WEB APPENDICES
for
"Psychological Pressure in Competitive Environments: Evidence from a Randomized Natural Experiment"

September 2009

Table of Contents

APPENDIX A. Dynamic Panel Data Analysis with Lagged Endogenous Variables.

As is well known, in the presence of lagged endogenous variables in linear models with additive effects the standard response is to consider instrumental-variables estimates which exploit the lack of correlation between lagged values of the variables and future errors in first differences. In non-linear models, however, very few results are available.[1] In this appendix we describe the Manuel Arellano and Raquel Carrasco (2003) model and how it is applied to the setting we study. We also report the results of the multinomial logit analysis applied to the penalty shoot-outs for which we have data on "goals," "misses" and "saves."

The basic idea of the Arellano-Carrasco model is to define conditional probabilities for every possible sequence of realizations of the state variables. Then, the estimator computes the probability of a given outcome along every possible path of past realizations of the endogenous regressors. The panel data structure allows the identification of the effect of individual unobserved heterogeneity since outcomes can be different even when teams share the same history of realizations of the state variables.

Consider two discrete outcomes (score, no score) denoted $y_{it} = \{1, 0\}$. The probability of each of them depends on the specific sequence of past outcomes and the state of the shoot-out tournament. Since outcomes can be different, different experiences change the information set and the expected realizations of future outcomes. To be more specific, the probability of a given outcome may depend on certain intrinsic characteristics of the teams involved in the shoot-out, as well as on their expectation on the realization of the final outcome. This can be written as follows:

$$y_{it} = \mathbf{1}\left\{\beta z_{it} + E\left(\eta_i \mid w_i^t\right) + \varepsilon_{it} \geq 0\right\},$$

$$\varepsilon_{it} \mid w_i^t \sim N\left(0, \sigma_t^2\right),$$

where z_{it} includes the set of time–invariant characteristics of the teams and the shoot-out, x_{it}, plus the state of the shoot-out and the previous outcomes $y_{i(t-1)}$. Denote

[1] For fixed effects the few available methods are case-specific (logit and Poisson) and, in practice, lead to estimators that do not converge at the usual $\sqrt{n}$-rate. In the case of random effects, the main difficulty is the so-called initial conditions problem: if one begins to observe subjects after the "process" in question is already in progress, it is necessary to isolate the effect of the first lagged dependent variable from the individual-specific effect and the distribution of the explanatory variables prior to the sample.

by $w_i^t = \{w_{i1}, \ldots, w_{it}\}$ the history represented by a sequence of realizations $w_{it} = \{x_{it}, y_{i(t-1)}\}$, and by η_i an individual effect (future outcome realization for team i) whose forecast is revised each period t as the information summarized by the history w_i^t accumulates.[2] The conditional distribution of the sequence of expectations $E(\eta_i \mid w_i^t)$ is left unrestricted, and hence the process of updating expectations as information accumulates is not explicitly modeled. This is the only aspect that makes the model semi–parametric. Given the history of past outcomes, since errors are normally distributed, the conditional probability of $y_{it} = 1$ at time t for any given history w_i^t is:

$$\Pr\left(y_{it} = 1 \mid w_i^t\right) = \Phi\left[\frac{\beta z_{it} + E\left(\eta_i \mid w_i^t\right)}{\sigma_t}\right].$$

Since the model has discrete support, any individual history can be summarized by a cluster of nodes $j = 1, \ldots, J$ representing the sequence of realizations for each vector of characteristics. Thus, the conditional probability can be rewritten as:

$$p_{jt} = \Pr\left(y_{it} = 1 \mid w_i^t = \phi_j^t\right) \equiv h_t\left(w_i^t = \phi_j^t\right), \qquad j = 1, \ldots, J.$$

The estimation relies on an intuitive idea. In order to remove the unobserved individual effect, we account for the proportion of teams with identical characteristics and history up to time t that realize a given outcome at time t. We then repeat this procedure for every cluster of combinations of demographics and histories in our data. For each cluster we compute the percentage of times that outcome $y_{it} = 1$ occurs. This provides a simple estimate of the unrestricted probability $\hat{p}_{jt}$ for each possible history in the sample. Then, by taking first differences of the inverse of the equation above we get:

$$\sigma_t \Phi^{-1}\left[h_t\left(w_i^t\right)\right] - \sigma_{t-1}\Phi^{-1}\left[h_{t-1}\left(w_i^{t-1}\right)\right] - \beta\left(x_{it} - x_{i(t-1)}\right) = \xi_{it},$$

and, by the law of iterated expectations, we have:

$$E\left[\xi_{it} \mid w_i^{t-1}\right] = E\left[E\left(\eta_i \mid w_i^t\right) - E\left(\eta_i \mid w_i^{t-1}\right) \middle| w_i^{t-1}\right] = 0.$$

This conditional moment condition serves as the basis of the GMM estimation of parameters β and σ_t (subject to the normalization restriction that $\sigma_1 = 1$). Arellano

[2] The specification of Arellano and Carrasco (2003) is more general in the sense that it also includes a time-varying component, γ_t, common to all individuals. In our case all "demographic" variables are time–invariant.

and Carrasco (2003) show that there is no efficiency loss in estimating these parameters by a two–step GMM method where in the first step the conditional probabilities p_{jt} are replaced by unrestricted estimates $\hat{p}_{jt}$, which in our case are the proportion of teams with given characteristics and a given history. Then:

$$\hat{h}_t\left(w_i^t\right) = \sum_{j=1}^{J} \mathbf{1}\left\{w_i^t = \phi_j^t\right\} \cdot \hat{p}_{jt},$$

can be used to define the sample orthogonality conditions of Arellano–Carrasco's GMM estimator:[3]

$$\frac{1}{N}\sum_{i=1}^{N} d_{it}\left\{\sigma_t \Phi^{-1}\left[\hat{h}_t\left(w_i^t\right)\right] - \sigma_{t-1}\Phi^{-1}\left[\hat{h}_{t-1}\left(w_i^{t-1}\right)\right] - \beta\left(x_{it} - x_{i(t-1)}\right)\right\} = 0, \quad t = 2, \ldots, T,$$

where d_{it} is a vector containing the indicators $\mathbf{1}\left\{w_i^t = \phi_j^t\right\}$ for $j = 1, \ldots, J$.

With respect to the magnitude of the effects, the marginal effects associated with the transition among different states can be computed as follows. Arellano and Carrasco (2003) show that the probability of a given outcome when we compare two states $z_{it} = z^0$ and $z_{it} = z^1$ changes by the proportion:

$$\hat{\triangle}_t = \frac{1}{N}\sum_{i=1}^{N}\left\{\Phi\left(\hat{\sigma}_t^{-1}\hat{\beta}\left(z^1 - z_{it}\right) + \Phi^{-1}\left[\hat{h}_t\left(w_i^t\right)\right]\right) - \Phi\left(\hat{\sigma}_t^{-1}\hat{\beta}\left(z^0 - z_{it}\right) + \Phi^{-1}\left[\hat{h}_t\left(w_i^t\right)\right]\right)\right\}.$$

Since this proportion depends on the history of past ω_i^t, these marginal effects are different for each partial score in the sample, and for each team.

Finally, a reason why the Arellano-Carrasco is preferred in our setting is that alternative fixed-effects approaches such as Bo Honoré and Arthur Lewbel (2002) and Honoré and Ekaterini Kyriazidou (2000) are far more demanding in terms of data. In particular, they require the exogenous regressors to vary over time, something that does not occur in our data. Honoré and Kyriazidou (2000) include one lagged dependent variable but require that the remaining explanatory variables should be strictly exogenous, thus excluding the possibility of a lagged dependent regressor. Further, their estimator does not converge at the usual $\sqrt{n}$–rate. Honoré and Lewbel (2002)

[3] We use the orthogonal deviations suggested by Arellano and Olympia Bover (1995) instead of first differences among past values of the state variables.

allow for additional predetermined variables but at the cost of requiring a continuous, strictly exogenous, explanatory variable that is independent of the individual effects. See Ivan Fernandez-Val (2009) for a characterization of the bias of fixed effect estimators in non-linear panel data models.

Finally, as indicated in the main text of the article, for a subset of all the penalty shoot-outs in the sample we have detailed information on whether the no-goals are due to "saves" by the goalkeeper or "misses" by the kicker. Table A1 reports the results of a multinomial logit specification with goals, misses and saves using these data.

[Table A1 here]

Panel A reports the raw data in scoring, misses and saving rates for the first and second team. We find that both teams have basically the same proportion of saves, and hence that the difference in scoring rates between the first and the second team basically corresponds to their difference in misses.

In Panel B we report the results of different regression specifications. We find that the coefficient on "Partial score -1" is positive and highly significant (beyond the 1 percent level) for misses, but insignificant for saves in all the specifications. This means that lagging in the score predicts more misses by the kicker but no more saves by the goalkeeper. The interaction with the "Second team" variable is negative and significant, which means that when the partial score is -1 the first kicking team is more likely to miss. This, as in Table 7, is likely the result of being in an objectively worse situation than the second team (it has had the same number of chances of scoring, whereas, at every kick, the second team has always had one less chance). No variable except the constant term is significant for saves. The variable "Penalty Kick Importance," although not significant at the conventional significance levels, has a fairly high t-statistic for misses.

We take these results as consistent with the idea that the decrease in the scoring rate for the second team, which is the one more likely to be behind in the partial score, can be mainly attributed to an increase in misses by the kicker rather than to an increase in saves by the goalkeeper of the opposing team.

REFERENCES

Arellano, Manuel, and Olympia Bover. 1995. "Another Look at the Instrumental Variable Estimation of Error–Components Models." *Journal of Econometrics,* 68(1): 29–51.

Arellano, Manuel, and Raquel Carrasco. 2003. "Binary Choice Panel Data Models with Predetermined Variables." *Journal of Econometrics,* 115(1): 125–157.

Fernandez-Val, Ivan. 2009. "Fixed Effects Estimation of Structural Parameters and Marginal Effects in Panel Probit Models." *Journal of Econometrics,* 150(1): 71–85.

Honoré, Bo, and Arthur Lewbel. 2002. "Semiparametric Binary Choice Panel Data Models without Strictly Exogenous Regressors." *Econometrica,* 70(5): 2053–2063.

Honoré, Bo, and Ekaterini Kyriazidou. 2000. "Panel Data Discrete Choice Models with Lagged Dependent Variables." *Econometrica,* 68(4): 839–874.

TABLE A.1– DETERMINANTS OF MISSES AND SAVES

Panel A: Proportions of Penalties Scored, Saved and Missed

	N (penalties)	Scored	No Scored: Saved	No Scored: Missed
First team	234	75.6%	17.1%	7.3%
Second team	220	68.6%	16.8%	14.6%
	Difference:	-7.0%	-0.3%	+7.3%

Panel B: Panel Data Analysis for Misses and Saves

	Misses [1]	Misses [2]	Saves [3]	Saves [4]
Constant	-1.57***	-1.70***	-0.98**	-0.67**
	(3.01)	(2.57)	(2.15)	(2.23)
Partial Score -1	**0.09*** **	**0.17*** **	**-0.17**	**-0.12**
	(2.65)	(2.85)	(0.44)	(0.36)
Partial Score 0	0.46	0.33	0.15	0.06
	(0.33)	(0.12)	(0.27)	(0.16)
Partial Score +1	0.09	0.17	0.88	-0.47
	(0.93)	(0.77)	(0.38)	(0.36)
Second Team	0.27	0.15	0.10	-0.22
	(0.67)	(0.49)	(0.55)	(0.68)
Partial Score -1 x Second Team	-0.009**	-0.005**	0.006	0.015
	(2.06)	(1.90)	(1.03)	(0.70)
Penalty Kick Importance	1.05	1.17	0.44	0.83
	(1.12)	(1.10)	(0.88)	(0.47)
Round Fixed Effect	No	Yes	No	Yes
Goalkeeper Fixed Effect	No	Yes	No	Yes
N	454	454	454	454

Notes: Misses by kicker are penalty kicks shot to the upright posts, the horizontal crossbar or outside the goal. Saves are penalty kicks stopped by the goalkeeper. All the specifications include Field, Competition and Match type fixed effects, as well as controls for "Potential Winner" and "Potential Loser" penalty kicks. Absolute, choice–biased sampling, heteroskedastic-consistent, *t*–statistics are reported in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

APPENDIX B. A Theoretical Framework

In this appendix we study a simple model that is consistent with the empirical evidence, *always* predicts a first-mover advantage, and relies on a reference point associated with the partial score (score at the time of kicking). After the presentation of the model, various extensions are discussed.

Let (s, r) , with $s \in \mathbb{Z}$, $r \geq 1$, denote the *score at the end of a round* r. The score measures the difference in goals between the team that kicks first F and the one that kicks second D. A round involves one penalty kick for F and one for D. The total number of rounds is n. The *partial score* for a team $\alpha \in \{F, D\}$ in round r is the difference between the goals scored by α and those scored by the opponent, immediately before team α is about to take its penalty kick in round r. That is, for team F the partial score at r is $(s, r-1)$, while, for team D, it is $(-s-x, r-1)$, where $x = 1$ if F scores in round r and $x = 0$ otherwise. In what follows we will use the terms team and player indistinctly.

Denote by $p \in [0,1]$ the probability of player α scoring a goal when the partial score is tied or positive for him, and by $q \in [0,1]$ the probability of player α scoring a goal when he is behind in the partial score by at least one goal. Under psychological pressure $p > q$, while under no psychological pressure $p = q$.

For any given (s, r) with $r < n$ there are exactly four possible outcomes at the end of round r: (i) both players score a goal, (ii) the first scores and the second fails, (iii) the first fails and the second scores, and (iv) both fail. The probability vectors associated to these outcomes depend on (s, r). There are three possible cases:

1. If $s = 0$, then $(p \cdot q, p(1-q), (1-p)p, (1-p)^2)$;
2. If $s > 0$, then $(p \cdot q, p(1-q), (1-p)q, (1-p)(1-q))$;
3. If $s < 0$, then $(q \cdot p, q(1-p), (1-q)p, (1-q)(1-p))$.

To simplify notation we write $a = p \cdot q$, $b = p(1-q)$, $c = (1-p)p$, $d = (1-p)^2$, $e = (1-p)q$, and $f = (1-p)(1-q)$. The above defines a Markov chain. Since we are interested in rank-order tournaments, we need to refine the notion of maximum and minimum scores. If n is even, the maximum and minimum scores are $\frac{n}{2} + 1$ and $-(\frac{n}{2}+1)$, while if n is odd the maximum and minimum scores are $\frac{n+1}{2}$ and $-(\frac{n+1}{2})$. The state space is formed by all possible scores $S = \{s_{max}, s_{max}-1, \ldots, -1, 0, 1, \ldots, s_{min}-1, s_{min}\}$ with s_{max} and s_{min} defined as above. Typical elements of S are denoted by s, t or $s_0, s_1, \ldots, s_n$. The transition matrix P follows from the single-step transition probabilities p_{st}:

$$p_{00} = a + d;\ \ p_{01} = b;\ \ p_{0,-1} = c$$

$$s \in \{s_{min}, s_{max}\},\ \ p_{ss} = 1$$

$$s \in S \setminus \{0, s_{min}, s_{max}\},\ \ p_{ss} = a + f$$

$$s \in S \setminus \{0, s_{min}, s_{max}\},\ \ p_{s,s+1} = p_{-s,-s-1} = b$$

$$s \in S \setminus \{0, s_{min}, s_{max}\},\ \ p_{s,s-1} = p_{-s,-s+1} = e$$

and $p_{st} = 0$ otherwise. The initial distribution μ puts all the probability mass in state 0. Denote by $T(n, P)$ the n-round sequential tournament between F and D with transition matrix P. Denote by $p_{st}^{(n)}$ the (s, t) entry in the n-th power of the transition matrix P. Since the Markov chain is stationary, $p_{st}^{(n)}$ represents the probability of reaching state t starting from state s in n rounds.

The probability that team F wins the n-round sequential tournament $T(n, P)$ is:

$$W(F, n) = \sum_{s=1}^{s=s_{max}} P(s, n),$$

with $P(s, n)$ denoting the probability of a final score s. To calculate $P(s, n)$ we have to correct for the probability of reaching a *final state* in some previous round. A final state is a pair (s, r) where there is no possibility of turning around the sign of the score s in the remaining time $n - r$. Then we have that:

$$P(1, n) = p_{01}^{(n)} - p_{02}^{(n-1)} p_{21},$$

$$P(s_{max}, n) = p_{0,s_{max}-1}^{(n+1-s_{max})} p_{s_{max}-1,s_{max}} + p_{0s_{max}}^{(n+1-s_{max})},$$

and for $1 < s < s_{max}$:

$$P(s, n) = p_{0,s-1}^{(n+1-s)} p_{s-1,s} + p_{0s}^{(n+1-s)} - p_{0,s+1}^{(n-s)} p_{s+1,s}.$$

In principle, these probabilities can be obtained using standard matrix algebra. The probability of team D winning at the end of the n-round contest $W(D, n)$ is obtained analogously.

Denote by $W(\alpha, r)$ the probability that α is either ahead of its opponent at the end of round $r \leq n$ or has already won the tournament by then. We are now ready to derive convenient formulations for $W(F, r)$ and $W(D, r)$.

Proposition 1. Let $T(n, P)$ be an n-round sequential tournament. Then, for every $r \leq n$, $W(F, r) = \frac{b}{b+c}(1 - p_{00}^{(r)})$ and $W(D, r) = \frac{c}{b+c}(1 - p_{00}^{(r)})$.

Proof: Take any path ending in state $s > 0$ in round r, and denote it by $s_0 s_1 \cdots s_{r-1} s_r$ with $s_0 = 0$ and $s_r = s$. The probability measure of such path is $p_{s_0 s_1} \cdots p_{s_{r-1}, s_r}$. We distinguish two cases:

(1) First, the path $s_0 s_1 \cdots s_{r-1} s_r$ does not reach a final state s' in some previous round $h < r$. We construct a unique symmetric path to the original one, ending in state $-s$. If $s_{r-1} = 0$, stop. Otherwise, proceed backwards until reaching a k, $0 \leq k \leq r-1$, such that $s_k = 0$. Clearly, such a k exists. Then, for every $l \geq k$ write $s'_l = -s_l$, and write $s'_l = s_l$ otherwise. It is immediate that the constructed path $s'_0 s'_1 \cdots s'_{r-1} s'_r$ starts with $s'_0 = 0$, ends in $s'_r = -s$, does not go through any final state, and has an associated probability measure of $p_{s'_0 s'_1} \cdots p_{s'_{r-1}, s'_r}$, where $p_{s'_l, s'_{l+1}} = p_{s_l, s_{l+1}}$ for every $l \neq k$, while $b = p_{s_k, s_{k+1}} \geq p_{s'_k, s'_{k+1}} = c$. That is, the difference in the probability measures between the two paths is $(b - c) p_{s_0 s_1} \cdots p_{s_{k-1}, s_k} p_{s_{k+1}, s_{k+2}} \cdots p_{s_{r-1}, s_r}$.

(2) Second, the path $s_0 s_1 \cdots s_{r-1} s_r$ does reach a final state s' in a previous round $h < r$. First, we modify the path $s_0 s_1 \cdots s_{r-1} s_r$ to correct for the sub-path following the final state s' by writing $s_h = s_{h+1} = \cdots = s_{r-1} = s_r = s'$, with associated probability measure $\bar{p}_{s_l, s_{l+1}} = p_{s_l, s_{l+1}}$ whenever $s_l \leq h - 1$, and $\bar{p}_{s_l, s_{l+1}} = 1$ whenever $s_l > h - 1$. Second, we apply the same argument as before to the modified path to show that there exists a unique symmetric path ending in final state $-s'$, and where the difference in the probability measures between the two paths is $(b - c) \bar{p}_{s_0 s_1} \cdots \bar{p}_{s_{k-1}, s_k} \bar{p}_{s_{k+1}, s_{k+2}} \cdots \bar{p}_{s_{r-1}, s_r}$.

Consequently, it is immediate that there exists a probability mass $\gamma(r)$ such that $W(A, r) = b\gamma(r)$ and $W(B, r) = c\gamma(r)$. Note that by definition of final states, $p_{00}^{(r)}$ does not reach any final state in some previous round $h \leq r$. Now since $W(A, r) + W(B, r) + p_{00}^{(r)} = 1$, it follows that $\gamma(r) = \frac{1}{b+c}(1 - p_{00}^{(r)})$, and hence $W(A, r) = \frac{b}{b+c}(1 - p_{00}^{(r)})$ and $W(B, r) = \frac{c}{b+c}(1 - p_{00}^{(r)})$.$\square$

Since under no psychological pressure $p = q$ implies $b = c$, it directly follows from

Proposition 1 that $W(F,n) = W(D,n)$. On the other hand, under psychological pressure, that is when $p > q$, we have that $b > c$, and hence Proposition 1 implies that $W(F,n) > W(D,n)$. We summarize the above in the following corollary:

Corollary. Let $T(n,P)$ be an n-round sequential tournament. Then, for every $p,q \in [0,1]$, and for every $r \leq n$, if $p = q$ then $W(F,r) = W(D,r)$, and when $p > q$, $W(F,r) > W(D,r)$.

The main merit of the model is that it only takes two parameters to predict a greater probability of the first team winning. This simplicity involves strong assumptions and is not without limitations which further generalizations of the model may address. For instance, in order to capture the possibility of order-dependent technologies, it is possible to consider the difference $p - q$ to be team-specific or alternatively a four-parameter model (p_α, q_α), $\alpha = F, D$. For example, a partial score of 0 is better news for the second team than for the first team, and this may have an impact on performance. These considerations will readily enrich the model and can be easily incorporated. In the latter case, for instance, it can be shown that the advantage of one team over the other depends on the parameter range of the (p_α, q_α) values, and that it is possible to find parameter values where *both* the first team has an advantage over the second and vice versa. Examples are available upon request.

Finally, the assumption that all team players use the same (p,q) technology can be generalized to introduce heterogeneity in player quality: $\{(p_h, q_h)\}_{h \in Q}$, where Q is the set of players of a team ordered by quality. This opens up the possibility of investigating the strategic placement of players throughout the tournament, a decision that is taken before the toss out. We have studied this extension by having the two teams with (the same) two types of players differing in their (p,q) technology compete in a tournament with just 2 rounds. We have found that, even in this simple case, it is possible to find parameters within an empirically sensible range of values that make *each* of the 4 possible combinations of the players a Nash equilibrium. That is, in theory, any combination could be optimal. This means that without precise information regarding the technologies of each player, it is not possible to have a sharp, testable implication on the strategic allocation of players.

For instance, consider the two teams F and D competing in a tournament with only 2 rounds. Each team has two players: H and L. Player H plays with technology

(p_h, q_h), $p_h \geq q_h$, and player L plays with (p_l, q_l), $p_l \geq q_l$, where $p_h > p_l$. Assume that one of the players is vulnerable to psychological pressure, and hence plays with $p_i > q_i$, and the other player is not vulnerable, and hence he plays with $p_j = q_j$; $i, j \in \{H, L\}$, $i \neq j$. Denote by $(\beta_\alpha, \gamma_\alpha)$ the order in which team $\alpha \in \{F, D\}$ places the players H and L in the 2 rounds. For example, $(\beta_F, \gamma_F) = (L, H)$ means that the first team F places player L in round 1 and player H in round 2. Let $W(\alpha, (\beta_F, \gamma_F), (\beta_D, \gamma_D))$ denote the probability of winning for team α when the first team chooses the order (β_F, γ_F) and the second team chooses the order (β_D, γ_D). It is immediate to show that:

$$W(F, (H, L), (H, L)) - W(F, (L, H), (H, L)) = (p_h - p_l)(1 - q_l)(p_h - q_h) \quad (1)$$

$$W(F, (H, L), (L, H)) - W(F, (L, H), (L, H)) = (p_h - p_l)(1 - q_h)(p_l - q_l) \quad (2)$$

$$\begin{aligned} W(D, (H, L), (H, L)) - W(D, (H, L), (L, H)) &= (p_h - p_l)(1 - p_h)(p_l - q_l) \\ &+ p_h(1 - p_l)(q_h p_l - q_l p_h) \quad (3) \end{aligned}$$

$$\begin{aligned} W(D, (L, H), (H, L)) - W(D, (L, H), (L, H)) &= (p_h - p_l)(1 - p_l)(p_h - q_h) \\ &+ p_l(1 - p_h)(q_h p_l - q_l p_h) \quad (4) \end{aligned}$$

It then follows that:

(i). When player L is the vulnerable player, and hence $p_h = q_h > p_l > q_l$, equation (1) is equal to 0, and equations (2), (3), and (4) are all strictly positive. Hence, there are two pure strategy Nash equilibria: $((H, L), (H, L))$ and $((L, H), (H, L))$.

(ii). When player H is the vulnerable player, and hence $p_h > q_h$ and $p_h > p_l = q_l$, equation (1) is strictly positive, equation (2) is equal to 0, equation (3) is strictly negative, and the sign of equation (4) is equal to the sign of $A = (p_h - p_l + p_h p_l (p_l - 1))$.[4] Consequently, $((H, L), (L, H))$ is a pure strategy Nash equilibrium, and $((L, H), (L, H))$ will be another pure strategy Nash equilibrium when $A \leq 0$.

Therefore, even in this simple version of a penalty shootout, the four possible placements of the players could be a pure strategy Nash equilibrium. Hence, without

[4] Empirically plausible parameters for each potential case can be readily found. For example, $A < 0$ when $p_h = 0.8$ and $p_l = 0.7$; $A > 0$ when $p_h = 0.85$ and $p_l = 0.65$; and $A = 0$ when $p_h = 0.65/0.7725 \simeq 0.841$ and $p_l = 0.65$.

precise information on who is the vulnerable player (and when H is the vulnerable player, on the specific values of p_h and p_l) no testable prediction on the strategic allocation of players obtains.