SIMULATION APPENDIX: A MONTE CARLO EXPERIMENT

This appendix describes a small Monte Carlo simulation experiment. The experiment was conducted to assess the importance of the assumption on the economic content of the market specific, time-invariant error component (=random effect). In particular, we wanted to assess the effect of falsely assuming a linearly separable random effect (corresponding to market specific unobservable entry cost), when the true market specific unobservable is due to imperfectly observed market size, and to assess the performance of our simulation estimator when the model is correct. Notice that we assume zero correlation between the explanatory variables and the random effect, i.e., we are not concerned with spurious vs. true state dependence here, even though we control for that with the actual data. The data generating process (DGP) is a panel data probit model which includes three exogenous explanatory variables (one being the constant term, or the fixed entry cost), and unobserved heterogeneity. We assumed zero autocorrelation:

(A.1) $y_{it} = 1((POP_{it} + rh_i)bPROF_{it} - F + e_{it} \geq 0)$ t=1,…,5 i = 1, …., 453.

Where e_{it} ~N(0,1) and h_i ~N(0,0.7071). The variance assumptions correspond to a situation where the variance share of unobserved heterogeneity is 1/3, and therefore ρ=0.5. Sample size corresponds to our actual data. The means and standard deviations of the exogenous variables are as follows: POP (2,1.2), F (3.5, 1) (the variance coming from e_{it}) and PROF (1.5, 0.5). The variable POP was created to mimic our actual POPulation variable, which exhibits a large variation between markets, but small variation within markets over time. The between markets variance was set to unity, while the within markets variance was set to 0.3. The average entry probability with this DGP was 0.0908, which is reasonably close to our data.

We estimated three different specifications: a standard probit, a simulated random effects probit with a linear random effect, and finally, a simulated random effects probit with the random effect placed into the S(.) function (corresponding to the true DGP):

(A.2) $y_{it} = 1((POP_{it})bPROF_{it} - F + e_{it} \geq 0)$

(A.3) $y_{it} = 1((POP_{it})bPROF_{it} - F + rh_i + e_{it} \geq 0)$

and (A.1). For identification, we normalized Var(e_{it}) ≡ 1 for all models. We use the same simulation estimator as with the actual data, i.e., a decomposition simulator, the number of simulation draws being

R = 40, and antithetics. The DGP was simulated 20 times. Estimation method is ML for standard probit, and MSL for the random effects specifications.

The results of the simulation exercise are given in Table A.I.

Table A.I.
Monte Carlo Results

Standard Probit	True Value	Mean	S.d.	Bias	MSE
F	3.5	2.79471	0.116768	-0.70529	1.133204
β	0.5	0.360472	0.046956	-0.13953	0.052384
ρ	0.5	0	0	-0.5	0.25
Linear decomp.	**True Value**	**Mean**	**S.d.**	**Bias**	**MSE**
F	3.5	3.37342	0.399397	-0.12658	0.668901
β	0.5	0.470211	0.061793	-0.02979	0.021187
ρ	0.5	0.283178	0.330067	-0.21682	0.210995
Nonlinear decomp.	**True Value**	**Mean**	**S.d.**	**Bias**	**MSE**
F	3.5	3.40612	0.246362	-0.09388	0.051897
β	0.5	0.483516	0.044727	-0.01648	0.001737
ρ	0.5	0.412275	0.206755	-0.08772	0.063712

As the results reveal, the standard probit and the linear decomposition (standard random effects) probit fare much worse than the nonlinear decomposition estimator. This was of course expected, as the latter corresponds to the true DGP in structure. What was not known was how biased the results would be if one made the wrong assumption concerning the error structure. The Monte Carlo experiment shows that although the linear decomposition estimator delivers reasonably accurate point estimates of the fixed entry costs, and the coefficient of market size times variable profits, the point estimate of the variance share of the random effect is severely downward biased, with a large standard error. In particular, we could not reject the Null that ρ is insignificantly different from zero. This is important as all the controls for unobserved heterogeneity that are used as standard, and are used in this paper with the reduced form models, are linearly separable.

Also, the performance of the nonlinear decomposition estimator was unknown. The Monte Carlo study indicates that the estimator performs well. The point estimates are closest to the true values; the standard errors are smaller than for the linear decomposition estimator as are the biases. The mean square errors are orders of magnitude smaller than those of the other estimators. As with the other two estimators, the biases are all downward; they are however much smaller. With this estimator, the point estimate of ρ is statistically significant.

This experiment naturally cannot indicate what the true interpretation of the (potential) unobserved heterogeneity in our data is. It underlines that whatever interpretation one chooses does matter for the results.