ONLINE APPENDIX – NOT FOR PUBLICATION

Contents

I Specification, Solution and Simulation of Life-Cycle Models

We assume that the utility function is intertemporally additive and the sub-utilities are iso-elastic. The problem of the generic household is:

$$\underset{C_t}{max}\ E_t \sum_{t=1}^{T} \beta^t u(C_t)$$

$$s.t.\ X_{t+1} = (1 + R_{t+1})(X_t - C_t) + Y_{t+1}$$

where C_t is nondurable consumption in period t, X_t is cash-on-hand (total financial and nonfinancial wealth) and Y_t is current labor income. We assume that durable consumption and leisure are separable from the nondurable consumption. The income process is assumed as follows:

$$Y_{t+1} = P_{t+1} \varpi_{t+1}$$

$$P_{t+1} = G P_t N_{t+1}$$

where G is predictable permanent income growth, P_t is permanent income which is subject to lognormally distributed shocks N_t with mean unity and variance $(e^{\sigma_n^2} - 1)$, current income Y_t equals permanent income multiplied by a transitory shock, ϖ_t , which is distributed lognormally with mean one and variance $(e^{\sigma_\varpi^2} - 1)$. The interest rate series is assumed to be generated by a stationary first order autoregressive process with long-run mean μ and autoregressive coefficient ρ. Interest rates shocks ε_{t+1} are assumed to be white noise with variance, σ_ε^2. The process is:

$$R_{t+1} = (1 - \rho)\mu + \rho R_t + \varepsilon_{t+1}.$$

The inter-temporal model described above does not have an analytical solution due to the assumed income uncertainty. Therefore we utilized the standard numerical dynamic programming methods to obtain a solution. Since the utility function is additive over the life cycle we solved the model recursively starting from the last period of life. We assume away any bequest motive so that consumption in period T is:

$$C_t(x_T) = x_T.$$

The problem is solved via policy function iteration using the terminal value condition. Having a nonstationary income process makes the problem harder to solve since the range of possible income values is too large. Instead, we redefine all relevant variables in terms of their ratios to permanent income and solve for the consumption to income ratio. By doing this we reduced the number of state variables to two, namely the cash on hand to income ratio and the interest rate. Moreover, we obtain an iid income process which can be approximated by standard Quadrature methods. Given the redefinition of the variables, the Euler equation can be written as:

$$\theta_t(\omega_t, R_t)^{-\gamma} - \beta E_t[(1 + R_{t+1})\theta_{t+1}(\omega_{t+1}, R_{t+1})^{-\gamma} n_{t+1}^{-\gamma} = 0$$

where $\theta_t = \frac{C_t}{P_t}$, $\omega_t = \frac{x_t}{P_t}$. At the terminal date T, consumption to income ratio is a function of only the cash on hand to income ratio and since the bequest motive is assumed away it follows that $\theta_T = \omega_T$.

For the income process, we use 10 point Gaussian Quadrature and we approximate the interest rate process by forming a 10 point first order discrete Markov process. We use a cubic spline to approximate the consumption function at each iteration. The agent is allowed to borrow the amount he can pay back with certainty. In practice this constraint will never bind because the functional form of the utility function implies that zero consumption results in infinite marginal utility. In models where we do not assume an explicit borrowing limit (model 1, 2, 3 and 4), the consumption functions are continuously differentiable. In fact, in our case where agents have iso-elastic preferences and income uncertainty, consumption functions are strictly concave.

In order to solve the problem, we define an exogenous grid for the cash on hand to income ratio: $\{x_j\}_{j=1}^{J}$. It is important to adjust the grid as the solution goes back in time. The algorithm finds the consumption level that makes the standard Euler equation hold for each value of x and r. We made the grid for x finer at lower levels in order to capture the curvature of the consumption function. After solving for the consumption function of a generic household for 80 periods, we simulate consumption paths for 10,000 ex-ante identical households facing the same interest rate realizations. We use only the middle 40 periods for the estimations. Table 1 presents the assumed parameter values for our experiments.

We also show that the second order approximation to the Euler equation does not provide a superior basis for estimation, and in fact suffers from a weak instrument problem: the standard instruments do not have useful predictive power for the second order term.

II Second Order Approximation

Following the steps in Carroll (2001) we get:

$$\Delta logC_{h,t+1} = \alpha + \frac{1}{\gamma} log(1 + R_{t+1}) + \frac{\gamma+1}{2} [\Delta logC_{h,t+1}]^2 + v_{h,t+1} \tag{1}$$

for the second order approximation.

The constant term α contains the discount rate (β) and the means of the higher order moments consumption growth and interest rates (for the second order approximation, this is the third and higher order moments.) In either case the residual term contains (i) the true innovation in marginal utility (or "expectation error") between t and $t+1$, ε_{t+1}, (ii) the measurement errors at t and $t+1$, and (iii) an approximation error composed of variation in the higher moments of consumption growth and interest rates (conditional on past information).

For the second order approximation , the instruments we consider are the ones used extensively in the literature: lagged interest rates, twice lagged consumption growth and lagged income and twice lagged consumption growth squared to the instrument set for the second order approximation

Note that because we know the true value of the preference parameter (γ) in our simulated data, we can use the true parameter value to construct the true residuals in the linearized Euler equation (plus a constant). That is, inverting equations (1) and evaluating at $\gamma = 4$ gives:

for the first order approximation and

$$\Delta logC_{h,t+1} - \frac{1}{4} log(1 + R_{t+1}) - \frac{4+1}{2} (\Delta logC_{h,t+1})^2 = \alpha + v_{h,t+1} \tag{2}$$

When there are two endogenous variables, a maximum relative bias of 10% compared to OLS, a 5% significance level and four instruments (as in our experiments with the second-order linearized Euler equation), the critical value is 7.56 for Linear GMM , and 4.72 for LIML.

Appendix Table A1 and A2 are parallel Tables 4 and 5 in the text, but for the second-order approximation of the Euler Equation. The first thing to note in Table A1 is that there are now strong signs of weak instrument problems (recall that critical value for two endogenous and four instruments at the 5% significance level is 7.56 for Linear GMM and 4.72 for LIML). The extra endogenous variable here is the square of consumption growth, and the extra instrument the square of lagged consumption growth, as is standard in the literature, but the instrument set has insufficient predictive power. In commenting on the poor performance of the second-order linearization to of the Euler Equation, Attanasio and Low (2004) postulated that the problem was the lack of a good instrument for the second order term. The statistical tests reported in the final column of Table A1 confirm that this is the case.

Turning to instrument validity, we see similar patterns as in Table 4. The lags of consumption growth (and its square), and lagged income, often are not. Problems of invalid instruments are greater in models with impatient agents but are somewhat ameliorated by measurement error in consumption (and the consequent use of second lags as instruments).

Table A2 reports the estimates. There are two coefficients to estimate in the second-order approximate Euler equation. The coefficient on the interest rate is the EIS, while the coefficient on the square of consumption growth is the coefficient of relative prudence.

There are two important questions here. First, do the estimates of the coefficient of relative prudence contain any useful information. The answer is plainly no. In almost all models the estimates are very imprecise and suffer from substantial mean bias. The second question is whether the inclusion of the second order term in consumption growth improves the estimate of the EIS, by reducing the scope for approximation bias. Comparing Table A2 with Table 5 suggest no strong evidence of improvement.

III Non-linear GMM Estimation

A summary of section 5 is that with realistic data structures, linearized Euler Equation estimation seems much less promising. An alternative strategy is estimating the full nonlinear Euler equation (Equation (3)) using the Generalized Method of Moments methodology introduced by Hansen (1982) , and hence eliminating problems associated with approximation. As discussed in section 2, the main problem of this strategy is that consistent estimation requires perfect data on consumption, whereas the available consumption data for households are almost certainly very noisy[1].

Appendix Table A3 presents the results of two alternative nonlinear GMM estimation on the consumption data with measurement error. The first non-linear GMM is Exact GMM (EGMM) introduced by Hansen (1982). The second estimator (GMM-D), is proposed by Alan, Attanasio and Browning (2009) takes account of the measurement error in consumption[2], and uses the following moment condition rather than Equation (3):

$$\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}(1+r_{t+1})\beta - \left(\frac{c_{t+2}}{c_t}\right)^{-\gamma}(1+r_{t+1})(1+r_{t+2})\beta^2 = 0 \tag{3}$$

For both of the estimators, we have 2 unknown parameters, CRRA (γ) and Discount Factor (β). We use two sets of instruments, first one contains constant and lagged interest rate and second set additionally includes twice lagged consumption growth [3].

For each of our estimations, we set N to 100 and T to 40. The results in Table 10 presents 999 Monte Carlo replications for four models . The results are indicating that as expected Exact Euler Equation perform poorly in identifying CRRA, mean biases are big and estimates are very imprecise. This is especially the case in just identified models (first row), but adding the twice lagged consumption growth in the instrument set makes a little improvement (second row of the Table A3). Results are also confirming that similar to linearized Euler Equations, non-linear GMM is performing better in the patient environments.The estimations of discount factor in all models show a downward bias, mean standard errors are relatively big, indicating also the imprecision of the estimates.

The last two rows of the Table A3 presents the results from GMM-D. It is clear from third row, in the just identified case we don't see a real improvement. Adding twice lagged consumption to instrument set is improving the performance of GMM-D esspecially in the patient models, but still it is not clearly superior to the linearized Euler Equation. This is clear in comparision of mean biases in Table A2 and A3. Discount factor in all models still show a significant downward bias.

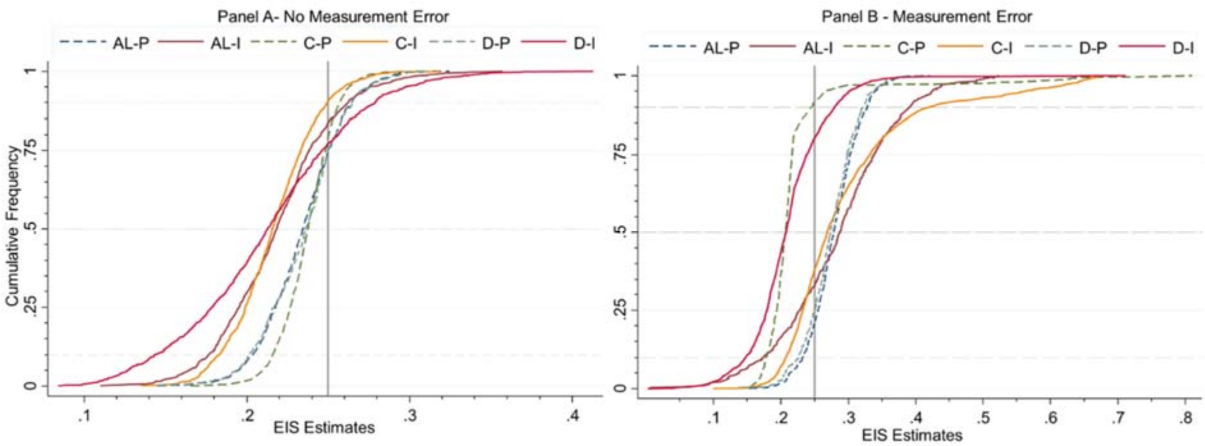
[1]It is also important to note that although large T is still required in non-linear GMM since estimation on synthetic cohorts requires an equation that is linear in parameters, Non-Linear GMMs cannot be used in synthetic panels

[2]Alan et al. (2009) has exploited the iso-elastic utility assumption in agents problem that gives a ratio of consumptions in the Euler equations. Authors showed that their proposed GMM-D estimator is performing better than EGMM in the environments where measurement error is stationary and multiplicative.

[3]We also used once lagged consumption growth instead of twice lagged, our results show very large mean biases.

Overall, the results of this subsection do not significantly alter the findings that emerged in the previous subsections. Specifically we find that in our environments non-linear GMM estimators do not out perform the linearized Euler Equation.

Figure A1: Cumulative Density Functions of EIS estimates



Notes: The vertical reference line is at the true value of EIS (0.25.) Detailed statistics of Monte Carlo results are presented in the column 1 of Table 5.

Table A1: Second Order Approximation

Instrument Validity and Relevance Results, 999 replications with Measurement Error

	Instrument Validity				Instrument Relevance (First Stage)
	<u>Second Order</u>: $\Delta logC_{t+1}^{h} - \frac{1}{4} log(1+R_{t+1}) - \frac{4+1}{2} \left(\Delta logC_{t+1}^{h}\right)^2 = \alpha + v_{t+1}^{h}$				
	Fraction of t stats with absolute value greater than 1.96				CDF stat
Model	***Lag Interest Rate***	***Twice Lag Consumption Growth***	***Lag Income***	***Twice Lag Consumption Growth Square***	[10% , 90%] Mean
AL-P	[-0.23, 2.04]	[-3.10, 9.26]	[-9.31, 1.14]	[-7.97, 2.36]	[0.24, 2.47]
	0.039	0.093	0.057	0.114	1.19
AL-I	[-0.10, 36.8]	[-9.33, -0.32]	[-55.5, 0.15]	[1.10, 4.63]	[3.63, 6.22]
	0.160	0.340	0.486	0.779	4.90
C-P	[-5.66, 7.30]	[-7.95, 2.87]	[-8.93, 5.8]	[-12.49, 3.56]	[0.13, 6.24]
	0.145	0.298	0.144	0.305	2.60
C-I	[-8.10, 6.93]	[-8.97, 2.74]	[-3.18,9.67]	[-6.68, 6.52]	[0.09, 25.32]
	0.176	0.491	0.268	0.368	5.81
D-P	[-1.23, 1.84]	[-2.49, 10.8]	[-6.45, 1.14]	[-5.53, 2.37]	[0.24, 2.41]
	0.029	0.083	0.046	0.107	1.18
D-I	[-0.45, 4.74]	[-4.82, 5.51]	[-5.21, 4.69]	[-3.98, 3.35]	[0.31, 4.09]
	0.156	0.309	0.103	0.197	1.92

Notes to Table A1:

1. All other numbers are result of 999 Monte Carlo replications. Panel length is 40 periods.
2. For the second order approximation we use the twice lagged consumption instruments with the lag interest rate and lag income.
3. Instrument validity test is a t-test obtained from the regression of constructed residuals on instruments. Instrument validity columns report for each instrument the fraction of t stats with absolute value greater than 1.96 (critical value at 5% significance level)
4. Instrument Relevance column reports the Cragg-Donald F (CDF) statistic from the first stage of IV. For the second order, endogenous variables are interest rate and lagged consumption growth square and instruments are lagged interest rate, twice lagged consumption growth, twice lagged consumption growth square and lagged income. Mean values of CDF are reported. CDF values at 10 and 90 percent are in parentheses. Stock and Yogo (2002) - test for weak instrument H_0 =bias of two stage estimation relative to OLS is greater than 10% . Critical Value at 5% significance level when the number of instruments is 4 and endogenous variable is 2 for Linear GMM = 7.56, for LIML =4.72

Appendix Table A2: Second Order Approximation, Estimates of EIS and Prudence, 999 replications, With Measurement Error

	True Value	EIS: $\left(\frac{1}{\gamma}\right) = 0.25$	PRUDENCE: $\left(\frac{\gamma+1}{2}\right) = 2.5$
Model			
AL-P	*mean coefficient*	0.293	0.38
	[2.5%, 97.5%]	[0.205 ,0.396]	[-11.32, 10.71]
	(mean bias as % of true parameter value)	(17.2)	(-84.8)
	mean standard error	0.121	3.92
AL-I	*mean coefficient*	0.183	6.59
	[2.5%, 97.5%]	[0.008, 0.311]	[4.27, 9.42]
	(mean bias as % of true parameter value)	(-26.8)	(163)
	mean standard error	0.189	3.02
C-P	*mean coefficient*	0.234	1.38
	[2.5%, 97.5%]	[0.150, 0.380]	[-3.18, 7.09]
	(mean bias as % of true parameter value)	(-6.4)	(-44.8)
	mean standard error	0.102	5.06
C-I	*mean coefficient*	0.321	1.09
	[2.5%, 97.5%]	[0.175, 0.686]	[-7.76, 9.35]
	(mean bias as % of true parameter value)	(28.4)	(-56.4)
	mean standard error	0.095	1.98
D-P	*mean coefficient*	0.292	0.832
	[2.5%, 97.5%]	[0.203, 0.387]	[-9.80, 10.71]
	(mean bias as % of true parameter value)	(16.8)	(-66.7)
	mean standard error	0.132	3.21
D-I			
	mean coefficient	0.156	5.85
	[2.5%, 97.5%]	[-0.153, 0.402]	[-0.80, 11.81]
	(mean bias as % of true parameter value)	(-37.6)	(134)
	mean standard error	0.134	5.42

Notes to Table A2:

1. Panel length is 40 periods.
2. For each model and estimation strategy, the first number is the mean value of the estimate of the EIS and Prudence in 999 Monte Carlo replications. The second number, in brackets, is the semi parametric confidence interval, third number in parenthesis is the mean bias as the percentage of true parameter value. The last number is mean robust standard errors clustered on period.

Appendix Table A3: Non-linear GMM
Exact Euler Equation and GMM-D, 999 replications, Long Panel With Measurement Error

		Model	AL-P		AL-I		C-P		C-I	
		Parameters	$1/\gamma$	β	$1/\gamma$	β	$1/\gamma$	β	$1/\gamma$	β
Estimator	Instruments	True Value	0.25	0.971	0.25	0.935	0.25	0.971	0.25	0.935
Exact Euler Equation	$(1+r_t)$									
		mean coefficient	0.203	0.951	0.211	0.789	0.214	0.947	0.213	0.924
		(mean standard error)	(0.454)	(0.064)	(0.620)	(0.708)	(0.432)	(0.998)	(0.173)	(0.029)
		(mean bias as % of true parameter value)	[-18.7]	[-5.3]	[-15.6]	[-15.6]	[-14]	[-2.5]	[-14.6]	[0.1]
		RMSE	0.46	0.067	0.621	0.72	0.433	1.00	0.177	0.031
Exact Euler Equation	$(1+r_t)$, (C_{t-1}/C_{t-2})									
		mean coefficient	0.202	0.919	0.372	0.861	0.211	0.908	0.419	0.924
		(mean standard error)	(0.388)	(0.003)	(0.134)	(0.131)	(0.373)	(0.002)	(0.373)	(0.029)
		(mean bias as % of true parameter value)	[-18.9]	[-2]	[48.5]	[-7.9]	[-15.6]	[-6.4]	[149]	[-1.2]
		RMSE	0.391	0.052	0.181	0.150	0.375	0.063	0.409	0.031
GMM-D	$(1+r_t)$									
		mean coefficient	0.223	0.895	0.236	0.926	0.245	0.896	0.247	0.870
		(mean standard error)	(0.13)	(0.021)	(0.172)	(0.005)	(0.135)	(0.361)	(0.271)	(0.922)
		(mean bias as % of true parameter value)	[-8.13]	[-7.8]	[-5.33]	[-0.9]	[-1.78]	[-7.7]	[-0.87]	[-6.9]
		RMSE	0.133	0.078	0.173	0.010	0.135	0.369	0.271	0.924
GMM-D	$(1+r_t)$, (C_{t-1}/C_{t-2})									
		mean coefficient	0.235	0.902	0.237	0.930	0.252	0.905	0.303	0.890
		(mean standard error)	(0.127)	(0.018)	(0.186)	(0.005)	(0.127)	(0.016)	(0.565)	(0.023)
		(mean bias as % of true parameter value)	[-5.95]	[-7.1]	[-5.32]	[-0.5]	[0.83]	[-6.8]	[21.2]	[-4.8]
		RMSE	0.128	0.071	0.186	0.007	0.127	0.068	0.567	0.050

Notes to Appendix Table A3

1. Panel length is 40 periods. Table reports the nonlinear GMM results using the simulated data with measurement error. The measurement error that we add to the simulated data is i.i.d log normal with a unit mean and a variance of 0.004. The data structure is a long panel (T=40)
2. All other numbers are result of 999 Monte Carlo replications. For each model and estimation strategy, the first number is the mean value of the estimate of the EIS in 999 Monte Carlo replications. The second number, in parenthesis, is the mean robust standard errors clustered on period. The third number which is in the square brackets is the mean bias as the percentage of true parameter value. RMSE reports the root mean squared error.

Appendix Table A4: Deaton Impatient Model - Instrument Validity and Relevance Results, 999 replications

	Instrument Validity *First Order*: $\Delta log C^{h}_{t+1} - \frac{1}{4} log(1+R_{t+1}) = \alpha + e^{h}_{t+1}$			Instrument Relevance (First Stage)
Model	***Lag Interest Rate***	***Twice Lag Consumption Growth***	***Lag Income***	F-Stat Mean [10%, 90%]
LONG PANEL MEASUREMENT ERROR				
D-I *Cons. Excluded*				
[2.5%, 97.5%]	[-3.9, 2.34]	[-2.84, 0.89]	[-4.06, 0.75]	6681
"Fraction of |t|>1.96"	0.130	0.224	0.139	[6572 , 6776]
Cons. Included				
[2.5%, 97.5%]	[-7, 4.35]	[-1.99, 2.10]	[-3.06, 1.95]	6967
"Fraction of |t|>1.96"	0.322	0.343	0.282	[6431 , 7102]
SHORT PANEL MEASUREMENT ERROR				
D-I *Cons. Excluded*				
[2.5%, 97.5%]	[-13.7, 0.64]	[-9.06, 7.41]	[-1.41, 7.87]	504
"Fraction of |t|>1.96"	0.319	0.706	0.234	[498, 511]
Cons. Included				
[2.5%, 97.5%]	[-15, -1.82]	[0.99, 5.10]	[1.02, 3.95]	502
"Fraction of |t|>1.96"	0.540	0.782	0.578	[496, 508]
SYNTHETIC PANEL MEASUREMENT ERROR				
D-I *Cons. Excluded*				
[2.5%, 97.5%]	[-3.02, 2.10]	[-2.34, 0.98]	[-2.06, 1.75]	6.36
"Fraction of |t|>1.96"	0.103	0.203	0.089	[3.17, 11.67]
Cons. Included				
[2.5%, 97.5%]	[-3.72, 0.05]	[-3.9, 1.80]	[-5.04, 3.45]	7.05
"Fraction of |t|>1.96"	0.240	0.332	0.430	[2.16, 12.8]

Notes to Table A4:

1. All other numbers are result of 999 Monte Carlo replications. The measurement error that we add to the simulated data is i.i.d log normal with a unit mean and a variance of 0.004
2. For the first order approximation instruments, we use the twice lagged consumption instrument with the lag interest rate and lag income.

3. Instrument validity test is a t-test obtained from the regression of constructed residuals on instruments.
    a. Instrument validity columns report for each instrument (i) t-statistic values at 2.5 and 97.5 percent , (ii) the fraction of t stats with absolute value greater than 1.96 (critical value at 5% significance level)
    b. Instrument Relevance column reports the Cragg-Donald F (CDF) statistic from the first stage of IV. For the first order approximation interest rate is the only endogenous variable and lagged interest rate, (twice)lagged consumption growth, lagged income are instruments. Mean values of CDF are reported. CDF values at 10 and 90 percent are in parentheses.
    c. Stock and Yogo (2002) - test for weak instrument H_0 =bias of two stage estimation relative to OLS is greater than 10%
        i. Critical Value at 5% significance level when the number of instruments is 3 and endogenous variable is 1 for Linear GMM = 9.08, for LIML=6.46.

Appendix Table A5: Deaton Impatient Model -Estimates of the EIS, 999 replications
First Order Approximation, With Measurement Error, Long Panel

True Value EIS: $\left(\frac{1}{\gamma}\right) = 0.25$

D-I Model	***Cons. Excluded***	***Cons. Included***
	LONG PANEL	
mean coefficient	0.209	0.187
[2.5%, 97.5%]	[0.16, 0.28]	[0.09, 0.31]
(mean std. error)	(0.194)	(0.189)
(mean bias as % of true parameter value)	(-13.4)	(-25.2)
	SHORT PANEL	
mean coefficient	0.183	0.122
[2.5%, 97.5%]	[0.032, 0.438]	[0.011, 0.214]
(mean std. error)	(0.251)	(0.315)
(mean bias as % of true parameter value)	(-26.8)	(-51.2)
	SYNTHETIC PANEL	
mean coefficient	0.2	0.167
[2.5%, 97.5%]	[0.054, 0.378]	[0.052, 0.36]
(mean std. error)	(0.401)	(0.392)
(mean bias as % of true parameter value)	(-20)	(-33.2)

1. Panel length is 40 periods.
2. For each column, the first number is the mean value of the estimate of the EIS in 999 Monte Carlo replications. The second number, in brackets is 2.5 and 97.5 percent of EIS estimates. Third number is the mean robust standard errors clustered on period. Fourth number is the mean bias as the percentage of true parameter value.

Appendix Table A6: First Order Approximation, Measurement Error, Different Data Structures, Confidence Intervals

Model		*Mean Coefficient*	*Mean Sd. Error*	% of CI that includes either 0 or 1	both 0 or 1
AL-P	*Long Panel*	0.279	(0.106)	15	0
	Short Panel	0.211	(0.336)	100	22
	Synthetic Cohort	0.218	(0.224)	100	15
AL-I	*Long Panel*	0.284	(0.245)	97	18
	Short Panel	0.204	(0.356)	100	24
	Synthetic Cohort	0.222	(0.339)	100	25
C-P	*Long Panel*	0.212	(0.098)	30	2
	Short Panel	0.332	(0.366)	98	32
	Synthetic Cohort	0.203	(0.232)	82	11
C-I	*Long Panel*	0.298	(0.189)	45	12
	Short Panel	0.345	(0.285)	100	40
	Synthetic Cohort	0.297	(0.335)	99	35
D-P	*Long Panel*	0.274	(0.105)	23	3
	Short Panel	0.208	(0.335)	100	27
	Synthetic Cohort	0.214	(0.223)	100	13
D-I	*Long Panel*	0.209	(0.194)	96	1
	Short Panel	0.183	(0.251)	100	16
	Synthetic Cohort	0.200	(0.401)	100	43