Incentives Work: Getting Teachers to Come to School

Esther Duflo, Rema Hanna, and Stephen Ryan

Web Appendix

Online Appendix: Estimation of model with AR(1) errors: Not for Publication

To estimate a model with this stochastic process on the outside option, we discretize the error term into 200 states and solve the dynamic programming problem defined by Equations 6 and 7 for an initial guess of our parameters.[40] We then simulate many work histories from this model, forming an unbiased estimate of the distribution of *sequences* of days worked at the beginning of each month. We simulate the model by drawing sequences $\epsilon = \{\epsilon_0, \ldots, \epsilon_t\}$ and following the optimal policy prescribed the dynamic program. For different sequences of ϵ's we obtain different work histories. Repeating this process many times results in unbiased estimates of the probabilities of all possible sequences. We then match the model's predicted set of probabilities over these sequences against their empirical counterparts.

To address the "incidental parameters" problem of not knowing the first ϵ, we treat the the draw on ϵ at $t = 1$ as coming from the limiting distribution of Equation 12, given by $\bar{F} = N(0, 1/(1-\rho^2))$. The key assumption here is that the dependence of the error term at the beginning of next month on the error term at the fifth day of the previous month is nearly zero, even for high values of ρ.[41] As a result, we can treat the error term next month as drawn from the unconditional distribution of ϵ.

Denoting a sequence of days worked as A, for each teacher-month sequence, we form a vector of moment conditions:

$$E[Pr(A_{im}; X_m) - \widehat{Pr}(A_{im}; X_m, \widehat{\theta})] = 0, \tag{15}$$

where $Pr(A_{im,}; X_m)$ is the empirical probability of observing a sequence of days worked conditioning on X_m, a vector containing the number of holidays and the maximum number of days in that month an agent could potentially work in that month.[42] The moments used in estimation sum across all months and all teachers to form the unconditional expectation of observing a sequence of days worked, $Pr(A)$. We form $\widehat{Pr}(A)$, the model's predicted

[40] The number of states for ϵ was determined by increasing the number of points in the discretization of the error term until there was no change in the expected distribution of outcomes. For alternative approaches to estimating dynamic discrete choice models with serially-correlated errors, see Keane and Wolpin (1994) and Stinebrickner (2000).

[41] For example, the dependence of an error term for the first day of next month, 22 periods in the future, on the error term from the fifth day of the current month is ρ^{22}, or 0.00039 for $\rho = 0.7$.

[42] This is necessary since the maximum payoff a teacher could obtain varies across months with the length of the month and the number of holidays in that month, which count as a day worked in the bonus payoff function if they fall on a workday.

counterpart, through Monte Carlo simulation:

$$\widehat{Pr}(A; X, \widehat{\theta}) = \frac{1}{N \cdot M \cdot NS} \sum_{i=1}^{N} \sum_{m=1}^{M} \sum_{j=1}^{NS} 1(A_{ijm} = A; X_m, \widehat{\theta}), \tag{16}$$

where A_{ijm} is the simulated work history associated with teacher i and simulation j, as derived from a dynamic program constructed in accordance with the parameters $\widehat{\theta}$ and the characteristics of the month m. The number of simulations used to form the expected probability of observing a sequence of days worked is denoted as NS. In all of our estimations we use $NS = 200,000$. Note that we are also drawing ϵ_1 anew from the distribution $\bar{F}$ for every simulated path, where we keep track of the seeding values in the random number generator, as to ensure that the function value is always the same for a given $\widehat{\theta}$. The unconditional moments are:

$$E[Pr(A; X_m) - \widehat{Pr}(A; X_m, \widehat{\theta})] = 0, \tag{17}$$

The objective function under the method of simulated moments is:

$$\min_{\theta} g(X_i, \theta)' \Omega^{-1} g(X_i, \theta), \tag{18}$$

where $g(X_i, \theta)$ is the vector of moments formed by stacking the $2^N - 1$ moments defined by Equation 17, and Ω^{-1} is the standard two-step optimal weighting matrix. For more details concerning the implementation and asymptotic theory of simulation estimators, see McFadden (1989) and Pakes and Pollard (1989).

Matching sequences of days worked from the first N days in each month produces $2^N - 1$ linearly independent moments, where we subtract one to correct for the fact that the probabilities have to sum to one. In our estimation, we match sequences of length $N = 5$, which generates 31 moments. Experimentation with shorter and longer sequences of days worked did not result in significant changes to the coefficients.[43]

We also relax the assumption that the outside option is equal across all agents by allowing

[43] There is also a related econometric problem: the more moments one has to match, the lower the number of observations corresponding to each sequence. As the number of moments gets large, the number of teachers who actually followed any specific sequence diminishes towards zero. The number of days we match reflects a tradeoff in the additional information embodied in a longer sequence of choice behavior against the empirical imprecision of measuring those moments. This is a conceptually separate problem from the computational burden of simulating the model probabilities precisely, which also contributes to noisy estimates. For example, using the first 16 days of the months, where we will start seing some teachers "in the money," generates 65,535 moments, which is more than double the number of observations in the data.

the outside option, μ_{im}, to be drawn anew for all teachers every month from a known parametric distribution $G(\mu)$. When forming moments in the MSM estimator in Equation 18 we need to integrate out this unobserved heterogeneity. The modification to the expected probability of observing a sequence of days worked for teacher i is then:

$$\widehat{Pr}(A_i; X, \widehat{\theta}) = \frac{1}{M \cdot NS \cdot U} \sum_{m=1}^{M} \sum_{i=1}^{NS} \sum_{u=1}^{U} 1(A_{im} = A; X_m, \widehat{\theta}_1, u_{im}), \tag{19}$$

where u_{im} is a draw of the mean level of the outside option from $G(\widehat{\theta}_2)$, the unknown distribution of heterogeneity in the population. In practice we set $U = 200$. For clarity, we have partitioned the set of unknown parameters into $\widehat{\theta}_1 = \{\beta, \rho\}$ and $\widehat{\theta}_2$, the set of parameters governing the distribution of unobserved heterogeneity. Note that this model is slightly different than the fixed-effects model considered in the i.i.d. case above, as it allows the draw of the outside option to vary across both months and agents.

We estimate models with two types of unobserved heterogeneity which differ through the specification of $G(\widehat{\theta}_2)$. In the first model $G(\widehat{\theta}_2)$ is distributed normally with mean and variance $\widehat{\theta} = \{\mu_1, \sigma_1^2\}$. In the second model, our preferred specification, we allow for a mixture of two types, where each type is distributed normally with proportion p and $(1-p)$ in the population.

Appendix Table 1: Empirical Sequences of Days Worked in the Last Five Days of a Month

In the Money (1)	Sequence of Days (2)	Frequency (3)	Percent (4)	Number of Days Worked (5)
No	00000	38	.4367816	0
No	01000	1	.0114943	1
No	10000	1	.0114943	1
No	00010	2	.0229885	1
No	00101	1	.0114943	2
No	00110	1	.0114943	2
No	01001	1	.0114943	2
No	00011	3	.0344828	2
No	11000	4	.045977	2
No	10011	1	.0114943	3
No	11010	1	.0114943	3
No	01101	1	.0114943	3
No	01011	1	.0114943	3
No	10101	1	.0114943	3
No	11100	2	.0229885	3
No	10110	2	.0229885	3
No	00111	3	.0344828	3
No	11001	3	.0344828	3
No	10111	2	.0229885	4
No	01111	3	.0344828	4
No	11110	4	.045977	4
No	11111	9	.1034483	5
Yes	00000	20	.019305	0
Yes	10000	2	.0019305	1
Yes	01000	2	.0019305	1
Yes	00010	3	.0028958	1
Yes	00100	7	.0067568	1
Yes	00001	9	.0086873	1
Yes	10010	2	.0019305	2
Yes	10001	2	.0019305	2
Yes	10100	2	.0019305	2
Yes	01010	3	.0028958	2
Yes	01001	4	.003861	2
Yes	00101	5	.0048263	2
Yes	00110	5	.0048263	2
Yes	01100	6	.0057915	2
Yes	11000	9	.0086873	2
Yes	00011	26	.0250965	2
Yes	10110	3	.0028958	3
Yes	10101	4	.003861	3
Yes	01101	5	.0048263	3
Yes	11010	5	.0048263	3
Yes	01110	6	.0057915	3
Yes	10011	8	.007722	3
Yes	01011	14	.0135135	3
Yes	11001	19	.0183398	3
Yes	11100	31	.0299228	3
Yes	00111	48	.046332	3
Yes	10111	23	.0222008	4
Yes	11110	48	.046332	4
Yes	11011	49	.0472973	4
Yes	11101	55	.0530888	4
Yes	01111	74	.0714286	4
Yes	11111	537	.5183398	5

Appendix Table 2: Fitted Moments

Sequence	Empirical	Fitted
(1)	(2)	(3)
0 0 0 0 0	0.062	0.066
0 0 0 0 1	0.009	0.01
0 0 0 1 0	0.004	0.006
0 0 0 1 1	0.021	0.016
0 0 1 0 0	0.004	0.004
0 0 1 0 1	0.001	0.003
0 0 1 1 0	0.006	0.003
0 0 1 1 1	0.045	0.024
0 1 0 0 0	0.005	0.006
0 1 0 0 1	0.005	0.004
0 1 0 1 0	0.001	0.002
0 1 0 1 1	0.012	0.008
0 1 1 0 0	0.007	0.005
0 1 1 0 1	0.012	0.008
0 1 1 1 0	0.008	0.008
0 1 1 1 1	0.105	0.058
1 0 0 0 0	0.015	0.008
1 0 0 0 1	0.005	0.01
1 0 0 1 0	0.002	0.004
1 0 0 1 1	0.018	0.02
1 0 1 0 0	0.002	0.004
1 0 1 0 1	0.002	0.006
1 0 1 1 0	0.006	0.006
1 0 1 1 1	0.042	0.041
1 1 0 0 0	0.015	0.013
1 1 0 0 1	0.016	0.016
1 1 0 1 0	0.009	0.008
1 1 0 1 1	0.048	0.038
1 1 1 0 0	0.011	0.023
1 1 1 0 1	0.036	0.039
1 1 1 1 0	0.055	0.064

Appendix Table 3: Alternative Specifications

Parameter	Appendix Model I	Appendix Model II
	(1)	(2)
β	0.043	0.016
	(0.002)	(0.000)
μ_1	1.236	-1.166
	(0.094)	(0.011)
σ_1^2	0.143	
	(0.039)	
μ_2	1.870	
	(0.672)	
σ_2^2	2.051	
	(0.305)	
p	0.003	
	(0.002)	
Yesterday Shifter		1.145
		(0.017)
Heterogeneity	RC	None
LLH		9654.887
ϵ_{Bonus}	1.371	0.722
	(0.436)	(0.075)
ϵ_{bonus_cutoff}	-6.06	-0.058
	(3.635)	(0.028)
Predicted Days Worked	14.98	17.264
	(1.541)	(0.360)
Days Worked BONUS=0	2.63	1.992
	(0.362)	(0.086)
Out of Sample Prediction	19.83	20.566
	(1.725)	(0.117)

Note: Both models are estimated using maximum likelihood.

Appendix Table 4: Model VI Cost-Minimization Policies

Days Worked	Bonus Cutoff	Bonus	Cost
(1)	(2)	(3)	(4)
14	13	25	642
15	21	75	686
16	18	50	698
17	20	75	782
18	19	75	877
19	23	175	967
20	17	75	1051
21	19	100	1134
22	21	200	1515