Quantal Response Equilibrium Strategy Frequency Estimation: A tractable statistical model for indefinitely-repeated games that doesn't work too well*

James R. Bland†

November 29, 2025

Abstract

This paper endogenizes strategy frequencies using a logit Quantal Response Equilibrium condition. The model has many fewer parameters than the classic strategy frequency estimation method, and can make out-of-sample predictions, which the classic strategy frequency estimation method cannot. What is less useful is that these predictions are hopeless. Either the model is wrong, or people don't play equilibrium strategies (in which case the model is still wrong). Researchers should probably stick to the classic strategy frequency estimation method insted of using the methods outlined below.
Keywords: Strategy frequency estimation, Quantal response equilibrium, Bayesian estimation, Prisoner's Dilemma, indefinitely-repeated games
JEL: C11, C18, C57, C72, C73, C92

1 Introduction

Understanding strategies is important to understanding how and if cooperation is supported in indefinitely repeated games. In direct-repsonse exeperiments, howerever, we do not observe strategies, we only observe *actions*. In this situation, the

*I thank the members of #EconTwitter and #EconSky for encouragement.
†The University of Toledo, 2801 Bancroft St, Toledo, OH 43606, USA, james.bland@utoledo.edu, orcid.org/0000-0002-7117-9998

Strategy Frequency Estimation Method (SFEM, Dal Bó and Fréchette, 2011; Fudenberg et al., 2012; Dvorak, 2023) is often used to estimate the importance of various strategies. One downside of the SFEM is that there is no sense of equilibrium. That is, players are just assumed to play strategies, but in no way are they responding to the incentives of the game or the strategies used by their opponents. This paper introduces a new technique: the Quantal Response Equilibrium Strategy Frequency Estimation Method (QRE-SFEM), to analyze indefinitely-repeated games. It merges two concepts in the econometrics of games literature. First, I simplify the infinity of strategies that players could be using by makeing a classic SFEM-like assumption that players choose one strategy from an assumed list. Secondly, I impose an equilibrium condition. That is, strategies are chosen *endogenously* (unlike the classic SFEM) as a response to the distribution of strategies in the population. To do this, I use a logit Quantal Response Equilibrium (QRE, McKelvey and Palfrey, 1995) condition that assumes players noisily best respond to the distribution of strategies.

The QRE-SFEM is potentially attractive becuase it scales well with the number of strategies. That is, there are just two parameters in the model, trembles and choice precision, irrespective of the number of strategies on the list. This is because strategy frequencies in the QRE-SFEM are determined endogenously. Compare this to the classic SFEM, where the number of parameters grows linearly with the number of strategies. As such, it allows the practitioner to entertain perhaps many more on-list strategies. Furthermore, the QRE-SFEM can pool data across treatments, whereas the classic SFEM is typically estimated one treatment at a time. This parsimony allows the model to make out-of-sample predictions about the strategies used in treatments that were not studied in the experiment.

I demonstrate the QRE-SFEM using data from Dal Bó and Fréchette (2011). I show that the QRE-SFEM generally performs poorly at organizing the data and in predicting strategy frequencies both within-sample and out-of-sample. If we take the classic SFEM estimates seriously, this suggests that participants do not play equilibrium strategies, or that the chosen equilibrium concept does not apply.

While I do not advise using the techniques outlined below, the replication archive for this paper can be found at `https://github.com/JamesBlandEcon/QRESFEM`. In it, you will find moderately cutting-edge code for implementing the SFEM and computing QRE efficiently in *Stan* (Carpenter et al., 2017). Separately they may be of use to someone. Please do not ever combine these techniques again.

Table 1: The Prisoner's Dilemma and the normalized Prisoner's Dilemma

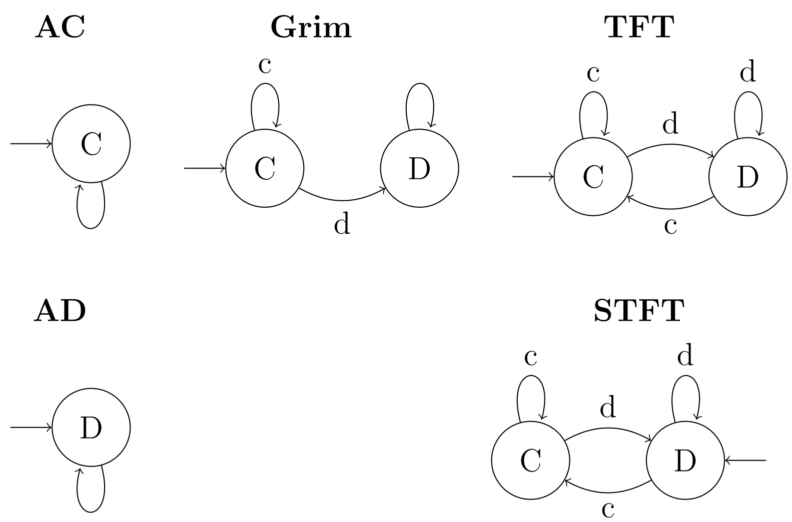
(a) Raw parameterization

	C	D
C	R, R	S, T
D	T, S	P, P

(b) normalized parameterization

	C	D
C	1, 1	$-l$, $1+g$
D	$1+g$, $-l$	0, 0

Figure 1: Some commonly used SFEM strategies



2 QRE with a list of strategies

Consider the Prisoner's Dilemma and its normalized parameterization, as shown in Table 1(b). In indefinitely-repeated games, this stage game is played with a random stopping rule described by *continuation probability* $\delta \in (0, 1)$. That is, players play another stage game with probability δ, and stop playing with probability $1 - \delta$. In the normalized parameterization (panel (b)), which I shall mainly work with below, g measures the gain from defecting when one's opponent cooperates, and l measures the loss from cooperating when one's opponent defects.

Strategies in indefinitely-repeated games can be complicated because in principle they can condition actions on the entire history of play. A popular work-around for estimating strategies is to assume that players use one of an assumed list of strategies, such as the one in Figure 1. Here, I restrict attention to 2-state strategies, which can be represented as seven bits of information (Zhang, 2018, see Appendix A):[1]

$$\begin{aligned}
s &= \{x_0, a_0, x_{0,0}, x_{0,1}, a_1, x_{1,0}, x_{1,1}\} \\
x_0 &= \text{starting state} \\
a_k &= \text{action played in state} \\
x_{k,j} &= \text{transition state if opponent plays action } j \\
x_0,\ a_k,\ s_{k,j} &\in \{0, 1\}
\end{aligned} \tag{1}$$

however the methodology can easily be extended to longer memory strategies, perhaps with substantial additional computational burden. Given that participants are restricted to playing one strategy $s \in \mathcal{S}$, let $\rho \in \Delta^{|\mathcal{S}|-1}$ be the *strategy frequencies*. That is, ρ_s is the fraction of players using strategy s. As with the classic SFEM, I also assume that playes *tremble* with probability $\epsilon \in (0, 0.5)$. That is, they choose the action *not* prescribed by their strategy with probability ϵ. Together, the set of strategies $\mathcal{S}$ and the tremble probability ϵ describes a $|\mathcal{S}| \times |\mathcal{S}|$ payoff matrix V, where $V_{s,t}$ is the discounted expected payoff from playing strategy s against an opponent who plays strategy t.

For two given strategies s and t, let H be the transition probabilities between states. $H_{i,j}$ is the probability that we enter state j given that we are in state i now. Let h denote the transition probabilities from states to outcomes. For two arbitrary strategies, we can compute this as follows:

$$h_{s,o} = (1-\epsilon)^F \epsilon^{2-F} \tag{2}$$

$$\text{where: } F = I(a_1 = o_1) + I(a_2 = o_2) \tag{3}$$

where $o = \{o_1, o_2\}$ indexes the outcomes of the game, and $I(\cdot)$ is the indicator function. That is $o \in \{\{DD\}, \{DC\}, \{CD\}, \{CC\}\}$. We can then construct H, the transition probability between states, using the strategies and h. Now, letting $u = \begin{pmatrix} 0 & 1+g & -l & 1 \end{pmatrix}^\top$, we can write the Bellman equation for the value function $U_{s,t}$ of playing strategy s against strategy s as follows:

$$\begin{aligned}
U_{s,t} &= (1-\delta)u + \delta H U_{s,t} && (4) \\
&= (1-\delta)\left[\mathcal{I} - \delta H\right]^{-1} u && (5)
\end{aligned}$$

Finally, the relevant element of $U_{s,t}$ is the starting states of the strategies:

$$V_{s,t} = U_{s,t}(x_0^s, x_0^t) \tag{6}$$

[1]For example the "Tit-for-tat" strategy in Figure 1 can be represented by the 7-bit string 1001101.

At this point, we have constructed the payoff matrix V, where $V_{s,t}$ is the expected discounted payoff from playing strategy s against strategy t. Now it is time to endogenize the strategy frequencies ρ. To do this, note that the expected utility from playing each strategy when playing against a population of players who choose strategies according to ρ is $V\rho$. Here, I use a logit Quantal Response Equilibrium condition, which assumes that the probabilistic best response to opponents mixing according to ρ is:

$$q(\lambda; V\rho) = \frac{\exp(\lambda V\rho)}{\sum_{j\in\mathcal{S}} \exp(\lambda[V\rho]_j)} \tag{7}$$

where $\lambda \geq 0$ is a *choice precision* parameter. When $\lambda = 0$ players randomize uniformly over the strategy space, and as $\lambda \to \infty$ the probabilistic best response function approaches a best response function. The equilibrium condition for QRE is that the probabilistic best response to ρ is equal to ρ:

$$\rho = q(\lambda; V\rho) \tag{8}$$

Solving (8) can be achieved using numerical continuation techniques developed in Turocy (2005) and Turocy (2010). See Bland and Turocy (2025) for a detailed discussion of implementing these techniques.

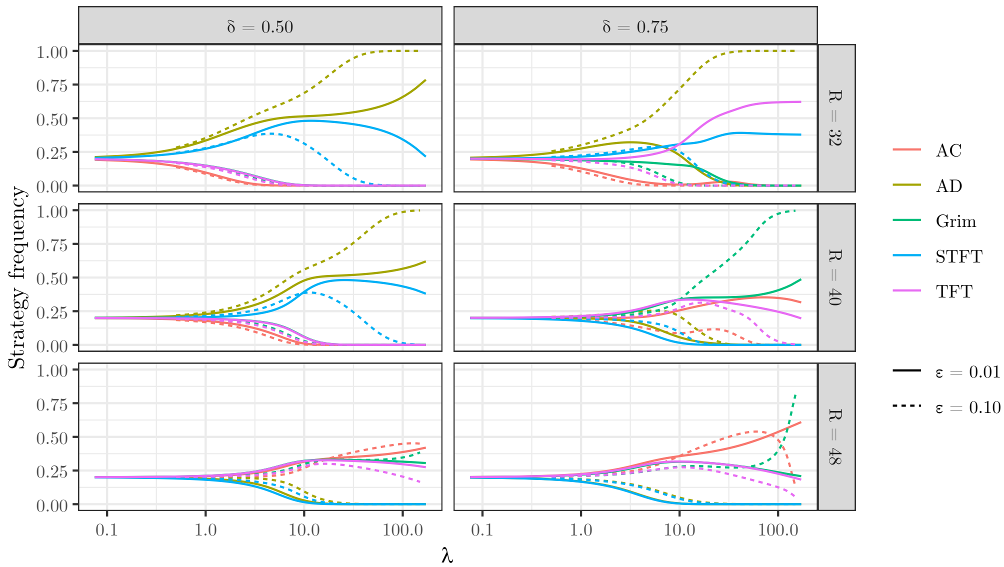
Estimation of the QRE-SFEM can be achieved using likelihood-based techniques, where the log-likelihood function is:

$$\log p(y, c \mid \lambda, \epsilon) = \sum_{i=1}^{N} \log\left(\sum_{s=1}^{|\mathcal{S}|} \rho_s(\lambda, \epsilon)(1-\epsilon)^{y_{i,s}} \epsilon^{c_i - y_{i,s}}\right) \tag{9}$$

where $y_{i,s}$ is a count of the number of times participant i chose the action prescribed by strategy s, and c_i is the total number of actions played by participant i. The likelihood function for the classic SFEM is identical to (9) except that ρ is a fundamental parameter of the model (rather than being a function of λ and ϵ). In the applications section below, I recover the parameters of the QRE-SFEM and SFEM using Bayesian techniques implemented in *Stan* (Carpenter et al., 2017).[2]

[2]See Bland (2023) for a reference on how to estimate QRE models using Bayesian techniques, and for a discussion of the merits of using Bayesian techniques for estimating QRE. Bland (2025b) also has chapters dedicated to estimating the classic SFEM using Bayesian techniques (Chapters 11 and 12) and computing and estimating QRE models (Chapters 13-16).

Figure 2: Logit Quantal Response Equilibrium strategy frequencies for two values of ϵ



3 An application

In order to showcase the QRE-SFEM, I turn to the data used for the SFEM's original application in indefinitely-repeated Prisoner's Dilemma experiments: Dal Bó and Fréchette (2011).[3] This experiment consisted of six treatments in a 2×3 factorial between-subjects design varying the continuation probability $\delta \in \{0.50, 0.75\}$ and the cooperation payoff $R \in \{32, 40, 48\}$ (see Table 1(a)). The other stage-game payoffs were held constant at $T = 50$, $S = 12$, and $P = 25$. Following the SFEM as implemented by Dal Bó and Fréchette (2011), I limit attention to data generated in supergames that started after at least 110 interactions.

For simplicity, I begin with estimating the QRE-SFEM using just the five strategies listed in Figure 1. Figure 2 traces out the principal branch of the QRE-SFEM for two values of tremble probability ϵ.[4] This provides us with some kind of idea about the relevant values of choice precision λ. Here we can see that for $\lambda < 0.1$

[3]SFEM estimations can also be found in Fudenberg et al. (2012), so given the long time it takes to publish in economics, it is reasonable to assume that these were done in parallel.

[4]As with other generalizations of Nash equilibrium, QRE is not necessarily unique. Instead, the literature typically focuses on the "principal branch", which can be found by tracing out QRE mixed strategies as a function of λ starting from $\lambda = 0$ and $\rho_s = \frac{1}{|\mathcal{S}|}$, which is always a logit QRE.

in all treatments the model is predicting very close to uniform randomization over the five strategies. Then as λ becomes larger than about 100, play seems to have converged for most of the treatments. I will use this to calibrate my prior for λ that is used in Bayesian estimation.[5] Specifically, since $\lambda \geq 0$, I will assume a log-normal prior for λ, and let the range of λ discussed above be the 95% prior Bayesian credible region for λ. That is:

$$\Pr(0.1 < \lambda < 100) \implies \log(100) = m + 1.96s \tag{10}$$

$$\log(0.1) = m - 1.96s \tag{11}$$

$$\log(\lambda) \sim N(1.15, 1.76^2) \tag{12}$$

Then, for ϵ I assume a uniform prior, truncated to the $(0, 0.5)$ interval. When I estimate a classic SFEM, I assume a Dirichlet($\mathbf{1}$) prior for ρ, where $\mathbf{1}$ is an $|\mathcal{S}|$-vector of ones.

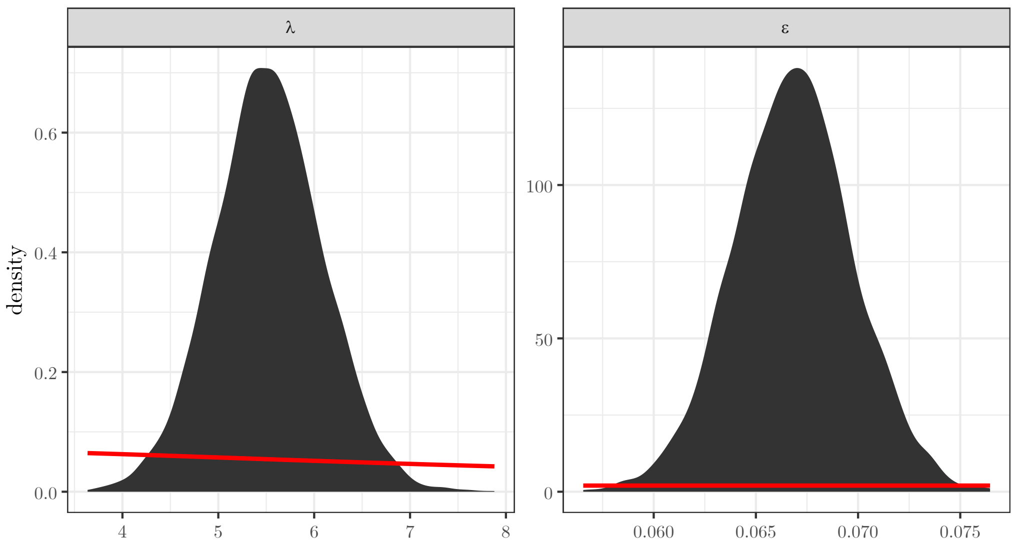
I start by estimating the QRE-SFEM once using all six treatments' data. Figure 3 shows the marginal posterior distributions (shaded desnities) of the QRE-SFEM's two fundamental parameters λ (choice precision) and ϵ (trembles). The prior density is also shown in red. Here we can see substantial updating in that the prior densities are much flatter than the posteriors. In other words, we have learned substantially from the Dal Bó and Fréchette (2011) data. While these parameters may be of interest on their own, what they imply about strategy frequencies is probably more interesting. A summary of the posterior distribution of strategy frequencies from this model is shown in Table 2. Importantly, all of these thirty values are functions of just two fundamental parameters of the model: λ and ϵ. This highlights the parsimony of the model: whereas the classic SFEM has $|\mathcal{S}|$ free parameters *per treatment*,[6] the QRE-SFEM just has two in total, and because of the equilibrium condition it can make predictions across different treatments. In principle, one could use the QRE-SFEM to predict strategy frequencies in treatments not studied in an experiment. In contrast, the classic SFEM cannot make out-of-sample predictions, becuase there is no equilibrium condition embedded into it. So how well does the QRE-SFEM do in this situation relative to the classic SFEM?[7] One Bayesian measure of a model's

[5] See Bland (2025a) for a discussion of how to choose priors for structural models in experimental economics. The following calibration falls under the "Priors informed by our understanding of transformations of the parameters" technique discussed in that article. Bland (2023) describes prior calibration specifically for QRE models.

[6] That is, $|\mathcal{S}| - 1$ free parameters go into defining the strategy frequencies for each treatment, and then there is ϵ.

[7] The posterior estimates for the classic SFEM are shown in the columns labeled "SFEM" in Table 4.

Figure 3: Posterior distribution of the fundamental parameters from the pooled QRE-SFEM model



Notes: Shaded curve shows kernel-smoothed posterior density. Red curve shows prior density.

Table 2: QRE-SFEM strategy frequency estimates from the fully pooled model

	$\delta = 0.50$			$\delta = 0.75$		
	$R = 32$	$R = 40$	$R = 48$	$R = 32$	$R = 40$	$R = 48$
AD	0.557	0.447	0.418	0.211	0.142	0.083
	(0.010)	(0.018)	(0.021)	(0.007)	(0.011)	(0.010)
AC	0.000	0.003	0.047	0.146	0.238	0.311
	(0.000)	(0.002)	(0.011)	(0.004)	(0.007)	(0.012)
Grim	0.017	0.125	0.119	0.261	0.257	0.262
	(0.006)	(0.012)	(0.012)	(0.010)	(0.008)	(0.005)
TFT	0.014	0.118	0.110	0.249	0.251	0.267
	(0.005)	(0.012)	(0.012)	(0.008)	(0.007)	(0.005)
STFT	0.411	0.307	0.306	0.132	0.111	0.077
	(0.003)	(0.008)	(0.015)	(0.010)	(0.012)	(0.011)

Notes: Posterior means with posterior standard deviations in parentheses.

performance relative to another is its posterior probability. I estimate a posterior probability for the QRE-SFEM in the order of about 10^{-29}.[8] From this, we can conclude "not well at all".

Before assessing how well the QRE-SFEM can perform making out-of-sample predictions, I give it its best possible chance to do as well as the classic SFEM by comparing the models' performances treatment-by-treatment. That is, for each treatment separately, I estimate the QRE-SFEM and the classic SFEM. By "best possible chance" I mean that here I am only asking the QRE-SFEM's two parameters to do similarly well to the classic SFEM's five parameters. That is, I am not pooling the QRE-SFEM over multiple treatments. Furthermore, I am assessing the QRE-SFEM's ability to predict strategy frequencies *within-sample* (rather than out-of-sample). That is, I only ask the QRE-SFEM to predict the strategy frequencies in the treatment on which it had data. The posterior estimates of these models are shown in Table 4. Here we can see an amazingly strong agreement between the models on the tremble probability! This is in terms of both the posterior mean and the posterior standard deviation. However for the values that practitioners actually

[8]This is computed using the `bridgesampling` library (Gronau et al., 2020) in *R* (R Core Team, 2025).

Table 3: Posterior probabilities of the QRE-SFEM

R	$\delta = 0.5$	$\delta = 0.75$
32	0.9874	0.1591
40	0.7436	0.0000
48	0.0000	0.0000

Notes: Posterior probability of the QRE-SFEM where the alternative model is the classic SFEM. Prior assigns equal probability to the two models. Models estimated separately by treatment.

care about, strategy frequencies, there is not much agreement, except for the AD strategy in some treatments. To illustrate this, Table 3 shows the QRE-SFEM model's posterior probabilities for each treatment (relative to the classic SFEM). Here we can see strong support for the QRE-SFEM in the $R = 32$, $\delta = 0.5$ treatment (where cooperation should not be supported in equilibrium), and mild support for it in the $R = 40$, $\delta = 0.5$ treatment. All other treatments favor the classic SFEM.

Next, I turn to assessing the ability of the QRE-SFEM to make out-of-sample predictions about strategy frequencies. To this end, I re-estimate the QRE-SFEM six more times, each time leaving out one treatment. I then compare the estimated strategy frequencies of the QRE-SFEM for the left-out treatment to the classic SFEM estimates for that treatment. This comparison is shown in Figure 4. The black dashed line is a 45° line. If the QRE-SFEM is performing well, then all of the strategies should fall on or close to this line. To the contrary, here we can see no real relationship between the SFEM estimates and the QRE-SFEM predictions.

Finally, I assess the QRE-SFEM's ability to organize a much larger list of strategies. In particular, I use all strategies that can be represented using the 7-bit representation described in (1). There are twenty-six such strategies.[9] I compare this model to the classic SFEM using just the five strategies in Figure 1.[10] Here, the QRE-SFEM has a model posterior probability in the order of about 10^{-126} compared to the classic SFEM with just five strategies. That is, it also performs *worse* than the QRE-SFEM with just five strategies.

[9] While at first it may seem that there should be $2^7 = 128$ of these strategies, substantial culling of this list can be achieved by noting that many groups of strategies are isomorphic.

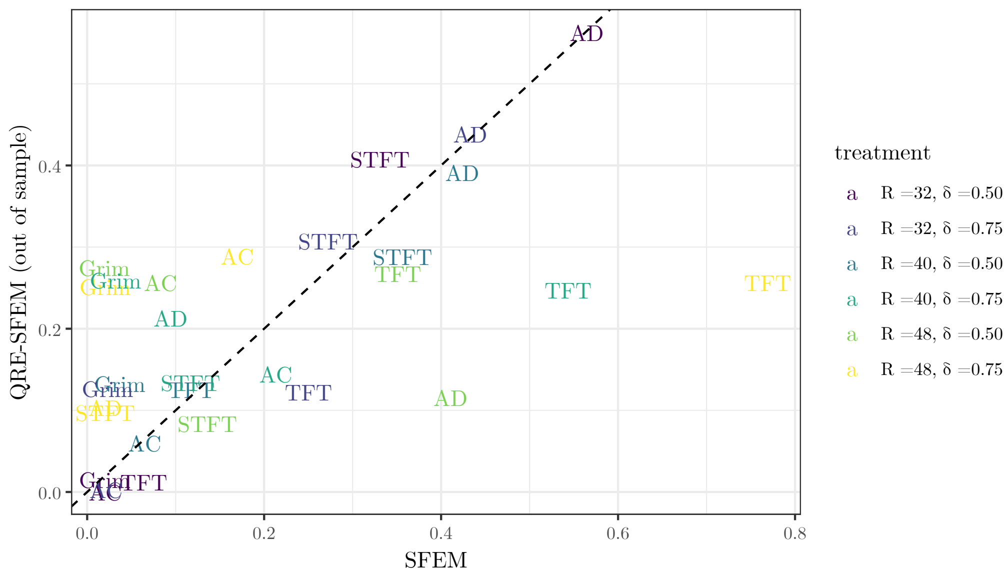
[10] This is because the classic SFEM is poorly identified for the full list of 26 strategies. I suspect that this is becuase the number of strategies is close enough to the number of participants in each treatment.

Table 4: QRE-SFEM and classic SFEM estimates estimating each treatment separately

	$\delta = 0.50$						$\delta = 0.75$					
	$R = 32$		$R = 40$		$R = 48$		$R = 32$		$R = 40$		$R = 48$	
	QRE	SFEM	QRE	SFEM	QRE	SFEM	QRE	SFEM	QRE	SFEM	QRE	SFEM
λ	5.270		5.831		0.930		4.561		3.350		21.993	
	(1.363)		(0.976)		(0.827)		(1.261)		(3.083)		(11.935)	
ϵ	0.038	0.038	0.118	0.117	0.068	0.067	0.078	0.080	0.061	0.064	0.033	0.034
	(0.007)	(0.007)	(0.009)	(0.009)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.005)	(0.005)
AD	0.519	0.565	0.465	0.424	0.195	0.411	0.442	0.433	0.198	0.094	0.004	0.020
	(0.019)	(0.115)	(0.038)	(0.084)	(0.005)	(0.069)	(0.043)	(0.074)	(0.023)	(0.045)	(0.007)	(0.020)
AC	0.002	0.020	0.031	0.065	0.204	0.083	0.010	0.021	0.176	0.213	0.451	0.170
	(0.003)	(0.020)	(0.016)	(0.037)	(0.004)	(0.040)	(0.012)	(0.021)	(0.018)	(0.088)	(0.052)	(0.084)
Grim	0.029	0.020	0.094	0.037	0.207	0.019	0.137	0.023	0.240	0.032	0.272	0.020
	(0.017)	(0.020)	(0.022)	(0.027)	(0.007)	(0.019)	(0.025)	(0.022)	(0.045)	(0.029)	(0.019)	(0.020)
TFT	0.025	0.064	0.083	0.118	0.206	0.351	0.122	0.250	0.233	0.544	0.269	0.769
	(0.015)	(0.036)	(0.021)	(0.045)	(0.006)	(0.068)	(0.025)	(0.062)	(0.035)	(0.100)	(0.025)	(0.089)
STFT	0.426	0.331	0.326	0.357	0.188	0.136	0.289	0.272	0.154	0.117	0.003	0.021
	(0.017)	(0.114)	(0.021)	(0.083)	(0.011)	(0.047)	(0.018)	(0.068)	(0.047)	(0.047)	(0.006)	(0.019)

Notes: Posterior means with posterior standard deviations in parentheses.

Figure 4: Out-of-sample predictions



Notes: Coordinates are at the posterior mean estimates. Black dashed line is a 45° line.

4 Conclusion

The QRE-SFEM is a tractable, 2-parameter statistical model of play in indefinitely-repeated games. It endogenizes strategy frequencies using a logit Quantal Response Equilibrium condition. As such, the model is highly portable, and can make out-of sample predictions. Unfortunately when taken to data, the model performs poorly on many counts relative to the classic SFEM.

If we are to take the classic SFEM estimates seriously, these results suggest that the logit Quantal Response Equilibrium condition is a bad way to endogenize strategy frequencies in indefinitely-repeated games. A possible take-away from this is that people are not playing equilibrium strategies, even after many rounds of play.

Possible extensions of the QRE-SFEM could include allowing for participant-specific trembles as introduced to the classic SFEM in Bland (2020), or strategy switching between supergames as modeled by Backhaus and Breitmoser (2024). However given the poor performance of the vanilla QRE-SFEM, it is advisable for researchers to reach for lower-hanging fruit. Perhaps an alternative to the QRE endogenization of strategy frequencies could be a level-k model (Stahl and Wilson, 1994).

References

Backhaus, T. and Y. Breitmoser (2024). Individual strategy choice in the repeated prisoner's dilemma. Technical report, Discussion Paper Series–CRC TR 224, 341.

Bland, J. R. (2020). Heterogeneous trembles and model selection in the strategy frequency estimation method. *Journal of the Economic Science Association 6*(2), 113–124.

Bland, J. R. (2023). Bayesian inference for quantal response equilibrium in normal-form games. *Games and Economic Behavior*.

Bland, J. R. (2025a). Some guidance for the choice of priors for bayesian structural models in economic experiments. *Journal of the Economic Science Association*, 1–10.

Bland, J. R. (2025b). Structural Bayesian techniques for experimental and behavioral economics, with applications in R and Stan.

Bland, J. R. and T. L. Turocy (2025). Quantal response equilibrium as a structural model for estimation: The missing manual. *Games and Economic Behavior*.

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. A. Brubaker, J. Guo, P. Li, and A. Riddell (2017). Stan: A probabilistic programming language. *Journal of statistical software 76*.

Dal Bó, P. and G. R. Fréchette (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review 101*(1), 411–29.

Dvorak, F. (2023). stratEst: a software package for strategy frequency estimation. *Journal of the Economic Science Association 9*(2), 337–349.

Fudenberg, D., D. G. Rand, and A. Dreber (2012). Slow to anger and fast to forgive: Cooperation in an uncertain world. *American Economic Review 102*(2), 720–49.

Gronau, Q. F., H. Singmann, and E.-J. Wagenmakers (2020). bridgesampling: An r package for estimating normalizing constants. *Journal of Statistical Software 92*, 1–29.

McKelvey, R. D. and T. R. Palfrey (1995). Quantal response equilibria for normal form games. *Games and economic behavior 10*(1), 6–38.

R Core Team (2025). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Stahl, D. O. and P. W. Wilson (1994). Experimental evidence on players' models of other players. *Journal of economic behavior & organization 25*(3), 309–327.

Turocy, T. L. (2005). A dynamic homotopy interpretation of the logistic quantal response equilibrium correspondence. *Games and Economic Behavior 51*(2), 243–263.

Turocy, T. L. (2010). Computing sequential equilibria using agent quantal response equilibria. *Economic Theory 42*(1), 255–269.

Zhang, H. (2018). Errors can increase cooperation in finite populations. *Games and Economic Behavior 107*, 203–219.