD to contribute to mainstream disciplinary research across various fields of social and behavioral sciences, modelers must be able to draw on the best available methods to estimate feedback-rich mechanism based models using quantitative data. We hope the introduction of indirect inference extends the toolbox of SD researchers and allows them to combine the benefits of broad model boundary and feedback richness, which traditional SD brings to understanding various phenomena, with the quantitative rigor of modern econometrics.

Acknowledgement

We would like to thank Dr. Kate McLaughlin for generously sharing her dataset with us. This research was funded by NIH/NIMH grant R21MH100515 (A.K.W., PI).

References:

Abela, J. Z., Brozina, K., & Haigh, E. (2002). An Examination of the Response Styles Theory of Depression in Third- and Seventh-Grade Children: A Short-Term Longitudinal Study. *Journal of Abnormal Child Psychology, 30*(5), 515-527. doi: 10.1023/a:1019873015594

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*(3), 411-423. doi: 10.1037/0033-2909.103.3.411

Coddington, R. D. (1972). The significance of life events as etiologic factors in the diseases of children: I—A survey of professional workers. *Journal of Psychosomatic Research, 16*(1), 7-18. doi: http://dx.doi.org/10.1016/0022-3999(72)90018-9

Duffie, D., & Singleton, K. J. (1993). Simulated Moments Estimation of Markov-Models of Asset Prices. *Econometrica, 61*(4), 929-952.

Durlauf, S. N., & Blume, L. E. (2008). The new Palgrave dictionary of economics. *Basingstoke, Hampshire New York: Palgrave Macmillan.*

Gourieroux, C., Monfort, A., & Renault, E. (1993). Indirect Inference. *J. of App. Econ., 8*, S85-S118.

Gouriéroux, C., Phillips, P. C., & Yu, J. (2010). Indirect inference for dynamic panel models. *Journal of Econometrics, 157*(1), 68-77.

Guvenen, F., & Smith, A. (2010). Inferring labor income risk from economic choices: An indirect inference approach: National Bureau of Economic Research.

Jalali, M., Rahmandad, H., & Ghoddusi, H. (2013). *Estimation of Unknown Parameters in System Dynamics Models Using the Method of Simulated Moments.* Paper presented at the Simulation Conference (WSC), 2013 Winter.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering, 82*(1), 35-45. doi: 10.1115/1.3662552

Kovacs, M. (1992). Children's Depression Inventory manual. *North Tonawanda, NY:Multi-Health Systems*.

Mcfadden, D. (1989). A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical-Integration. *Econometrica, 57*(5), 995-1026.

Mealli, F., & Rampichini, C. (1999). Estimating binary multilevel models through indirect inference. *Computational statistics & data analysis, 29*(3), 313-324.

Michl, L. C., McLaughlin, K. A., Shepherd, K., & Nolen-Hoeksema, S. (2013). Rumination as a mechanism linking stressful life events to symptoms of depression and anxiety: Longitudinal evidence in early adolescents and adults. *Journal of Abnormal Psychology, 122*(2), 339-352. doi: 10.1037/a0031994

Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology, 100*(4), 569-582. doi: 10.1037/0021-843x.100.4.569

Nolen-Hoeksema, S., Larson, J., & Grayson, C. (1999). Explaining the gender difference in depressive symptoms. *Journal of Personality and Social Psychology, 77*(5), 1061-1072. doi: 10.1037/0022-3514.77.5.1061

Nolen-Hoeksema, S., Stice, E., Wade, E., & Bohon, C. (2007). Reciprocal relations between rumination and bulimic, substance abuse, and depressive symptoms in female adolescents. *Journal of Abnormal Psychology, 116*(1), 198-207. doi: 10.1037/0021-843x.116.1.198

Pakes, A., & Pollard, D. (1989). SIMULATION AND THE ASYMPTOTICS OF OPTIMIZATION ESTIMATORS. *Econometrica, 57*(5), 1027-1057. doi: 10.2307/1913622

Rahmandad, H., & Sabounchi, N. (2011). *Building and estimating a dynamic model of weight gain and loss for individuals and populations*. Paper presented at the the 29th International Conference of the System Dynamics Society, Washington, DC.

Smith, A. A. (1993). Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics, 8*(S1), S63-S84. doi: 10.1002/jae.3950080506

Smith, G. L., Schmidt, S. F., McGee, L. A., Aeronautics, U. S. N., & Administration, S. (1962). *Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity on Board a Circumlunar Vehicle*: National Aeronautics and Space Administration.

Snijders, T. A. B. (2001). The Statistical Evaluation of Social Network Dynamics. *Sociological Methodology, 31*(1), 361-395. doi: 10.1111/0081-1750.00099

Wood, S. N. (2010). Statistical inference for noisy nonlinear ecological dynamic systems. [10.1038/nature09319]. *Nature, 466*(7310), 1102-1104. doi: http://www.nature.com/nature/journal/v466/n7310/abs/nature09319.html#supplementary-information