Journal of Agribusiness 38, 2 (Fall 2020)

Sample Selection Bias in Hedonic Pricing Models of Thoroughbred Broodmares

Matthew Muntifering and John N. Ng'ombe

An issue with ordinary least squares estimations in hedonic pricing model literature is that they do not account for sample selection bias. In broodmare auctions, the purchase decision and whether a price is realized or zero is endogenous. This paper contributes to the hedonic broodmare price analysis literature by implementing the Heckman sample selection regression to estimate a hedonic pricing model using data from the January 2020 Keenland Sale. Many published papers do not accommodate this selection process and may have biased coefficients. This paper further contributes methodologically to the thoroughbred broodmare literature by being the first to deliver a useful empirical application of a Bayesian Heckman selection model. The broodmare prospect, age, square of age, domestic status, and the day of the session are significant for broodmare pricing. These may be implemented within a profit-maximizing purchasing and breeding strategy.

Key words: Bayesian Methods, Hedonic Models, Sample Selection Bias, Thoroughbred Broodmares

American Pharoah (2015) and Justify (2018) claimed the Triple Crown of Thoroughbred Racing after a draught since Affirmed (1978) took the title. Thirteen racers have won the prestigious award in history, with some earning more than $10 million in today's dollars.

The American Horse Council Foundation (AHCF) estimates 7.2 million horses are in the United States consuming 32 million acres of owned land and another 49 million acres of leased land. Further, AHCF estimates that the direct effect of the horse industry on the domestic economy is $50 billion. The direct employment total reaches near one million jobs earning roughly $38 million in various accounts. The ripple effects from this industry are estimated to be $122 billion and 1.7 million jobs. The high stakes associated with thoroughbred horseracing makes understanding the determinants of prices economically important for both buyers and sellers. According to Chizum and Wimmer (1997) and Wimmer and Chizum (2006), asymmetric information and adverse selection prevail in Thoroughbred markets. These issues may make the empirical findings of this paper useful for increasing market efficiency.

Vickner (2018) points out that, among all the hedonic price models applied to Thoroughbreds, majority study yearlings. Only Neibergs (2001), Maynard and Stoeppel

---

Matthew Muntifering is a Ph.D. candidate and John N. Ng'ombe is a visiting assistant professor, both in the Department of Agricultural Economics and Rural Sociology, Auburn University, Auburn, Alabama.

(2007), and Dority et al. (2016) focus on broodmares. Chezum and Wimmer (1997), Vickner and Koch (2001), Robbins and Kennedy (2001), Wimmer and Chezum (2006), Parson and Smith (2008), Plant and Stowe (2013), and Marion and Stowe (2016) all focus predominately on yearlings. Stowe and Ajello (2010) perform ordinary least squares (OLS) in their hedonic pricing model of stud fee determinants, while Stowe (2013) extends this model to include fixed effects. Taylor et al. (2006) uses the Heckman model within the horse literature on data about Quarter horses. This paper contributes to the relatively scarce literature on broodmare pricing by applying a Bayesian Heckman selection model to account for sample selection bias. Failure to consider the endogenous selection process biases parameter estimates and would misinform prospective buyers, sellers, bloodline agents, policymakers, and fellow scientists.

Vickner (2018) suggests that a Bayesian Heckman selection model would be an interesting future contribution to Thoroughbred literature. Heckman et al. (2013) provide the details for extending the classical selection model to a Bayesian framework. Ng'ombe and Boyer (2019) point out that a Bayesian inference is desirable as it is exact for any sample size.

Our results suggest there is significant sample selection in broodmare pricing which promotes using sample selection techniques when modeling hedonic prices of Thoroughbred broodmares. Broodmare prosect, age, square of age, domestic status, and the day of the session were found to have meaningful effects on the average broodmare selling price.

The rest of the paper is organized as follows. The next section presents methods employed, which are then followed by the data section. Description of data is followed by results and discussion. The last section concludes the paper.

Conceptual Model

A Heckman sample selection framework is applied to a hedonic pricing model for broodmares at the January 2020 Keenland Sale. The Heckman selection model is a two-step procedure. The initial model is

$$y_i = X_i\beta + \varepsilon_i \tag{1}$$

where y_i is the dependent variable, X_i denotes explanatory variables, β denotes parameters to be estimated, and ε_i denotes the error term assumed to follow a Gaussian distribution with mean zero and constant variance. The second equation is shown as

$$Z_i\gamma + \epsilon_i > 0 \tag{2}$$

where Z_i denotes independent variables that explain selection and may overlap with those that are in X_i, γ denotes parameters to be estimated in the selection equation, and ϵ_i is the error term that is assumed to follow the standard Gaussian distribution. Because the two models are related through the error terms, the correlation that exists between them is shown as

$$corr(\varepsilon_i, \epsilon_i) = \rho \tag{3}$$

It is worth pointing out that equations (1) through (3) can be estimated so long samples are larger than zero using various models that are Tobit-type. Among them, we select the Heckman selection model out of preference and because it provides consistent, asymptotically efficient estimates of the parameters (Wooldridge, 2002, Green, 2003; Xu et al., 2017). Using this information, the likelihood function of the Heckman selection model is

$$L = \prod_c \times \left[1 - \Phi(Z_i\gamma)\right] \cdot \prod_{uc} \Phi\left(\frac{Z_i\gamma + \rho(y_i - X_i\beta)/\sigma}{\sqrt{1-\rho^2}}\right) \cdot \frac{\phi((y_i - X_i\beta)/\sigma}{\sigma} \tag{4}$$

where $\prod_c, \prod_c$ are, respectively, products over censored and uncensored samples; and ϕ and Φ are the standard Gaussian and cumulative distributions, respectively. The frequentist approach would involve using maximum likelihood estimations that would require maximizing the log-likelihood form of equation (4) to obtain Heckman selection model parameter estimates. More details of doing so can be found in Heckman (1979), Wooldridge (2002), and Greene (2003).

This paper employs a Bayesian approach to estimate the model. The Bayesian approach allows us to make robust and informative statements about our findings by using credible intervals—equivalent to confidence intervals in frequentist-based econometrics (Gelman et al., 2013; Ng'ombe et al., 2020). A Bayesian credible interval is interpreted as the probability that a given value falls in that range given the model and data, notwithstanding the scarcity of data (McElreath, 2020). This is more intuitive especially that statistical inference is valid regardless of the sample size (Gelman et al., 2013; Ng'ombe and Boyer, 2019; McElreath, 2020). Motivated by these observations and the small nature of our dataset, this paper employs a Bayesian Heckman selection model.

The parameters to be estimated can be vectorized as $\kappa = [\beta, \gamma, \rho, \tau]'$, where β, γ, and ρ are as previously defined and τ precision (i.e., $1/\sigma^2$). Next, we need the likelihood function for the model (equation (4)) $p(y|\kappa)$ and the prior distribution of the parameters

$p(\kappa)$ to obtain the posterior distribution of the parameters by Bayes' Theorem. The posterior distribution can be specified as

$$(5) \qquad p(\kappa|y) \propto p(y|\kappa) \cdot p(\kappa)$$

Estimation of equation (5) is not trivial. Recent computer revolution (Ng'ombe and Brorsen, 2020) has resulted in the powerful tool of the Markov chain Monte Carlo (MCMC) that makes it easier to implement. More details about MCMC can be found in Gelman et al. (2013) and Gill (2013).

Empirical Model

Empirically, we model broodmare selling prices by rewriting equation (1) as

$$(6) \qquad y = \sum_{b=1}^{5} \beta_b X_b + X_g + \sum_{m=1}^{5} \beta_m X_m + u_1$$

where y is the natural log of the broodmare selling price; X_b represents breeding characteristics; and X_g, and X_m are genetic, and market characteristics, respectively. The error term u_1 is

$$(7) \qquad u_1 \sim N(0, \sigma)$$

The covariates in this model are motivated by Maynard and Stoeppel (2007) and Dority et al. (2016). They argue that breeding, genetic, and market characteristics are relevant in explaining Thoroughbred broodmare auction prices. The breeding characteristics in the model are a dummy = 1 if a broodmare is a prospect, age is in years, color dummy =1 if the broodmare is black. Other variables include sire earning and sire stud fee. Stowe (2013) finds that sire stud fee is highly explained by the progeny sale price.

Poerwanto and Stowe (2010) find a positive relationship between the number of foals produced by a sire and the sire's yearlings' average selling prices. Therefore, sire representation is included into the model as a genetic characteristic. Market characteristics include a dummy = 1 if the sire is domestic and dummies for the days of the auction. In the case of broodmare auctions, each individual broodmare is not sold, and the price is only observed if the selection equation is satisfied, that is

$$(8) \qquad \gamma_s z_s + u_2 > 0$$

where z_s is a dummy = 1 if the sire has won a Triple Crown race. This can be the Kentucky Derby, Preakness Stakes, or Belmont Stakes. The error term u_2 is

(9) $$u_2 \sim N(0,1)$$

(10) $$corr(u_1, u_2) = \rho$$

In terms of priors, we impose the usual conjugate and diffuse priors for all the parameters. That is

(11) $$\beta \sim N(b_0, B_0), \quad \tau \sim Gamma(\delta_0, \delta_1)$$

where b_0 and B_0 are hyper-priors assumed to be 0 and 1,000, respectively; and δ_0 and δ_1 are shape and scale hyper-priors which we set to 0.001 so as to impose a higher prior variance. Because we impose larger prior variances, it implies that our priors would have a negligible effect on our results. We then use MCMC to sample from the posterior. Our simulations were conducted in Stata software using the Random-walk Metropolis-Hastings algorithm (StataCorp, 2019). Our MCMC techniques involved two Markov chains with a burn-in phase of 5,000 to allow the chains to forget their initial regions (Gelman et al., 2013; Ng'ombe et al., 2020). To obtain high-quality posterior distributions, the total number of iterations were 25,000 per chain. To determine whether our chains converged to their target posterior distributions, for brevity, we checked trace and autocorrelation plots of the Markov chains. Trace plots with good mixing indicate successful convergence while autocorrelation plots that die away are by convention a sign of successful convergence of the MCMC (Gelman et al., 2013; Gill, 2013; Ng'ombe et al., 2020).

Estimating hedonic pricing models via OLS in the existence of this error correlation causes estimates to be biased because they violate the assumption of random sampling. Dority et al. (2016) does not account for sample selection processes. Heteroskedasticity may also arise. Maynard and Stoeppel (2007) account for heteroskedasticity using a Box-Cox transformation. Marion and Stowe (2016) use a Breusch Pagan test and reject the null hypothesis of heteroskedasticity. Nonetheless, our paper's methodological contribution is an empirical application of a Bayesian Heckman selection model applied to Thoroughbred broodmare auctions.

Data and Descriptive Statistics

Data on broodmare sales prices and characteristics were obtained from the January 2020 Keenland Sale at Keenland Association in Lexington, Kentucky. The sire nationality and performance data were obtained from the Blood-Horse Stallion Register and matched to corresponding broodmares. Table 1 presents the descriptive statistics.

Table 1. Descriptive Statistics.

Variable	Mean	Standard Deviation	Minimum	Maximum
Price ($)	27670.23	63181.87	0	640000
Prospect	0.511	0.5	0	1
Age in years	5.742	2.519	2	16
Black	0.468	0.499	0	1
Sire stud fee ($)	63812.34	67347.84	2000	250000
Sire earnings ($)	2080000	2260000	32400	1.05E+07
Representation	10.439	6.488	1	25
Domestic sire	0.95	0.217	0	1
Triple Crown	0.168	0.374	0	1
Session 1	0.225	0.418	0	1
Session 2	0.261	0.439	0	1
Session 3	0.187	0.39	0	1
Session 4	0.171	0.377	0	1
Session 5	0.154	0.361	0	1
Observations	524			

The sample contains 524 unique broodmares with 323 (61.6%) of those being sold. The other sale prices are recorded as zero. The average price conditional on being sold is $44,889 and ranges from $1,000 to $640,000. Broodmare prospects account for roughly 51% of the sample and average prospects have a price of $28,079 versus $27,241 of the average non-prospects. The difference, however, is statistically insignificant, with a p-value of 0.879 as shown in Figure 1.

The average broodmare in the sample is approximately six years of age. The average sire earned $2.08 million, has a stud fee of about $63,812, and is being represented 10 times. Ninety-five percent of the sires are domestic, and 16.8% of the total sires have won a Triple Crown race. A broodmare of a domestic sire on average sold for $28,379.92

versus $14,076.92. The difference is not statistically significant, with a p-value of 0.261. Figure 2 shows this comparison.

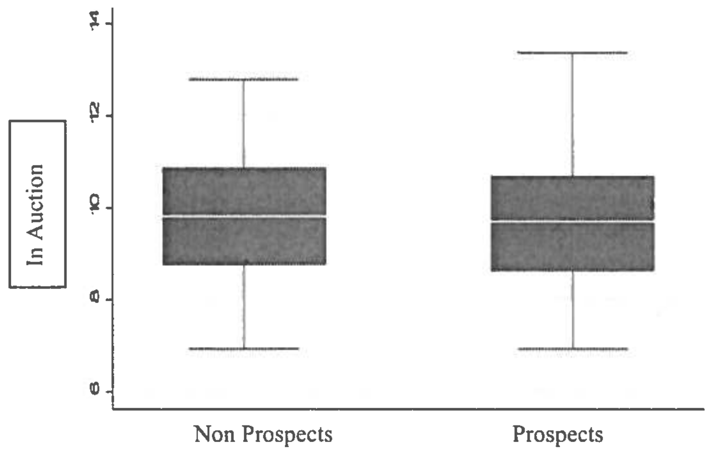

Figure 1. Price of Non-Prospects vs. Prospects.

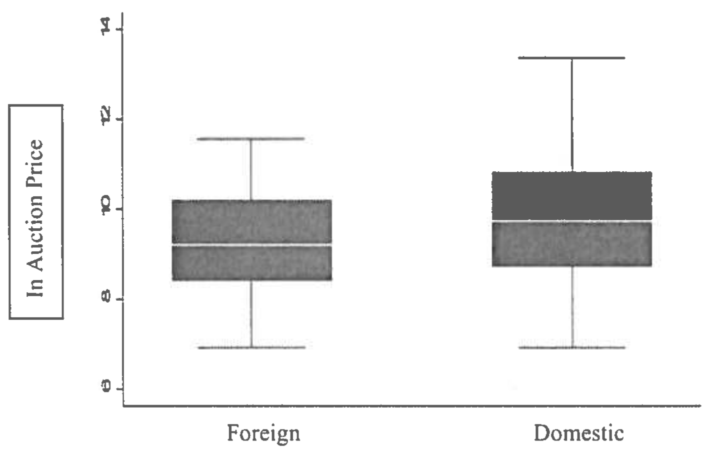

Figure 2. Price Difference Between Sire's Domestic Status.

Table 2 presents a pairwise comparison of mean price across different auction sessions. There are statistically significant differences in prices between session 1 to sessions 3, 4, and 5, respectively, as well as between 2 and 3, 4, and 5, respectively. The signs on each of these differences are negative and have management implications. Buyers may be able to receive a discounted price if they are willing to delay their purchase by attending a later auction. This inference is consistent whether using Bonferroni, Sidek, Sheffe, Tukey, SNK, Duncan, or Dunnet adjustments. Figure 3 visualizes this relationship. Dority et al. (2016) find that the longer buyers are willing to wait, the lower price that they can receive.

Table 2. Pairwise Comparisons of Mean Price Across Sessions with Bonferroni Adjustment.

Session	Contrast	Standard Error	P > t
2_vs_1	6935.513	7710.558	1
3_vs_1	-21575.94	8390.539	0.104
4_vs_1	-28729.55	8591.845	0.009
5_vs_1	-31674.73	8858.497	0.004
3_vs_2	-28511.46	8122.274	0.005
4_vs_2	-35665.06	8330.065	0
5_vs_2	-38610.25	8604.833	0
4_vs_3	-7153.605	8963.17	1
5_vs_3	-10098.79	9219.088	1
5_vs_4	-2945.185	9402.672	1

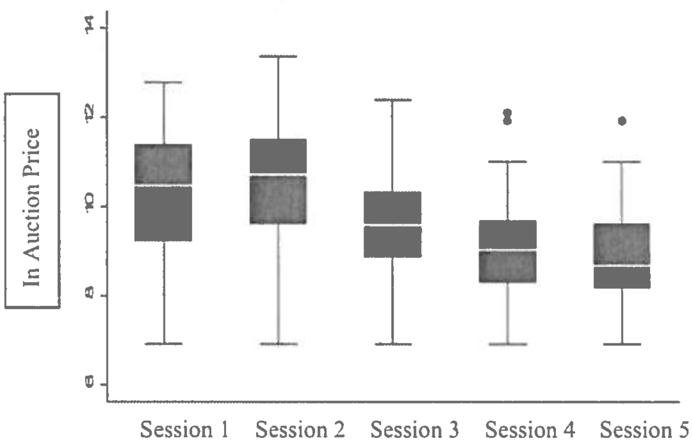

Figure 3. Price Over Session.

Results and Discussion

To show that our MCMC chains converged to their target posterior distributions, we show convergence diagnostics for variable *age* only to save space. Convergence diagnostics are shown in Figure 4. As shown in Figure 4, the trace plot indicates that each MCMC chain exhibits good mixing, which suggests successful convergence. The autocorrelation plot indicates that the terms of the chain decline toward zero as lags are increased, which also suggests successful convergence. The histogram and density plots are graphical representations of the posterior distribution of the coefficient of age for each chain.

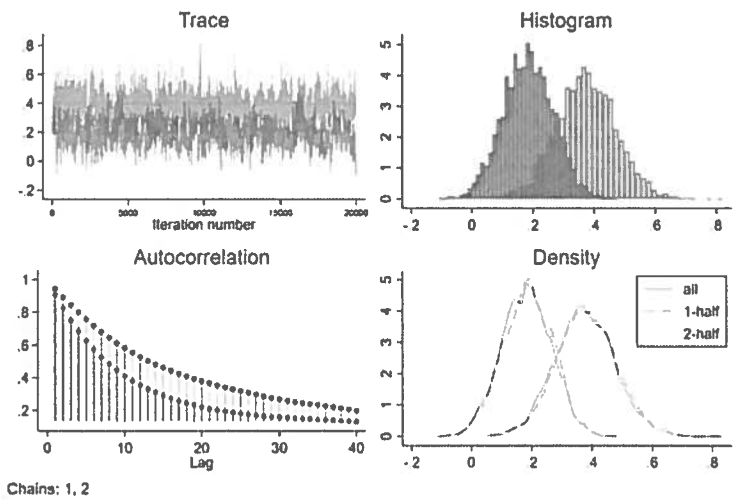

Figure 4. Convergence Diagnostic Plots for Variable Age.

Table 3 presents the posterior summary statistics from the Bayesian Heckman selection model. In terms of significance, results in Table 3 are significant if the 95% credible interval does not include the value of zero. Broodmare prospect, age, square of age, domestic status, and the session are found to have significant effects on average broodmare selling prices. *Ceteris paribus*, a broodmare prospect would peg a price of 54% more than otherwise. In terms of age, age of the broodmare has a positive effect on its selling price though the variable age square has a negative effect. This finding shows an inverted-U relationship between age and broodmare price. This implies that younger broodmares would be costlier, but, as they become older (i.e., with the turning point of 7.34 years), their price would eventually decline. This is plausible because a younger broodmare's investment portfolio would be expected to be higher in its younger age.

Table 3. Bayesian Heckman Results with Sample Selection.

Variable/Statistic	Mean	Standard Deviation	95% Credible Interval	
Dep. variable: log of price of broodmare				
Prospect	0.536	0.385	0.125	1.012
Age in years	0.279	0.164	0.043	0.539
Square of age	-0.019	0.009	-0.039	-0.004
Black	-0.041	0.306	-0.434	0.357
Log of sire earnings	0.018	0.058	-0.089	0.123
Log of sire fee	0.149	0.147	-0.072	0.353
Representation	0.017	0.025	-0.022	0.058
Domestic sire	0.966	0.274	0.556	1.394
Session 2	0.005	0.236	-0.347	0.265
Session 3	-1.145	0.505	-2.421	-0.639
Session 4	-1.518	0.198	-3.109	-1.216
Session 5	-2.266	1.006	-4.450	-1.327
Intercept	7.491	0.619	6.908	8.311
Siretcwinner	-0.077	0.124	-0.318	0.720
Intercept	0.293	0.059	0.180	0.409
Model statistics				
Athrho	-1.304	0.439	-2.238	-0.381
Log sigma	0.520	0.173	0.235	0.724
Rho	-0.829	0.172	-0.952	-0.363

Sire stud fee results are positive, which matches our expectation that a sire's stud fee is expected to increase the selling price of the broodmare. While this result is qualitatively consistent with Neiberg (2001) and Dority et al. (2016), our result is not significant. A domestic sire is associated with a 96.6% on average broodmare selling price relative to non-domestic ones, all else equal. This result ranges between 56% and 139% with a 95% probability. For managers, this means that including a domestic sire in your bloodline may increase future broodmare returns.

The result for sire representation also has management relevance. When deciding the number of mares for a sire to service, managers must trade off short-term earnings with the possibility of decreasing future value of the sire due to the possibility of an inadequate foal. Based on the insignificant results, the relationship between the sire representation and broodmare price is inconclusive.

The dummies indicating the session are all statistically significant except for the session 2 dummy. This is consistent with Dority et al. (2016) who find what they describe as buyer fatigue in these auctions. They point out that it is customary for the highest quality broodmares to be auctioned the earliest and that there may be a psychological notion that the best lot has been sold. This evidence suggests potential buyers who wait until later auction sessions incur additional risk. Managers and potential buyers should seek to attend the earliest sessions. Notice that results presented in Table 3 indicate significant sample selection. This is evidenced by the significant ρ value in the last row. This result shows that ρ lies between -0.363 and -0.952 with a 95% chance which implies that a sire winning a Triple Crown race is negatively correlated with the price of a broodmare. This result provides further evidence for the necessity to model sample selection processes. The corrected model gives managers better estimates of possible returns to sire earnings in the breeding market.

Without accounting for sample selection in such hedonic models as presented here, one would introduce bias in their results. To show that results in Table 3 are more appropriate, we estimated another model in which we tested for sample selection formally. In this model (results presented in Table A1 in appendix), we imposed $\rho = 0$ to imply that both the selection and outcome models can be estimated separately. Upon estimation, we computed a Bayes factor—a ratio of the model's marginal likelihoods (Gelman et al., 2013; Jarosz and Wiley, 2014). We found the inverse of the Bayes factor of 20.43 which suggests that model results in Table 3 are more appropriate than those reported in Table A1 (Gelman et al., 2013; Jarosz and Wiley, 2014). Stated differently, this finding suggests a very strong preference for the presence of sample selection in these data and, therefore, a Bayesian Heckman selection model.

Conclusion

This paper contributes to Thoroughbred literature by being the first to estimate a Bayesian Heckman sample selection model to the January 2020 Keenland Sales data to account for the sample selection process underlying broodmare sales. Given the documented asymmetric information and adverse selection in the Thoroughbred industry, an unbiased hedonic pricing model of broodmares stands to inform buyers of the characteristics important in determining price. This evidence may alleviate some inefficiency associated with the information gap and market failure. In an industry with roughly $175 billion in economic impact, the welfare loss from this inefficiency is likely nontrivial. Failure to account for the selection process prevalent by omitting broodmares with prices of zero from the sample will bias coefficient estimates and misinform prospective buyers, breeders, and racers. This estimation procedure, combined with the exactness of Bayesian inference, can be used in future Thoroughbred hedonic pricing analyses, whether for broodmares or yearlings.

A broodmare prospect, age, square of age, domestic status, and the day of the auction session are all significant factors in broodmare prices. Managers can implement this information into their buying and breeding strategies. Further studies may examine other variables, such as dam characteristics, sprinting speed, or breeder characteristics for significance, but should be aware of the modelling issues addressed in this paper. Additionally, it is noteworthy to mention that the current study uses few control variables due to data limitations. Admittedly, this is an important caveat. Thus, future studies using econometric methods employed in this paper should also consider including more control variables than used here.

Appendix

Table A1. Bayesian Heckman Results without Sample Selection.

Variable/Statistic	Mean	Standard Deviation	95% Credible Interval	
Dep. variable: log of price of broodmare				
Prospect	0.015	0.377	-0.462	0.538
Age in years	-0.036	0.288	-0.355	0.339
Square of age	-0.002	0.016	-0.023	0.016
Black	0.144	0.137	-0.137	0.412
Log of sire earnings	-0.100	0.110	-0.239	0.047
Log of sire fee	0.263	0.079	-0.116	0.405
Representation	-0.007	0.019	-0.040	0.024
Domestic sire	0.567	0.217	0.157	1.003
Session 2	0.141	0.178	-0.189	0.489
Session 3	-0.523	0.273	-0.928	-0.073
Session 4	-1.251	0.227	-1.559	-0.827
Session 5	-1.524	0.252	-2.007	10.870
Intercept	8.663	2.308	6.548	10.870
Siretcwinner	0.015	0.377	-0.318	0.282
Intercept	-0.036	0.288	0.178	0.414
Model statistics				
Athrho	0.000	0.009	-0.018	0.020
lnsigma	0.213	0.046	0.129	0.298

References

Chezum, B., and B. Wimmer. (1997). "Roses or lemons: adverse selection in the market for thoroughbred yearlings." *Review of Economics and Statistics* 79(3): 521-526.

Dority, B.L, D. Larreau, and F. Tenkorang. (2016). "Hedonic Price Analysis of Non-Barren Broodmares." *Review of Economic and Business Studies* 9(2): 61-75.

Gelman, A., J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin. (2003). Bayesian data analysis. CRC Press: Cleveland, Ohio.

Gill, J. (2014). Bayesian methods: A social and behavioral sciences approach. Vol. 20. CRC Press: Cleveland, Ohio.

Greene, W.H. (2003). Econometric analysis. Pearson Education: London.

Heckman, J.J. (1979). "Sample selection bias as a specification error." *Econometrica* 47: 153-161.

Heckman, J.J. (1976). "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models." *Annals of Economic and Social Measurement* 5(4): 475-492.

Heckman, J.J., H.F. Lopes, and R. Piatek. (2014). "Treatment effects: A Bayesian perspective." *Econometric Reviews* 33(1-4). Available at https://www.horsecouncil.org/resources/economics/.

Jarosz, A.F., and J. Wiley. (2014). "What are the odds? A practical guide to computing and reporting Bayes factors." *The Journal of Problem Solving* 7(1): 2.

McElreath, R. (2020). Statistical rethinking: A Bayesian course with examples in R and Stan. CRC Press: Cleveland, Ohio.

Marsh, T.L. (2006). "Show quality quarter horse auctions: price determinants and buy-back practices." *Journal of Agricultural and Resource Economics.* pp. 595-615.

Maynard, L.J., and K.M. Stoeppel. (2007). "Hedonic price analysis of thoroughbred broodmares in foal." *Journal of Agribusiness* 25(2): 1-15.

Neibergs, J.S. (2001). "A hedonic price analysis of thoroughbred broodmare characteristics." *Agribusiness: An International Journal* 17(2): 299-314.

Ng'ombe, J.N., and B.W. Brorsen. (2020). "Using Hamiltonian Monte Carlo to Estimate Crop Response Functions with Stochastic Plateaus." (2020). Proceedings of the Selected Presentations at the 2020 Southern Agricultural Economics Association Annual Meeting, Louisville, Kentucky, February 1-4.

Ng'ombe, J.N., and T.A. Boyer. (2019). "Determinants of earthquake damage liability assignment in Oklahoma: A Bayesian Tobit censored approach." *Energy Policy* 131: 422-433.

Ng'ombe, J.N., M.C. Tembo, and B. Masasi. (2020). "Are They Aware, and Why?" Bayesian Analysis of Predictors of Smallholder Farmers' Awareness of Climate Change and Its Risks to Agriculture. *Agronomy*, 10(3): 376.

Parsons, C., and I. Smith. (2008). "The price of thoroughbred yearlings in Britain." *Journal of Sports Economics* 9(1): 43-66.

Plant, E.J., and C.J. Stowe. (2013). "The price of disclosure in the thoroughbred yearling market." *Journal of Agricultural and Applied Economics* 45(2): 243-257.

Poerwanto, D.A., and C.J. Stowe. (2010). "The relationship between sire representation and average yearling prices in the Thoroughbred industry." *Journal of Agribusiness* 28 (1): 61-74.

Robbins, M., and P.E. Kennedy. (2001). "Buyer behaviour in a regional thoroughbred yearling market." *Applied Economics* 33(8): 969-977.

Robert, M., and C.J. Stowe. (2016). "Ready to run: price determinants of thoroughbreds from 2 year olds in training sales." *Applied Economics* 48(48): 4690-4697.

StataCorp. (2019). Stata Statistical Software: Release 16. College Station, Texas: StataCorp LLC.

Stowe, C.J. (2013). "Breeding to sell: a hedonic price analysis of leading Thoroughbred sire stud fees." *Applied Economics* 45(7): 877-885.

Stowe, C.J., and B. Ajello. (2010). "A Hedonic price analysis of differentiated products of unknown quality: Freshman sire stud fees in the thoroughbred breeding industry." *Journal of Agribusiness* 28(1): 19-30.

Taylor, M.R., K.C. Dhuyvetter, T.L. Kastens, M. Douthit, and S.S. Vickner. (2018). "Show quality quarter horse auctions: price determinants and buy-back practices." *Journal of Agricultural and Resource Economics* 31(3): 595-615.

Vickner, S.S., and S.I. Koch. (2001). "Hedonic pricing, information, and the market for thoroughbred yearlings." *Journal of Agribusiness* 19 (2): 1-17.

Wimmer, B.S., and B. Chezum. (2006). "Adverse selection, seller effort, and selection bias." *Southern Economic Journal* (73)1: 201-218.

Wooldridge, J.M. (2010). Econometric analysis of cross section and panel data. MIT Press: Cambridge, MA.