Heterogeneity, Risk Sharing and the Welfare Costs of Idiosyncratic Risk

INTERNET APPENDIX

Sam Schulhofer-Wohl*

September 7, 2007

Abstract

This appendix contains four sections. Section A reports results from running regressions of earnings on GDP using data from the PSID, for comparison with the results using HRS data in the body of the paper. Section B gives technical details on the Markov Chain Monte Carlo estimation employed in table 2 of the paper and reports the complete parameter estimates for the regressions summarized in that table. Section C reports results comparable to tables 2 and 3 of the paper using only Social Security covered earnings instead of the combination of Social Security and W-2 earnings. Section D shows that the Kiefer (1980) estimator for a regression model with an interaction between individual and time effects extends easily to an unbalanced panel with individual-specific intercepts and trends.

*Department of Economics and Woodrow Wilson School of Public and International Affairs, Princeton University, 363 Wallace Hall, Princeton, NJ 08544. E-mail: sschulho@princeton.edu. Phone: (609) 258-7392.

A An experiment with earnings and aggregate shocks in the PSID

I investigate how top-coding in the HRS data affects my results on the relationship between income and aggregate shocks by estimating the Mincer equation with GDP on data from the Panel Study of Income Dynamics, where top-coding is minimal. I study the labor earnings of male heads of household in the 1968 to 1997 waves of the PSID. To make the analysis as comparable as possible to the HRS results, I restrict the sample to men born in 1924 through 1947, the birth cohorts that the HRS covers. I drop observations where the worker's reported age on the annual PSID file differs by more than one year from the worker's age calculated from the reported birth year in the last file where the worker appears. As with the HRS, I drop workers who have fewer than five nonzero earnings observations. I deflate earnings by the Consumer Price Index. Table A.1 gives summary statistics.

I estimate the model by ordinary least squares using data in levels and in differences. If cointegration of GDP and earnings generates spurious results, the estimates in differences – which cannot be calculated for the HRS due to top-coding – should differ from the estimates in levels. I also create a dataset where I top-code all earnings observations above the 70th percentile of earnings in a given year and estimate a tobit model on this dataset. These results should be similar to the OLS estimates if the tobit model adequately accounts for top-coding.

I weight the workers in two ways. One, I weight each worker by the degrees of freedom in his wage regression, since these degrees of freedom are inversely proportional to the variance of the coefficient estimates for the worker. Two, I multiply the degrees of freedom by the worker's permanent income, calculated as the exponential of the worker's mean log income. The second set of weights cannot be constructed in the HRS due to top-coding.

Table A.2 shows the results for the PSID data. OLS coefficients in levels, OLS coefficients in differences and tobit coefficients are similar, and all of the unweighted coefficient estimates are similar to the estimates obtained for the HRS. These findings suggest that the HRS results would not have changed much if it had been possible to estimate the model in differences and if there had been no top-coding. Also, the GDP coefficient in the PSID is much larger than one when workers are weighted by regression degrees of freedom, but statistically indistinguishable from one when workers are reweighted by permanent income. These findings support the hypothesis that the HRS coefficients exceed one because the HRS analysis weights workers by regression degrees of freedom rather than by permanent income.

Table A.1: Summary statistics for PSID labor earnings sample.

Variable	mean	s.d.
Variables that are constant for each worker		
Age in 1992	55.0	7.21
Observations on worker	17.8	7.51
Permanent income[a]	23,421	15,985
Variables that change over time		
Age	42.8	8.97
Annual labor income[b]	27,270	24,368
Log(annual income)[b]	9.97	0.76
Number of men	2,002	
Number of observations	35,687	

Male heads of household in PSID born in 1924 through 1947 with at least five nonzero annual labor earnings observations in 1968 through 1997 survey waves. See appendix A for additional sample restrictions. [a]Calculated as exponential of time series mean of worker's log income. [b]Deflated by CPI; 1982-1984 dollars.

Table A.2: Regressions of log(real annual labor income) on GDP, male household heads ages 23-61 in PSID.

	(1)	(2)	(3)	(4)
	A. Individual OLS regressions in levels			
coefficient on	2.062	1.374	1.342	1.160
log(GDP)	(0.383)	(0.123)	(0.184)	(0.126)
workers	2,002	2,002	2,002	2,002
	B. Individual OLS regressions in differences			
coefficient on	1.793	1.343	1.388	1.099
log(GDP)	(0.280)	(0.122)	(0.153)	(0.112)
workers	2,002	2,002	2,002	2,002
	C. Individual tobit regressions in levels			
coefficient on	2.210	1.473	1.430	1.121
log(GDP)	(0.501)	(0.264)	(0.345)	(0.322)
workers	1,588	1,588	1,588	1,588
	D. Pooled regressions in differences			
coefficient on	1.314	-	1.090	
log(GDP)	(0.168)	-	(0.197)	
observations	35,687	-	35,687	
workers	2,002	-	2,002	
R-squared	0.005	-	0.006	
weights	None	degrees of freedom	permanent income	perm. inc. * d.f.

Panels A, B and C show means of coefficients from worker-by-worker regressions of log income (deflated by the CPI) on log GDP (per capita, chained 2000 dollars), with the standard error of the mean in parentheses. Estimates in levels include age and age squared as controls; estimates in differences include age as a control. Panel D shows coefficients from regressions pooling the data on all workers, with heteroscedasticity-robust standard errors adjusted for clustering by year in parentheses; a time trend is included as a control. In panel C, the highest 30 percent of earnings observations in each year were top-coded at the 70th percentile of observed earnings before running the regressions, and workers were dropped if the tobit estimation did not converge.

B Technical details and complete parameter estimates for table 2

The models estimated in Table 2 are of the form

$$\log\left(earnings_{it}\right) = a_i + \pi_{1j}\log\left(aggregate_t\right) + \mathbf{x}_{it}'\mathbf{\Pi}_j + u_{jt} + e_{it},$$

where i indexes workers, $j \in \{low, high\}$ indexes risk-tolerance groups and t indexes years.

I assume:

- The idiosyncratic errors e_{it} are i.i.d. normally distributed across individuals and dates with mean zero and variance $\sigma^2_{e,j}$. Define the precision of e_{it} to be $\tau_{e,j} = 1/\sigma^2_{e,j}$.
- The individual random effects a_i are i.i.d. normally distributed across individuals with mean zero and variance $\sigma^2_{a,j}$. Define $\tau_{a,j} = 1/\sigma^2_{a,j}$.
- The year random effects u_{jt} are i.i.d. normally distributed over years and across the two risk-tolerance groups with mean zero and variance $\sigma^2_{u,j}$. Define $\tau_{u,j} = 1/\sigma^2_{u,j}$. I include the year random effects in case the data are not independent across workers, in the spirit of reporting clustered standard errors in a frequentist framework. In practice, it turns out that the estimated variance of the year random effects is much smaller than that of the individual and idiosyncratic effects and that including year random effects has little influence on the results.

I adopt the following priors:

- Uniform (uninformative) prior for $\mathbf{\Pi}_j$ and π_{1j}.
- $\Gamma(1,1)$ prior for each of $\tau_{a,j}$, $\tau_{e,j}$ and $\tau_{u,j}$. Given the large number of observations, these priors have little influence on the results.

I follow algorithms 10 and 16 in Chib (2001) to obtain draws from the posterior distribution of $\{(\pi_{1j}, \mathbf{\Pi}_j, \tau_{a,j}, \tau_{e,j}, \tau_{u,j})\}_{j=low,high}$, accounting for the fact that earnings are censored at the Social Security taxable maximum. The algorithms employ Gibbs sampling. I run the Gibbs sampler for 2,000 iterations and discard results from the first 1,000 iterations, then report statistics for the remaining 1,000 iterations.

Table 2 contains six blocks of estimates, corresponding to two samples (all men, and white native-born men with a high school education) and three aggregate variables (GDP, personal consumption expenditures and aggregate wages and salaries). The following six tables report the complete set of estimated parameters for each block. The time variable t is normalized as $year - 1951$.

Table B.1: Pooled random-intercept tobit regression of log(real annual earnings) on log(real GDP per capita), all men ages 23-61 in Health and Retirement Study (combined Social Security and W-2 earnings).

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
log(aggregate)	mean	2.624	2.863	0.239
	s.d.	0.054	0.077	0.097
	2.5 percentile	2.521	2.718	0.047
	97.5 percentile	2.728	3.020	0.425
education	mean	0.076	0.082	0.007
	s.d.	0.004	0.005	0.007
	2.5 percentile	0.067	0.072	-0.007
	97.5 percentile	0.084	0.093	0.020
experience	mean	0.084	0.082	-0.002
	s.d.	0.003	0.004	0.004
	2.5 percentile	0.079	0.075	-0.010
	97.5 percentile	0.089	0.089	0.007
experience^2	mean	-0.001	-0.002	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.002	-0.002	0.000
	97.5 percentile	-0.001	-0.001	0.000
white	mean	0.451	0.370	-0.081
	s.d.	0.038	0.053	0.065
	2.5 percentile	0.379	0.271	-0.202
	97.5 percentile	0.527	0.479	0.054
immigrant	mean	0.009	0.145	0.136
	s.d.	0.058	0.075	0.095
	2.5 percentile	-0.104	0.000	-0.058
	97.5 percentile	0.129	0.297	0.316
veteran	mean	-0.027	0.028	0.055
	s.d.	0.031	0.042	0.052
	2.5 percentile	-0.087	-0.054	-0.044
	97.5 percentile	0.033	0.110	0.153
t	mean	-0.029	-0.032	-0.003
	s.d.	0.005	0.007	0.008
	2.5 percentile	-0.038	-0.044	-0.020
	97.5 percentile	-0.019	-0.019	0.013

continued on next page

Table B.1: Pooled random-intercept tobit regression of log(real annual earnings) on log(real GDP per capita), all men ages 23-61 in Health and Retirement Study (combined Social Security and W-2 earnings).

continued from previous page

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
t^2	mean	-0.001	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.001	0.000
	97.5 percentile	-0.001	-0.001	0.000
constant	mean	-17.204	-19.461	-2.256
	s.d.	0.513	0.731	0.910
	2.5 percentile	-18.220	-20.909	-3.960
	97.5 percentile	-16.232	-18.068	-0.476
s.d.(e_{it})	mean	0.838	0.897	0.060
	s.d.	0.003	0.004	0.005
	2.5 percentile	0.832	0.890	0.051
	97.5 percentile	0.843	0.905	0.069
s.d.(u_{jt})	mean	0.201	0.246	0.045
	s.d.	0.024	0.032	0.041
	2.5 percentile	0.160	0.192	-0.035
	97.5 percentile	0.254	0.317	0.133
s.d.(a_i)	mean	0.665	0.720	0.055
	s.d.	0.011	0.014	0.018
	2.5 percentile	0.644	0.692	0.020
	97.5 percentile	0.688	0.749	0.089

Table B.2: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate personal consumption expenditure per capita), all men ages 23-61 in Health and Retirement Study (combined Social Security and W-2 earnings).

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
log(aggregate)	mean	2.485	2.592	0.107
	s.d.	0.066	0.087	0.110
	2.5 percentile	2.353	2.418	-0.110
	97.5 percentile	2.615	2.748	0.328
education	mean	0.075	0.082	0.006
	s.d.	0.004	0.006	0.007
	2.5 percentile	0.067	0.071	-0.007
	97.5 percentile	0.084	0.094	0.021
experience	mean	0.083	0.081	-0.002
	s.d.	0.003	0.004	0.005
	2.5 percentile	0.077	0.074	-0.011
	97.5 percentile	0.088	0.088	0.007
experience^2	mean	-0.001	-0.002	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.002	-0.002	0.000
	97.5 percentile	-0.001	-0.001	0.000
white	mean	0.450	0.371	-0.079
	s.d.	0.038	0.053	0.066
	2.5 percentile	0.375	0.265	-0.209
	97.5 percentile	0.523	0.472	0.049
immigrant	mean	0.005	0.145	0.139
	s.d.	0.060	0.073	0.095
	2.5 percentile	-0.115	-0.004	-0.039
	97.5 percentile	0.121	0.281	0.328
veteran	mean	-0.025	0.030	0.055
	s.d.	0.033	0.043	0.054
	2.5 percentile	-0.086	-0.053	-0.053
	97.5 percentile	0.040	0.116	0.162
t	mean	-0.031	-0.032	-0.001
	s.d.	0.005	0.007	0.008
	2.5 percentile	-0.041	-0.045	-0.017
	97.5 percentile	-0.021	-0.019	0.016

continued on next page

Table B.2: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate personal consumption expenditure per capita), all men ages 23-61 in Health and Retirement Study (combined Social Security and W-2 earnings).

continued from previous page

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
t^2	mean	-0.001	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.001	0.000
	97.5 percentile	-0.001	-0.001	0.000
constant	mean	-14.696	-15.664	-0.968
	s.d.	0.596	0.791	1.000
	2.5 percentile	-15.858	-17.065	-2.934
	97.5 percentile	-13.541	-14.065	1.040
s.d.(e_{it})	mean	0.837	0.897	0.060
	s.d.	0.003	0.004	0.005
	2.5 percentile	0.832	0.889	0.050
	97.5 percentile	0.843	0.905	0.070
s.d.(u_{jt})	mean	0.205	0.249	0.043
	s.d.	0.024	0.030	0.039
	2.5 percentile	0.164	0.196	-0.033
	97.5 percentile	0.259	0.312	0.122
s.d.(a_i)	mean	0.664	0.719	0.054
	s.d.	0.011	0.015	0.018
	2.5 percentile	0.644	0.690	0.017
	97.5 percentile	0.685	0.749	0.092

Table B.3: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate wages and salaries per capita), all men ages 23-61 in Health and Retirement Study (combined Social Security and W-2 earnings).

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
log(aggregate)	mean	2.266	2.578	0.312
	s.d.	0.049	0.070	0.087
	2.5 percentile	2.170	2.434	0.143
	97.5 percentile	2.358	2.716	0.485
education	mean	0.076	0.083	0.007
	s.d.	0.004	0.006	0.007
	2.5 percentile	0.068	0.073	-0.007
	97.5 percentile	0.084	0.094	0.021
experience	mean	0.083	0.081	-0.002
	s.d.	0.002	0.002	0.003
	2.5 percentile	0.079	0.077	-0.007
	97.5 percentile	0.086	0.086	0.004
experience^2	mean	-0.001	-0.002	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.002	-0.002	0.000
	97.5 percentile	-0.001	-0.001	0.000
white	mean	0.452	0.371	-0.081
	s.d.	0.037	0.052	0.062
	2.5 percentile	0.378	0.264	-0.199
	97.5 percentile	0.526	0.471	0.037
immigrant	mean	0.007	0.141	0.134
	s.d.	0.059	0.076	0.096
	2.5 percentile	-0.115	-0.003	-0.061
	97.5 percentile	0.126	0.290	0.310
veteran	mean	-0.026	0.027	0.053
	s.d.	0.029	0.043	0.051
	2.5 percentile	-0.079	-0.055	-0.044
	97.5 percentile	0.033	0.115	0.149
t	mean	-0.033	-0.041	-0.008
	s.d.	0.004	0.006	0.008
	2.5 percentile	-0.042	-0.054	-0.023
	97.5 percentile	-0.025	-0.029	0.007

continued on next page

Table B.3: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate wages and salaries per capita), all men ages 23-61 in Health and Retirement Study (combined Social Security and W-2 earnings).

continued from previous page

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
t^2	mean	0.000	0.000	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.001	0.000
	97.5 percentile	0.000	0.000	0.000
constant	mean	-12.318	-15.027	-2.710
	s.d.	0.444	0.621	0.773
	2.5 percentile	-13.196	-16.235	-4.243
	97.5 percentile	-11.462	-13.737	-1.173
s.d.(e_{it})	mean	0.838	0.898	0.060
	s.d.	0.003	0.004	0.005
	2.5 percentile	0.832	0.891	0.050
	97.5 percentile	0.843	0.905	0.069
s.d.(u_{jt})	mean	0.197	0.242	0.045
	s.d.	0.023	0.033	0.040
	2.5 percentile	0.158	0.190	-0.028
	97.5 percentile	0.246	0.319	0.130
s.d.(a_i)	mean	0.666	0.723	0.057
	s.d.	0.011	0.015	0.019
	2.5 percentile	0.644	0.694	0.019
	97.5 percentile	0.688	0.753	0.092

Table B.4: Pooled random-intercept tobit regression of log(real annual earnings) on log(real GDP per capita), white, U.S.-born men ages 23-61 with exactly 12 years of education in Health and Retirement Study (combined Social Security and W-2 earnings).

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
log(aggregate)	mean	2.474	2.861	0.387
	s.d.	0.101	0.149	0.184
	2.5 percentile	2.275	2.581	0.039
	97.5 percentile	2.662	3.158	0.753
experience	mean	0.066	0.062	-0.005
	s.d.	0.005	0.007	0.009
	2.5 percentile	0.057	0.048	-0.023
	97.5 percentile	0.076	0.077	0.013
experience^2	mean	-0.002	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.002	-0.002	0.000
	97.5 percentile	-0.001	-0.001	0.000
veteran	mean	-0.087	0.070	0.157
	s.d.	0.051	0.078	0.090
	2.5 percentile	-0.188	-0.085	-0.022
	97.5 percentile	0.016	0.223	0.333
t	mean	-0.014	-0.030	-0.016
	s.d.	0.009	0.012	0.015
	2.5 percentile	-0.031	-0.056	-0.047
	97.5 percentile	0.002	-0.007	0.013
t^2	mean	-0.001	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.001	0.000
	97.5 percentile	-0.001	0.000	0.001
constant	mean	-14.234	-17.786	-3.553
	s.d.	0.956	1.398	1.735
	2.5 percentile	-16.067	-20.537	-7.045
	97.5 percentile	-12.376	-15.173	-0.248
s.d.(e_{it})	mean	0.780	0.801	0.021
	s.d.	0.005	0.007	0.009
	2.5 percentile	0.770	0.788	0.004
	97.5 percentile	0.790	0.816	0.038

continued on next page

Table B.4: Pooled random-intercept tobit regression of log(real annual earnings) on log(real GDP per capita), white, U.S.-born men ages 23-61 with exactly 12 years of education in Health and Retirement Study (combined Social Security and W-2 earnings).

continued from previous page

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
s.d.(u_{jt})	mean	0.204	0.222	0.019
	s.d.	0.025	0.034	0.042
	2.5 percentile	0.164	0.170	-0.060
	97.5 percentile	0.264	0.306	0.109
s.d.(a_i)	mean	0.642	0.640	-0.001
	s.d.	0.020	0.029	0.035
	2.5 percentile	0.602	0.582	-0.068
	97.5 percentile	0.683	0.697	0.067

Table B.5: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate personal consumption expenditure per capita), white, U.S.-born men ages 23-61 with exactly 12 years of education in Health and Retirement Study (combined Social Security and W-2 earnings).

coefficient	posterior statistic	risk tolerance low	risk tolerance high	risk tolerance difference
log(aggregate)	mean	2.356	2.541	0.185
	s.d.	0.134	0.195	0.238
	2.5 percentile	2.082	2.166	-0.265
	97.5 percentile	2.623	2.927	0.663
experience	mean	0.066	0.060	-0.005
	s.d.	0.005	0.008	0.010
	2.5 percentile	0.055	0.044	-0.025
	97.5 percentile	0.075	0.076	0.014
experience^2	mean	-0.002	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.002	-0.002	0.000
	97.5 percentile	-0.001	-0.001	0.001
veteran	mean	-0.088	0.068	0.156
	s.d.	0.053	0.080	0.094
	2.5 percentile	-0.193	-0.092	-0.026
	97.5 percentile	0.013	0.227	0.348
t	mean	-0.018	-0.030	-0.011
	s.d.	0.008	0.013	0.015
	2.5 percentile	-0.035	-0.057	-0.041
	97.5 percentile	-0.004	-0.005	0.020
t^2	mean	-0.001	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.001	0.000
	97.5 percentile	-0.001	0.000	0.001
constant	mean	-11.969	-13.540	-1.571
	s.d.	1.222	1.784	2.190
	2.5 percentile	-14.380	-17.091	-5.923
	97.5 percentile	-9.509	-10.165	2.576
s.d.(e_{it})	mean	0.780	0.801	0.021
	s.d.	0.005	0.007	0.009
	2.5 percentile	0.771	0.787	0.004
	97.5 percentile	0.789	0.817	0.038

continued on next page

Table B.5: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate personal consumption expenditure per capita), white, U.S.-born men ages 23-61 with exactly 12 years of education in Health and Retirement Study (combined Social Security and W-2 earnings).

continued from previous page

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
s.d.(u_{jt})	mean	0.209	0.227	0.018
	s.d.	0.027	0.036	0.046
	2.5 percentile	0.166	0.176	-0.072
	97.5 percentile	0.270	0.323	0.120
s.d.(a_i)	mean	0.643	0.639	-0.004
	s.d.	0.019	0.028	0.035
	2.5 percentile	0.608	0.584	-0.073
	97.5 percentile	0.683	0.695	0.066

Table B.6: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate wages and salaries per capita), white, U.S.-born men ages 23-61 with exactly 12 years of education in Health and Retirement Study (combined Social Security and W-2 earnings).

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
log(aggregate)	mean	2.125	2.445	0.319
	s.d.	0.091	0.139	0.166
	2.5 percentile	1.941	2.178	-0.019
	97.5 percentile	2.299	2.718	0.643
experience	mean	0.066	0.060	-0.006
	s.d.	0.003	0.005	0.006
	2.5 percentile	0.060	0.050	-0.018
	97.5 percentile	0.072	0.070	0.007
experience^2	mean	-0.002	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.002	-0.002	0.000
	97.5 percentile	-0.001	-0.001	0.000
veteran	mean	-0.090	0.064	0.153
	s.d.	0.049	0.074	0.087
	2.5 percentile	-0.182	-0.082	-0.019
	97.5 percentile	0.011	0.210	0.323
t	mean	-0.019	-0.035	-0.016
	s.d.	0.008	0.012	0.014
	2.5 percentile	-0.033	-0.058	-0.043
	97.5 percentile	-0.005	-0.012	0.012
t^2	mean	0.000	0.000	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.001	0.000
	97.5 percentile	0.000	0.000	0.001
constant	mean	-9.505	-12.196	-2.691
	s.d.	0.821	1.225	1.468
	2.5 percentile	-11.088	-14.644	-5.588
	97.5 percentile	-7.846	-9.811	0.281
s.d.(e_{it})	mean	0.780	0.802	0.021
	s.d.	0.005	0.007	0.009
	2.5 percentile	0.771	0.788	0.005
	97.5 percentile	0.789	0.817	0.039

continued on next page

Table B.6: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate wages and salaries per capita), white, U.S.-born men ages 23-61 with exactly 12 years of education in Health and Retirement Study (combined Social Security and W-2 earnings).

continued from previous page

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
s.d.(u_{jt})	mean	0.207	0.223	0.016
	s.d.	0.028	0.034	0.045
	2.5 percentile	0.161	0.172	-0.070
	97.5 percentile	0.272	0.307	0.113
s.d.(a_i)	mean	0.645	0.643	-0.002
	s.d.	0.020	0.031	0.039
	2.5 percentile	0.607	0.583	-0.075
	97.5 percentile	0.687	0.709	0.082

C Estimates using only Social Security earnings

Tables 2 and 3 in the paper, and section B of this appendix, report estimates of the relationship between preferences and income processes using income data that combine Social Security covered earnings in years before 1980 with W-2 earnings in 1980 and later years. The following tables repeat the analysis using Social Security covered earnings in all years. Tables C.1 and C.2 are equivalent to tables 2 and 3, respectively, in the paper. Tables C.3 to C.8 are equivalent to tables B.1 to B.6 in this appendix.

Discussion

The results using Social Security earnings are qualitatively similar to those using the combined earnings series and generally continue to support the hypothesis that more risk-tolerant people have incomes that carry more idiosyncratic and more aggregate risk.

There are a few exceptions. In table C.1, which estimates common coefficients for each risk-tolerance group, personal consumption expenditures have the same effect on the incomes of both groups if the sample is restricted to white, U.S.-born high school graduates. However, personal consumption expenditures have a larger effect on the more risk-tolerant group when considering all men or when estimating a different coefficient for each man (as in table C.2), and other aggregate variables have a larger effect on the more risk-tolerant group in all specifications. Also in table C.1, when the sample is restricted to white, U.S.-born high school graduates, both risk-tolerance groups appear to face the same idiosyncratic risk. However, idiosyncratic risk is larger for the more risk-tolerant group when we consider all men or when we estimate separate coefficients for each man. Finally, the coefficients on the aggregate variable are generally larger when using Social Security earnings than when using the combined earnings series.

Table C.1: Pooled random-intercept tobit regressions of log(real annual earnings) on aggregate variables, men ages 23-61 in Health and Retirement Study (Social Security earnings only).

Aggregate variable:	GDP[a]		Personal consumption[a]		Wages/salaries[b]	
	Risk tolerance		Risk tolerance		Risk tolerance	
Parameter	low	high	low	high	low	high
			A. All men			
coefficient on	2.781	3.018	2.524	2.607	2.478	2.676
log(aggregate shock)	(0.050)	(0.074)	(0.062)	(0.082)	(0.042)	(0.064)
95% CI for difference	[0.062, 0.418]		[−0.110, 0.286]		[0.049, 0.343]	
standard deviation of	0.814	0.895	0.814	0.895	0.814	0.895
idiosyncratic error	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
95% CI for difference	[0.072, 0.091]		[0.072, 0.090]		[0.072, 0.091]	
observations	67,814	40,058	67,814	40,058	67,814	40,058
			B. White, U.S.-born men with exactly 12 years of education			
coefficient on	2.744	2.980	2.557	2.551	2.368	2.558
log(aggregate shock)	(0.096)	(0.140)	(0.116)	(0.168)	(0.086)	(0.129)
95% CI for difference	[−0.092, 0.583]		[−0.412, 0.421]		[−0.111, 0.506]	
standard deviation of	0.760	0.759	0.760	0.758	0.760	0.759
idiosyncratic error	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)
95% CI for difference	[−0.018, 0.016]		[−0.019, 0.015]		[−0.018, 0.015]	
observations	21,837	9,911	21,837	9,911	21,837	9,911

Regressions also include individual and year random effects and control for experience, experience squared, time trend, time trend squared and an indicator for veteran status. Regressions in panel A also include years of education and indicators for immigrants and whites. See tables C.3 to C.8 for estimated coefficients on controls. Table shows posterior means of parameters, with posterior standard deviations in parentheses. Figures in square brackets are 2.5th and 97.5th percentiles of the posterior distribution for the difference between low- and high-risk-tolerance groups. [a]Per capita, chained 2000 dollars. [b]Per capita, deflated by GDP deflator for personal consumption.

Table C.2: Mean parameter estimates from individual tobit regressions of log(real annual earnings) on aggregate variables, men ages 23-61 in Health and Retirement Study (Social Security earnings only).

Aggregate variable:	GDP[a]		Personal consumption[a]		Wages/salaries[b]	
	Risk tolerance		Risk tolerance		Risk tolerance	
Parameter	low	high	low	high	low	high
A. All men						
coefficient on	2.377	2.751	2.315	2.798	1.852	2.234
log(aggregate shock)	(0.156)	(0.225)	(0.190)	(0.260)	(0.134)	(0.191)
test for difference	$t = 1.37$, $p = 0.09$		$t = 1.50$, $p = 0.07$		$t = 1.64$, $p = 0.05$	
standard deviation of	0.484	0.546	0.482	0.542	0.482	0.542
idiosyncratic error	(0.007)	(0.010)	(0.007)	(0.010)	(0.007)	(0.010)
test for difference	$t = 5.18$, $p = 0.00$		$t = 4.99$, $p = 0.00$		$t = 4.97$, $p = 0.00$	
workers	2,234	1,354	2,224	1,355	2,226	1,343
B. White, U.S.-born men with exactly 12 years of education						
coefficient on	2.528	2.855	2.651	2.773	1.852	2.111
log(aggregate shock)	(0.280)	(0.392)	(0.353)	(0.434)	(0.238)	(0.331)
test for difference	$t = 0.68$, $p = 0.25$		$t = 0.22$, $p = 0.41$		$t = 0.64$, $p = 0.26$	
standard deviation of	0.440	0.465	0.437	0.464	0.437	0.464
idiosyncratic error	(0.013)	(0.018)	(0.013)	(0.018)	(0.013)	(0.018)
test for difference	$t = 1.14$, $p = 0.13$		$t = 1.25$, $p = 0.11$		$t = 1.24$, $p = 0.11$	
workers	658	311	658	307	661	309

Regressions also control for experience and experience squared. Workers were dropped if tobit estimation did not converge. Means weighted by number of uncensored observations on worker's earnings. Standard error of mean in parentheses. P-values are for test of null hypothesis that mean does not depend on risk tolerance, against alternative that mean is higher in high-risk-tolerance group. [a]Per capita, chained 2000 dollars. [b]Per capita, deflated by GDP deflator for personal consumption.

Table C.3: Pooled random-intercept tobit regression of log(real annual earnings) on log(real GDP per capita), all men ages 23-61 in Health and Retirement Study (Social Security earnings only).

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
log(aggregate)	mean	2.781	3.018	0.237
	s.d.	0.050	0.074	0.089
	2.5 percentile	2.682	2.872	0.062
	97.5 percentile	2.879	3.168	0.418
education	mean	0.044	0.055	0.012
	s.d.	0.005	0.007	0.008
	2.5 percentile	0.034	0.042	-0.004
	97.5 percentile	0.054	0.068	0.029
experience	mean	0.067	0.069	0.002
	s.d.	0.003	0.004	0.005
	2.5 percentile	0.062	0.062	-0.007
	97.5 percentile	0.072	0.077	0.012
experience^2	mean	-0.001	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.002	0.000
	97.5 percentile	-0.001	-0.001	0.000
white	mean	0.485	0.355	-0.130
	s.d.	0.045	0.060	0.072
	2.5 percentile	0.398	0.236	-0.270
	97.5 percentile	0.576	0.469	0.008
immigrant	mean	0.036	0.192	0.157
	s.d.	0.069	0.084	0.108
	2.5 percentile	-0.102	0.028	-0.045
	97.5 percentile	0.173	0.352	0.378
veteran	mean	-0.023	0.040	0.063
	s.d.	0.037	0.050	0.063
	2.5 percentile	-0.096	-0.052	-0.055
	97.5 percentile	0.048	0.140	0.190
t	mean	-0.030	-0.036	-0.006
	s.d.	0.006	0.007	0.009
	2.5 percentile	-0.041	-0.050	-0.024
	97.5 percentile	-0.019	-0.020	0.013

continued on next page

Table C.3: Pooled random-intercept tobit regression of log(real annual earnings) on log(real GDP per capita), all men ages 23-61 in Health and Retirement Study (Social Security earnings only).

continued from previous page

coefficient	posterior statistic	risk tolerance low	risk tolerance high	risk tolerance difference
t^2	mean	-0.001	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.001	0.000
	97.5 percentile	-0.001	-0.001	0.000
constant	mean	-18.291	-20.551	-2.260
	s.d.	0.483	0.680	0.831
	2.5 percentile	-19.235	-21.882	-3.899
	97.5 percentile	-17.337	-19.237	-0.614
s.d.(e_{it})	mean	0.814	0.895	0.081
	s.d.	0.003	0.004	0.005
	2.5 percentile	0.809	0.887	0.072
	97.5 percentile	0.819	0.903	0.091
s.d.(u_{jt})	mean	0.221	0.254	0.033
	s.d.	0.027	0.033	0.044
	2.5 percentile	0.175	0.200	-0.050
	97.5 percentile	0.280	0.332	0.126
s.d.(a_i)	mean	0.774	0.803	0.030
	s.d.	0.013	0.017	0.021
	2.5 percentile	0.748	0.771	-0.011
	97.5 percentile	0.797	0.836	0.072

Table C.4: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate personal consumption expenditure per capita), all men ages 23-61 in Health and Retirement Study (Social Security earnings only).

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
log(aggregate)	mean	2.524	2.607	0.083
	s.d.	0.062	0.082	0.100
	2.5 percentile	2.410	2.442	-0.110
	97.5 percentile	2.653	2.770	0.286
education	mean	0.043	0.054	0.011
	s.d.	0.005	0.006	0.008
	2.5 percentile	0.033	0.042	-0.005
	97.5 percentile	0.053	0.067	0.027
experience	mean	0.066	0.069	0.003
	s.d.	0.003	0.004	0.005
	2.5 percentile	0.060	0.060	-0.007
	97.5 percentile	0.072	0.077	0.013
experience^2	mean	-0.001	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.002	0.000
	97.5 percentile	-0.001	-0.001	0.000
white	mean	0.484	0.355	-0.128
	s.d.	0.044	0.060	0.072
	2.5 percentile	0.399	0.239	-0.276
	97.5 percentile	0.573	0.470	0.014
immigrant	mean	0.030	0.189	0.159
	s.d.	0.074	0.084	0.111
	2.5 percentile	-0.114	0.030	-0.063
	97.5 percentile	0.180	0.360	0.374
veteran	mean	-0.024	0.040	0.064
	s.d.	0.036	0.051	0.063
	2.5 percentile	-0.092	-0.052	-0.056
	97.5 percentile	0.049	0.147	0.184
t	mean	-0.030	-0.032	-0.002
	s.d.	0.006	0.007	0.009
	2.5 percentile	-0.042	-0.047	-0.021
	97.5 percentile	-0.019	-0.018	0.015

continued on next page

Table C.4: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate personal consumption expenditure per capita), all men ages 23-61 in Health and Retirement Study (Social Security earnings only).

continued from previous page

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
t^2	mean	-0.001	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.001	0.000
	97.5 percentile	-0.001	-0.001	0.000
constant	mean	-14.644	-15.426	-0.782
	s.d.	0.566	0.748	0.916
	2.5 percentile	-15.799	-16.959	-2.669
	97.5 percentile	-13.589	-13.962	1.004
s.d.(e_{it})	mean	0.814	0.895	0.081
	s.d.	0.003	0.004	0.005
	2.5 percentile	0.808	0.888	0.072
	97.5 percentile	0.819	0.903	0.090
s.d.(u_{jt})	mean	0.225	0.255	0.030
	s.d.	0.026	0.034	0.042
	2.5 percentile	0.180	0.200	-0.047
	97.5 percentile	0.282	0.332	0.119
s.d.(a_i)	mean	0.773	0.803	0.030
	s.d.	0.012	0.017	0.021
	2.5 percentile	0.749	0.771	-0.012
	97.5 percentile	0.797	0.837	0.071

Table C.5: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate wages and salaries per capita), all men ages 23-61 in Health and Retirement Study (Social Security earnings only).

coefficient	posterior statistic	risk tolerance low	risk tolerance high	risk tolerance difference
log(aggregate)	mean	2.478	2.676	0.198
	s.d.	0.042	0.064	0.076
	2.5 percentile	2.392	2.555	0.049
	97.5 percentile	2.556	2.800	0.343
education	mean	0.044	0.056	0.012
	s.d.	0.005	0.006	0.008
	2.5 percentile	0.034	0.045	-0.004
	97.5 percentile	0.055	0.068	0.028
experience	mean	0.066	0.069	0.003
	s.d.	0.002	0.003	0.003
	2.5 percentile	0.063	0.064	-0.004
	97.5 percentile	0.070	0.074	0.009
experience^2	mean	-0.001	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.002	0.000
	97.5 percentile	-0.001	-0.001	0.000
white	mean	0.487	0.354	-0.132
	s.d.	0.044	0.060	0.077
	2.5 percentile	0.398	0.234	-0.283
	97.5 percentile	0.571	0.474	0.020
immigrant	mean	0.031	0.192	0.161
	s.d.	0.071	0.089	0.115
	2.5 percentile	-0.106	0.015	-0.065
	97.5 percentile	0.171	0.362	0.392
veteran	mean	-0.024	0.040	0.064
	s.d.	0.037	0.049	0.061
	2.5 percentile	-0.098	-0.057	-0.050
	97.5 percentile	0.046	0.137	0.184
t	mean	-0.035	-0.041	-0.005
	s.d.	0.005	0.007	0.008
	2.5 percentile	-0.045	-0.054	-0.022
	97.5 percentile	-0.025	-0.027	0.011

continued on next page

Table C.5: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate wages and salaries per capita), all men ages 23-61 in Health and Retirement Study (Social Security earnings only).

continued from previous page

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
t^2	mean	0.000	0.000	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.001	0.000
	97.5 percentile	0.000	0.000	0.000
constant	mean	-13.788	-15.562	-1.774
	s.d.	0.383	0.571	0.678
	2.5 percentile	-14.528	-16.728	-3.041
	97.5 percentile	-12.989	-14.443	-0.434
s.d.(e_{it})	mean	0.814	0.895	0.082
	s.d.	0.003	0.004	0.005
	2.5 percentile	0.809	0.887	0.072
	97.5 percentile	0.819	0.903	0.091
s.d.(u_{jt})	mean	0.208	0.241	0.032
	s.d.	0.026	0.033	0.041
	2.5 percentile	0.166	0.188	-0.046
	97.5 percentile	0.265	0.314	0.120
s.d.(a_i)	mean	0.776	0.806	0.030
	s.d.	0.013	0.017	0.022
	2.5 percentile	0.751	0.775	-0.012
	97.5 percentile	0.802	0.840	0.071

Table C.6: Pooled random-intercept tobit regression of log(real annual earnings) on log(real GDP per capita), white, U.S.-born men ages 23-61 with exactly 12 years of education in Health and Retirement Study (Social Security earnings only).

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
log(aggregate)	mean	2.744	2.980	0.236
	s.d.	0.096	0.140	0.169
	2.5 percentile	2.546	2.719	-0.092
	97.5 percentile	2.933	3.259	0.583
experience	mean	0.060	0.053	-0.007
	s.d.	0.005	0.007	0.009
	2.5 percentile	0.050	0.038	-0.024
	97.5 percentile	0.069	0.068	0.010
experience^2	mean	-0.001	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.002	-0.002	0.000
	97.5 percentile	-0.001	-0.001	0.000
veteran	mean	-0.094	0.008	0.102
	s.d.	0.061	0.087	0.104
	2.5 percentile	-0.210	-0.158	-0.097
	97.5 percentile	0.022	0.176	0.309
t	mean	-0.021	-0.037	-0.016
	s.d.	0.009	0.014	0.017
	2.5 percentile	-0.038	-0.064	-0.047
	97.5 percentile	-0.004	-0.009	0.017
t^2	mean	-0.001	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.001	0.000
	97.5 percentile	-0.001	0.000	0.001
constant	mean	-16.774	-18.811	-2.037
	s.d.	0.907	1.318	1.605
	2.5 percentile	-18.575	-21.432	-5.155
	97.5 percentile	-14.907	-16.416	1.117
s.d.(e_{it})	mean	0.760	0.759	-0.001
	s.d.	0.005	0.007	0.009
	2.5 percentile	0.751	0.745	-0.018
	97.5 percentile	0.770	0.773	0.016

continued on next page

Table C.6: Pooled random-intercept tobit regression of log(real annual earnings) on log(real GDP per capita), white, U.S.-born men ages 23-61 with exactly 12 years of education in Health and Retirement Study (Social Security earnings only).

continued from previous page

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
s.d.(u_{jt})	mean	0.220	0.253	0.033
	s.d.	0.030	0.043	0.055
	2.5 percentile	0.174	0.190	-0.069
	97.5 percentile	0.289	0.367	0.151
s.d.(a_i)	mean	0.739	0.710	-0.028
	s.d.	0.024	0.032	0.039
	2.5 percentile	0.691	0.651	-0.100
	97.5 percentile	0.785	0.775	0.044

Table C.7: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate personal consumption expenditure per capita), white, U.S.-born men ages 23-61 with exactly 12 years of education in Health and Retirement Study (Social Security earnings only).

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
log(aggregate)	mean	2.557	2.551	-0.006
	s.d.	0.116	0.168	0.207
	2.5 percentile	2.314	2.238	-0.412
	97.5 percentile	2.781	2.880	0.421
experience	mean	0.059	0.052	-0.007
	s.d.	0.005	0.008	0.009
	2.5 percentile	0.048	0.036	-0.025
	97.5 percentile	0.069	0.067	0.011
experience^2	mean	-0.001	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.002	-0.002	0.000
	97.5 percentile	-0.001	-0.001	0.000
veteran	mean	-0.093	0.010	0.103
	s.d.	0.062	0.087	0.106
	2.5 percentile	-0.217	-0.168	-0.110
	97.5 percentile	0.028	0.184	0.315
t	mean	-0.023	-0.031	-0.008
	s.d.	0.009	0.014	0.016
	2.5 percentile	-0.041	-0.058	-0.040
	97.5 percentile	-0.005	-0.004	0.023
t^2	mean	-0.001	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.001	0.000
	97.5 percentile	0.000	0.000	0.001
constant	mean	-13.784	-13.591	0.193
	s.d.	1.070	1.527	1.892
	2.5 percentile	-15.833	-16.537	-3.749
	97.5 percentile	-11.584	-10.701	3.856
s.d.(e_{it})	mean	0.760	0.758	-0.002
	s.d.	0.005	0.007	0.009
	2.5 percentile	0.750	0.744	-0.019
	97.5 percentile	0.770	0.772	0.015
s.d.(u_{jt})	mean	0.223	0.251	0.028

continued on next page

Table C.7: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate personal consumption expenditure per capita), white, U.S.-born men ages 23-61 with exactly 12 years of education in Health and Retirement Study (Social Security earnings only).

continued from previous page

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
	s.d.	0.031	0.037	0.047
	2.5 percentile	0.172	0.193	-0.061
	97.5 percentile	0.293	0.334	0.133
s.d.(a_i)	mean	0.738	0.708	-0.030
	s.d.	0.023	0.031	0.038
	2.5 percentile	0.698	0.650	-0.103
	97.5 percentile	0.784	0.773	0.049

Table C.8: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate wages and salaries per capita), white, U.S.-born men ages 23-61 with exactly 12 years of education in Health and Retirement Study (Social Security earnings only).

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
log(aggregate)	mean	2.368	2.558	0.190
	s.d.	0.086	0.129	0.159
	2.5 percentile	2.202	2.316	-0.111
	97.5 percentile	2.536	2.822	0.506
experience	mean	0.058	0.051	-0.007
	s.d.	0.003	0.005	0.006
	2.5 percentile	0.051	0.041	-0.020
	97.5 percentile	0.065	0.061	0.005
experience^2	mean	-0.001	-0.001	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.002	-0.002	0.000
	97.5 percentile	-0.001	-0.001	0.000
veteran	mean	-0.094	0.010	0.104
	s.d.	0.059	0.082	0.100
	2.5 percentile	-0.216	-0.156	-0.092
	97.5 percentile	0.023	0.166	0.297
t	mean	-0.022	-0.038	-0.015
	s.d.	0.009	0.012	0.015
	2.5 percentile	-0.040	-0.062	-0.045
	97.5 percentile	-0.006	-0.014	0.013
t^2	mean	0.000	0.000	0.000
	s.d.	0.000	0.000	0.000
	2.5 percentile	-0.001	-0.001	0.000
	97.5 percentile	0.000	0.000	0.001
constant	mean	-11.662	-13.150	-1.488
	s.d.	0.775	1.159	1.443
	2.5 percentile	-13.180	-15.527	-4.384
	97.5 percentile	-10.107	-11.012	1.337
s.d.(e_{it})	mean	0.760	0.759	-0.001
	s.d.	0.005	0.007	0.009
	2.5 percentile	0.750	0.745	-0.018
	97.5 percentile	0.769	0.773	0.015

continued on next page

Table C.8: Pooled random-intercept tobit regression of log(real annual earnings) on log(real aggregate wages and salaries per capita), white, U.S.-born men ages 23-61 with exactly 12 years of education in Health and Retirement Study (Social Security earnings only).

continued from previous page

	posterior	risk tolerance		
coefficient	statistic	low	high	difference
s.d.(u_{jt})	mean	0.215	0.244	0.029
	s.d.	0.030	0.040	0.048
	2.5 percentile	0.168	0.186	-0.065
	97.5 percentile	0.282	0.340	0.134
s.d.(a_i)	mean	0.740	0.710	-0.029
	s.d.	0.023	0.030	0.038
	2.5 percentile	0.697	0.655	-0.107
	97.5 percentile	0.785	0.770	0.047

D Factor model estimation in unbalanced panels

The model to be estimated, as in the paper, is

$$y_{it} = a_i + b_i d_t + r_i t + g x_{it} + e_{it}. \tag{1}$$

Assume that $\mathrm{Var}[e_{it}] = \sigma_i^2$ for all t, that $\mathrm{Cov}[e_{is}, e_{it}] = 0$ for all $s \neq t$, and that $\mathrm{E}[x_{is} e_{it} | \boldsymbol{\iota}_i] = \mathrm{E}[\boldsymbol{\theta}_i e_{it} | \boldsymbol{\iota}_i] = \mathrm{E}[e_{it} | \boldsymbol{\iota}_i] = 0$ for all s and t. Also assume that the data are i.i.d. over i and that all of the variables have finite moments up to fourth order. Assume that g is an element of a known compact set. Finally, assume that at least some individuals are observed in all periods.

The estimator is

$$(\hat{\mathbf{d}}, \hat{g}) = \min_{\tilde{\mathbf{d}}, \tilde{g}} \frac{1}{N} \sum_{i=1}^{N} (\mathbf{u}_i^y - \tilde{g}\mathbf{u}_i^x)' \left[\mathbf{I}_i - \tilde{\boldsymbol{\Delta}}_i^d \left(\tilde{\boldsymbol{\Delta}}_i^{d\prime} \hat{\boldsymbol{\Delta}}_i^d \right)^{-1} \tilde{\boldsymbol{\Delta}}_i^{d\prime} \right] (\mathbf{u}_i^y - \tilde{g}\mathbf{u}_i^x), \tag{2}$$

subject to appropriate normalizations on $\tilde{\mathbf{d}}$.

Proposition. $(\hat{\mathbf{d}}, \hat{g})$ *converge in probability to* $(\mathbf{d}, g)$ *as* N *goes to infinity with* T *fixed.*

Proof. Because the data are i.i.d. with finite fourth moments, a uniform law of large numbers applies and, as $N \to \infty$, the objective function in (2) converges uniformly in probability to

$$f(\tilde{\mathbf{d}}, \tilde{g}) = \mathrm{E}\left[(\mathbf{u}_i^y - \tilde{g}\mathbf{u}_i^x)' \left[\mathbf{I}_i - \tilde{\boldsymbol{\Delta}}_i^d \left(\tilde{\boldsymbol{\Delta}}_i^{d\prime} \tilde{\boldsymbol{\Delta}}_i^d \right)^{-1} \tilde{\boldsymbol{\Delta}}_i^{d\prime} \right] (\mathbf{u}_i^y - \mathbf{u}_i^x) \right]. \tag{3}$$

Standard results on extremum estimators (e.g., Newey and McFadden, 1994, theorem 2.1) say that if the parameter space is compact and the true parameters $\mathbf{d}$ uniquely minimize f, then the estimator $\hat{\mathbf{d}}$ converges in probability to $\mathbf{d}$ as $N \to \infty$. The normalizations restrict $\mathbf{d}$ to a compact set, and g is contained in a compact set by assumption. I now show that the true parameters uniquely minimize f.

Sort households into groups so that all households with the same $\boldsymbol{\iota}_i$ are in the same group. The matrices $\mathbf{I}_i$ and $\hat{\boldsymbol{\Delta}}_i^d$ are the same for all households in a group and can be indexed by groups. For a given group j, define

$$f_j(\tilde{\mathbf{d}}, \tilde{g}) = \mathrm{E}\left[(\mathbf{u}_i^y - \tilde{g}\mathbf{u}_i^x)' \left[\mathbf{I}_j - \tilde{\boldsymbol{\Delta}}_j^d \left(\tilde{\boldsymbol{\Delta}}_j^{d\prime} \tilde{\boldsymbol{\Delta}}_j^d \right)^{-1} \tilde{\boldsymbol{\Delta}}_j^{d\prime} \right] (\mathbf{u}_i^y - \tilde{g}\mathbf{u}_i^x) \middle| i \in \text{group } j \right]. \tag{4}$$

Then $f(\tilde{\mathbf{d}}, \tilde{g}) = \sum_j \Pr(i \in \text{group } j) f_j(\tilde{\mathbf{d}}, \tilde{g})$.

Consider the group $j = 1$ whose members are observed in all periods. Write $\mathbf{V}_1(\tilde{\mathbf{d}}) = \mathbf{I} - \tilde{\boldsymbol{\Delta}}_1^d (\tilde{\boldsymbol{\Delta}}_1^{d\prime} \tilde{\boldsymbol{\Delta}}_1^d)^{-1} \tilde{\boldsymbol{\Delta}}_1^{d\prime}$. Also write $\boldsymbol{\theta}_i = [b_i \; r_i]'$. Let $\mathrm{E}_1(\cdot)$ denote $\mathrm{E}(\cdot | i \in \text{group } 1)$. Then, using the stated assumptions,

$$
\begin{aligned}
f_1(\tilde{\mathbf{d}}, \tilde{g}) &= \mathrm{E}_1\left[(\boldsymbol{\Delta}_1^{d\prime}\boldsymbol{\theta}_i + \mathbf{u}_i^e + (g-\tilde{g})\mathbf{u}_i^x)' V_1(\tilde{\mathbf{d}})(\boldsymbol{\Delta}_1^{d\prime}\boldsymbol{\theta}_i + \mathbf{u}_i^e + (g-\tilde{g})\mathbf{u}_i^x)\right] \\
&= \mathrm{E}_1\left[(\boldsymbol{\Delta}_1^{d\prime}\mathrm{E}_1\boldsymbol{\theta}_i + (g-\tilde{g})\mathrm{E}_1\mathbf{u}_i^x)' V_1(\tilde{\mathbf{d}})(\boldsymbol{\Delta}_1^{d\prime}\mathrm{E}_1\boldsymbol{\theta}_i + (g-\tilde{g})\mathrm{E}_1\mathbf{u}_i^x)\right] \\
&\quad + \mathrm{E}_1\left[(\boldsymbol{\Delta}_1^{d\prime}(\boldsymbol{\theta}_i - \mathrm{E}_1\boldsymbol{\theta}_i) + (g-\tilde{g})(\mathbf{u}_i^x - \mathrm{E}_1\mathbf{u}_i^x)' V_1(\tilde{\mathbf{d}})(\boldsymbol{\Delta}_1^{d\prime}(\boldsymbol{\theta}_i - \mathrm{E}_1\boldsymbol{\theta}_i) + (g-\tilde{g})(\mathbf{u}_i^x - \mathrm{E}_1\mathbf{u}_i^x)'\right] \\
&\quad + \mathrm{E}_1\left[\mathbf{e}_i{}'\left(\mathbf{I} - \frac{\boldsymbol{\iota}\boldsymbol{\iota}'}{\boldsymbol{\iota}'\boldsymbol{\iota}}\right) V_1(\tilde{\mathbf{d}})\left(\mathbf{I} - \frac{\boldsymbol{\iota}\boldsymbol{\iota}'}{\boldsymbol{\iota}'\boldsymbol{\iota}}\right)\mathbf{e}_i\right]
\end{aligned} \tag{5}
$$

Since e_i is serially uncorrelated and homoscedastic, the value of the third term is $\mathrm{E}_1[\sigma_i^2] \cdot \mathrm{tr}\left[\left(\mathbf{I} - \frac{\boldsymbol{\iota}\boldsymbol{\iota}'}{\boldsymbol{\iota}'\boldsymbol{\iota}}\right) V_1(\tilde{\mathbf{d}}) \left(\mathbf{I} - \frac{\boldsymbol{\iota}\boldsymbol{\iota}'}{\boldsymbol{\iota}'\boldsymbol{\iota}}\right)\right]$. Since the matrix involved is idempotent, its trace equals its rank, which is $T-3$ for any $\tilde{\mathbf{d}}$ not collinear with t and a constant, and $T-2$ otherwise. The true $\mathbf{d}$ is not collinear with t and a constant and therefore minimizes the third term.

Furthermore, $V_1(\tilde{\mathbf{d}})$ is symmetric and positive semidefinite, so the first two terms are non-negative. The second term is zero if and only if

$$
(\boldsymbol{\Delta}_1^{d\prime}(\boldsymbol{\theta}_i - \mathrm{E}_1\boldsymbol{\theta}_i) + (g-\tilde{g})(\mathbf{u}_i^x - \mathrm{E}_1\mathbf{u}_i^x) \propto \tilde{\boldsymbol{\Delta}}_1^d \tag{6}
$$

for all realizations of $\boldsymbol{\theta}_i$ and $\mathbf{x}_i$. Because $\mathbf{x}_i$ is not a constant vector for at least some i, this requires $\tilde{g} = g$ and $\tilde{\boldsymbol{\Delta}}_1^d = k\boldsymbol{\Delta}_1^d$ for some constant k. Since one column of $\tilde{\boldsymbol{\Delta}}_1^d$ is the known time trend, we must have $k = 1$ and $\tilde{\boldsymbol{\Delta}}_1^d = \boldsymbol{\Delta}_1^d$. When $\tilde{g} = g$ and $\tilde{\boldsymbol{\Delta}}_1^d = \boldsymbol{\Delta}_1^d$, the first term is zero as well. Thus f_1 is minimized when $\tilde{g} = g$ and $\tilde{\mathbf{d}} = \mathbf{d}$.

The same algebra applied to any group $j \neq 1$ shows that f_j is minimized when $\tilde{g} = g$ and $\tilde{\boldsymbol{\Delta}}_j^d = \boldsymbol{\Delta}_j^d$. We could have $\tilde{\boldsymbol{\Delta}}_j^d = \boldsymbol{\Delta}_j^d$ without having $\tilde{\mathbf{d}} = \mathbf{d}$ because $\tilde{\mathbf{d}}$ and $\mathbf{d}$ could differ at dates where group j is not observed. However, this would strictly increase f_1 and would not decrease f_j. Therefore, $\mathbf{d}$ and g uniquely minimizes $f(\tilde{\mathbf{d}}, g) = \sum_j \Pr(i \in j) f_j(\tilde{\mathbf{d}}, g)$. $\square$

References

Chib, Siddhartha, 2001, "Markov Chain Monte Carlo Methods: Computation and Inference," in Heckman, James J. and Edward Leamer, editors, *Handbook of Econometrics*, volume V, pp. 3569–3649, Amsterdam: Elsevier.

Kiefer, Nicholas M., 1980, "Estimation of Fixed Effect Models for Time Series of Cross-Sections With Arbitrary Intertemporal Covariance," *Journal of Econometrics* 14(2), 195–202.

Newey, Whitney K. and Daniel McFadden, 1994, "Large Sample Estimation and Hypothesis Testing," in Engle, Robert F. and Daniel L. McFadden, editors, *Handbook of Econometrics*, volume IV, pp. 2111–2245, Amsterdam: Elsevier.