Online Appendix
Wealth Distribution and Social Mobility in the U.S.: A Quantitative Approach

Jess Benhabib, Alberto Bisin, Mi Luo

A. Methods

A.1 Numerical solution

We solve the model for value functions and policy functions with the *Collocation method* in Miranda and Fackler (2004).

Each agent's recursive problem in the baseline case is

$$
\begin{aligned}
V_t(a,r,w) &= \max_c \mathbf{1}\{t<T\}\{u(c)+\beta V(a',r,w,t+1)\}+\mathbf{1}\{t=T\}\{u(c)+e(a')\} \\
s.t. & \\
a' &= (1+r)(a-c)+w \\
c &\leq a \\
c &\geq 0
\end{aligned}
$$

where we have explicitly allowed for the dependence on (r,w).

The problem can be written as

$$
\begin{aligned}
V_1(a,r,w) &= \max_{c\in[0,a]} u(c)+\beta V_2((1+r)(a-c)+w,r,w) \\
V_2(a,r,w) &= \max_{c\in[0,a]} u(c)+\beta V_3((1+r)(a-c)+w,r,w) \\
&\vdots \\
V_{T-1}(a,r,w) &= \max_{c\in[0,a]} u(c)+\beta V_T((1+r)(a-c)+w,r,w) \\
V_T(a,r,w) &= \max_{c\in[0,a]} u(c)+e((1+r)(a-c)+w)
\end{aligned}
$$

The parameters are: $\{\beta,T,u(c),e(a)\}$. Set $T=6$ for simplicity and we can decrease β to account for the longer length of periods.

The state space is $s=(a,z)$; $z=(r,w)$ is the exogenous state which has transition matrix $P=P_r\otimes P_w$

across generations, but is constant for each generation. The state space for z is discrete and so is enumerated $k = 1, \ldots, K$, where $K = N_r \times N_w$. Let $s = (s_1, s_2)$ and the choice variable $x = c$. The choice is consumption $x \in B(s)$, where

$$B(s) = [0, a]$$

Re-writing this as a system of six value functions

$$\begin{aligned}
V_1(s) &= \max_{x \in B(s)} F_1(s, x) + \beta V_2([(1+r)(s_1 - x) + w, s_2]) \\
&\vdots \\
V_T(s) &= \max_{x \in B(s)} F_2(s, x)
\end{aligned}$$

This is the system we will solve.

Approximation: Take $V_1, \ldots, V_T$ and approximate them on J collocation nodes $s_1, \ldots, s_J$ with a spline with J coefficients $c^1 = (c_1^1, \ldots, c_J^1)$, $c^2, \ldots, c^T$ and linear basis ϕ_j.

$$\begin{aligned}
V_1(s_i) &= \sum_{j=1}^{J} c_j^1 \phi_j(s_i) \\
&\vdots \\
V_T(s_i) &= \sum_{j=1}^{J} c_j^T \phi_j(s_i)
\end{aligned}$$

Let $c = (c^1, \ldots, c^T)$ and let $v_1(c^1) = [V_1(s_1), \ldots, V_1(s_J)]'$ and $v_2(c^2), \ldots, v_T(c^T)$ similarly defined for a given c. With $v(c) = [v_1(c^1)', \ldots, v_J(c^J)']'$ then

$$\begin{aligned}
v_1(s) &= \Phi c^1 \\
&\vdots \\
v_T(s) &= \Phi c^T
\end{aligned}$$

this is the *Collocation equation.*

Substituting the interpolants into the value functions

$$\begin{aligned}
\sum_{j=1}^{J} c_j^1 \phi_j(s_i) &= \max_{x \in B(s_i)} F_1(s_i, x) + \beta \sum_{j=1}^{J} c_j^2 \phi_j([(1+r)(s_{i,1} - x) + w, s_{i,2}]) \\
\sum_{j=1}^{J} c_j^2 \phi_j(s_i) &= \max_{x \in B(s_i)} F_1(s_i, x) + \beta \sum_{j=1}^{J} c_j^3 \phi_j([(1+r)(s_{i,1} - x) + w, s_{i,2}])
\end{aligned}$$

$$
\vdots
$$

$$
\sum_{j=1}^{J} c_j^T \phi_j(s_i) = \max_{x \in B(s_i)} F_2(s_i, x)
$$

The stacked system of value functions is

$$
\begin{aligned}
\Phi(s)c^1 &= F_1(s, x(s)) + \beta\Phi([(1+r)(s_1 - x(s)) + w, s_2])c^2 =: v_1(c^2) \\
\Phi(s)c^2 &= F_1(s, x(s)) + \beta\Phi([(1+r)(s_1 - x(s)) + w, s_2])c^3 =: v_2(c^3) \\
&\vdots \\
\Phi(s)c^T &= F_2(s, x(s))
\end{aligned}
$$

The zero system would be $\tilde{\Phi}(s)c - v(c) = 0$, where $\tilde{\Phi}$ is a block diagonal matrix of $\Phi's$.

A.2 Estimation

The estimation procedure we use, described below, is adapted from Guvenen (2016). The global stage is a multi-start algorithm where candidate parameter vectors are uniform Sobol (quasi-random) points. We typically take about 10,000 initial Sobol points for pre-testing and select the best 200 points (i.e., ranked by objective value) for the multiple restart procedure. The local minimization stage is performed with the Nelder-Mead's downhill simplex algorithm (which is slow but performs well on non-linear objectives). Within one evaluation, we draw 100,000 individuals randomly and simulate their entire wealth process initiated with zero wealth and the lowest earnings profile.

A.3 Standard errors

We use numerical derivatives to calculate the standard errors for the parameters in all the estimates. The procedure is standard. The variance-covariance matrix for parameter estimates is given by

$$
Q(W) = \left[\frac{\partial b(\theta_0)}{\partial(\theta)}' W \frac{\partial b(\theta_0)}{\partial(\theta)} \right]^{-1},
$$

where $\frac{\partial b(\theta_0)}{\partial(\theta)}$ is the derivative of the vector of moments with respect to the parameter vector. We calculate the derivatives numerically, i.e. perturbing θ and calculating new vector moments. Standard errors will then be the square roots of the diagonal elements of $Q(W)$.

We use bootstrapping to generate the standard errors for the statistics related to the return process, e.g. its mean, standard deviation, and autocorrelation coefficient. The procedure is standard.

We take the parameter values for generating the return process as given, i.e. the values for the five Markov states and the diagonal matrix of the transition matrix (hence the whole Markov transition matrix), then generate the return process a sufficiently large number of times. We then calculate the mean, standard errors and the autocorrelation coefficient directly using these series of the return processes.

B. Data

B.1 Labor earnings

The labor earnings data we use are adapted from the PSID, as cleaned by Heathcote, Perri and Violante (2010) - Sample C in their labeling.

We adopt the following procedure to obtain life-cycle age profiles, conditioning out year effects: i

1. Import household labor earnings - Sample C in Heathcote, Perri and Violante (2010). The exact variable is *redlabinc*, i.e. household labor income (head + wife for couples). Keep households with head aged between 25 and 60 (inclusive). Label this variable *inc*.

2. Take log of the household labor earnings, $\log(inc)$. Drop all the observations with zero labor earnings. Record the mean of $\log(inc)_{it}$ in the initial year (2002) as $\overline{\log(inc)}_{2002}$.

3. Regress $\log(inc)_{it}$ against a full set of year dummies, denoting residuals ϵ_{it}:

$$\log(inc)_{it} = \overline{\log(inc)}_{2002} + year_{1967-2001} + \epsilon_{it}.$$

4. Generate predicted log earnings as:

$$\widehat{\log(inc)}_{it} = \overline{\log(inc)}_{2002} + \epsilon_{it};$$

   and predicted earnings as:

$$\exp(\widehat{\log(inc)})_{it}.$$

5. Construct, with the generated predicted earnings, age-dependent decile values as follows:[1]

   (a) Calculate decile values of earnings for each age;

   (b) Calculate average decile earnings for each six-year age bin.

[1]This procedure maintains the distributional ranking of households across the life cycle and allows them to move across bins during the life-cycle.

B.2 Inter-generational labor earnings transitions

Chetty et al. (2014) construct a 100 by 100 transition matrix linking parental family income with child's income - see `http://equality-of-opportunity.org/images/online_data_tables.xls`, Online Table 1. The main sample they use is the Statistics of Income (SOI) annual cross-sections from 1980 to 1982 cohorts for children, linking children to their parents by using population tax records spanning 1996-2012. We in turn collapse this matrix into a 10 by 10 transition matrix, which we associate to labor earnings.[2]

B.3 Inter-generational wealth mobility

The alternative inter-generational wealth mobility matrix we use in Section 4.4.2 is estimated from the 2007-2009 SCF 2-year panel. For comparison we report on another matrix we obtain with the same underlying method but exploiting an age-independent wealth mobility matrix by Kennickell and Starr-McCluer (1997), who estimate a seven-state (bottom 25, 25-49, 50-74, 75-89, 90-94, top 2-5, top 1) six-year age-independent transition matrix from the 1983-89 SCF panel - Table 7:

$$T_{KS,6} = \begin{bmatrix} 0.672 & 0.246 & 0.063 & 0.018 & 0.001 & 0.000 & 0.000 \\ 0.246 & 0.495 & 0.190 & 0.042 & 0.019 & 0.007 & 0.000 \\ 0.066 & 0.192 & 0.480 & 0.208 & 0.037 & 0.016 & 0.000 \\ 0.021 & 0.082 & 0.329 & 0.418 & 0.113 & 0.036 & 0.002 \\ 0.011 & 0.071 & 0.212 & 0.301 & 0.225 & 0.177 & 0.004 \\ 0.000 & 0.028 & 0.164 & 0.104 & 0.180 & 0.430 & 0.094 \\ 0.000 & 0.031 & 0.024 & 0.061 & 0.045 & 0.247 & 0.593 \end{bmatrix}$$

As in the text, we use the model assumption that $a_0^n = a_T^{n-1}$ and reduce the problem to compute the intra-generational matrix whose component are transitions of the form $\Pr(a_T^{n-1} \in p \mid a_0^{n-1} \in p')$, which we obtain by raising to the power of 6 the age-independent 6 years matrix: [3]

[2]Chetty et al. (2014) also construct average income levels for both parent and child at age 29-30 - Online Table 2 - but do not provide life cycle data.

[3]We refer to a previous draft of this paper, Benhabib et al. (2015), NBER WP 21721, at http://www.nber.org/papers/w21721 for an estimate of our model matching this mobility matrix, with similar results to those obtained in Section 4.4.2 and in the baseline.

$$T_{KS,36} = \begin{bmatrix} 0.316 & 0.278 & 0.222 & 0.118 & 0.037 & 0.024 & 0.005 \\ 0.276 & 0.263 & 0.240 & 0.137 & 0.044 & 0.031 & 0.009 \\ 0.224 & 0.242 & 0.263 & 0.163 & 0.054 & 0.042 & 0.012 \\ 0.196 & 0.229 & 0.274 & 0.176 & 0.061 & 0.051 & 0.013 \\ 0.179 & 0.219 & 0.275 & 0.181 & 0.066 & 0.061 & 0.020 \\ 0.150 & 0.198 & 0.271 & 0.185 & 0.074 & 0.082 & 0.040 \\ 0.112 & 0.166 & 0.252 & 0.182 & 0.085 & 0.121 & 0.083 \end{bmatrix}$$

A similar procedure, producing similar results, can be adopted exploiting instead the age-independent wealth mobility matrix by Klevmarken, Lupton and Stafford (2003), who estimate a five-state (quintiles) five-year transition matrix from 1994- 1999 PSID data - Table 6.

C. Additional results

C.1 Full transition matrix for r in the baseline

The parameterization of the stochastic process for r we use is defined by 5 states r_i and 5 diagonal transition probabilities, $P(r^n = r_i \mid r^{n-1} = r_i)$, $i = 1, \ldots, 5$, restricting instead the 5×5 transition matrix as follows: $P(r^n = r_i \mid r^{n-1} = r_j) = P(r^n = r_i \mid r^{n-1} = r_i)e^{-\lambda j}$, $i = 1, 2, 3, 4$, $j \neq i$, λ such that $\sum_{j=1}^{5} P(r^n = r_i \mid r^{n-1} = r_j) = 1$; and $P(r^n = r_5 \mid r^{n-1} = r_j) = \frac{1}{4}\left(1 - P(r^n = r_5 \mid r^{n-1} = r_5)\right)$. For readers' convenience, we report here the full transition matrix for the return process r in the baseline estimation.

$$\begin{bmatrix} 0.0338 & 0.5013 & 0.2600 & 0.1349 & 0.0700 \\ 0.2876 & 0.2676 & 0.2876 & 0.1129 & 0.0443 \\ 0.1158 & 0.3163 & 0.1360 & 0.3163 & 0.1158 \\ 0.0446 & 0.1136 & 0.2894 & 0.2630 & 0.2894 \\ 0.2448 & 0.2448 & 0.2448 & 0.2448 & 0.0208 \end{bmatrix}$$

C.2 Counterfactual estimates

Appendix C - Table 1: Parameter estimates: Constant r

	preferences				
	σ	μ	A	β	T
	[2]	0.5827	0.0012	[0.97]	[36]
		(0.2204)	(0.5436)		
	rate of return process				
$\mathbb{E}(r)$	2.89%				
	(0.95%)				

Notes: Standard errors in (); fixed parameters in [].

Appendix C - Table 2: Parameter estimates: Constant w

	preferences				
	σ	μ	A	β	T
	[2]	0.5300	0.0055	[0.97]	[36]
		(0.0140)	(0.0011)		
	rate of return process				
state space	0.0083	0.0146	0.0240	0.0489	0.0740
	(0.0008)	(0.0011)	(0.0002)	(0.0021)	(0.0190)
transition diagonal	0.0943	0.0062	0.2249	0.4761	0.0981
	(0.2967)	(0.0225)	(1.0593)	(0.7110)	(0.2833)
statistics	$\mathbb{E}(r)$	$\sigma(r)$	$\rho(r)$		
	3.13%	2.34%	0.160		
	(1.65%)	(1.48%)	(0.008)		

Notes: Standard errors in (); fixed parameters in [].

Appendix C - Table 3: Parameter estimates: $\mu = 2$

	preferences				
	σ	μ	A	β	T
	[2]	2	0.0360	[0.97]	[36]
		-	(0.0779)		
	rate of return process				
state space	0.0033	0.0127	0.0205	0.0531	0.0975
	(0.0195)	(0.0081)	(0.0068)	(0.0159)	(0.0187)
transition diagonal	0.0762	0.5291	0.0068	0.0166	0.2912
	(0.0005)	(0.0008)	(0.0015)	(0.0026)	(0.0138)
statistics	$\mathbb{E}(r)$	$\sigma(r)$	$\rho(r)$		
	2.99%	2.97%	0.112		
	(%)	(%)	()		

Notes: Standard errors in (); fixed parameters in [].

C.3 Complete wealth mobility matrices

We report the complete wealth mobility matrix in the baseline:

$$\hat{T}_{36} = \begin{bmatrix} .349 & .216 & .197 & .131 & .108 \\ .175 & .197 & .245 & .233 & .149 \\ .180 & .193 & .201 & .253 & .173 \\ .151 & .207 & .201 & .210 & .231 \\ .150 & .183 & .157 & .171 & .340 \end{bmatrix}$$

The corresponding complete matrices for all the three counterfactual cases are:

1. constant r:

$$\hat{T}_{36,const\ r} = \begin{bmatrix} 0.258 & 0.246 & 0.182 & 0.174 & 0.140 \\ 0.224 & 0.265 & 0.190 & 0.178 & 0.143 \\ 0.196 & 0.233 & 0.271 & 0.171 & 0.129 \\ 0.175 & 0.166 & 0.248 & 0.244 & 0.167 \\ 0.153 & 0.101 & 0.106 & 0.222 & 0.418 \end{bmatrix}$$

2. constant w:

$$\hat{T}_{36,const\ w} = \begin{bmatrix} 0.564 & 0.403 & 0.022 & 0.004 & 0.006 \\ 0.040 & 0.579 & 0.380 & 0.002 & 0.000 \\ 0.002 & 0.002 & 0.489 & 0.381 & 0.126 \\ 0.113 & 0.006 & 0.022 & 0.430 & 0.429 \\ 0.265 & 0.015 & 0.099 & 0.183 & 0.438 \end{bmatrix}$$

3. $\mu = 2$:

$$\hat{T}_{36,\mu=2} = \begin{bmatrix} 0.258 & 0.214 & 0.200 & 0.178 & 0.151 \\ 0.205 & 0.271 & 0.203 & 0.173 & 0.148 \\ 0.193 & 0.220 & 0.242 & 0.201 & 0.143 \\ 0.173 & 0.181 & 0.199 & 0.250 & 0.198 \\ 0.171 & 0.117 & 0.153 & 0.200 & 0.360 \end{bmatrix}$$

The complete wealth mobility matrix in the estimate with r dependent on wealth is:

$$\hat{T}_{36,r(a)} = \begin{bmatrix} 0.267 & 0.227 & 0.186 & 0.165 & 0.155 \\ 0.222 & 0.221 & 0.225 & 0.172 & 0.160 \\ 0.203 & 0.201 & 0.236 & 0.205 & 0.155 \\ 0.168 & 0.178 & 0.189 & 0.231 & 0.234 \\ 0.143 & 0.172 & 0.163 & 0.226 & 0.296 \end{bmatrix}$$

The complete mobility matrix in the alternative social mobility exercise is

$$\hat{T}_{36} = \begin{bmatrix} 0.228 & 0.216 & 0.170 & 0.201 & 0.101 & 0.042 & 0.038 & 0.005 \\ 0.225 & 0.207 & 0.201 & 0.178 & 0.101 & 0.044 & 0.039 & 0.005 \\ 0.206 & 0.203 & 0.200 & 0.192 & 0.107 & 0.042 & 0.040 & 0.009 \\ 0.193 & 0.203 & 0.212 & 0.193 & 0.111 & 0.041 & 0.042 & 0.006 \\ 0.188 & 0.185 & 0.228 & 0.199 & 0.102 & 0.048 & 0.042 & 0.008 \\ 0.171 & 0.175 & 0.223 & 0.207 & 0.127 & 0.048 & 0.043 & 0.005 \\ 0.164 & 0.140 & 0.221 & 0.210 & 0.130 & 0.072 & 0.047 & 0.015 \\ 0.151 & 0.130 & 0.245 & 0.187 & 0.158 & 0.065 & 0.029 & 0.036 \end{bmatrix}$$

Finally, the complete wealth mobility matrix in the estimate allowing for non-stationary transitional dynamics is:

$$\hat{T}_{36,ns} = \begin{bmatrix} 0.334 & 0.167 & 0.167 & 0.167 & 0.167 \\ 0.327 & 0.171 & 0.168 & 0.167 & 0.167 \\ 0.315 & 0.174 & 0.171 & 0.169 & 0.172 \\ 0.265 & 0.207 & 0.176 & 0.170 & 0.181 \\ 0.190 & 0.173 & 0.180 & 0.181 & 0.276 \end{bmatrix}$$

C.4 Efficient Method of Simulated Moments Estimate

The following describes the procedure we used to produce an optimal weighting matrix for the second step estimation of the two-step Method of Simulated Moments (MSM).

Optimal weighting matrix. We follow Gourieroux, Monfort, and Renault (1993) and calculate the variance-covariance matrix of the data moments by bootstrapping, respectively for the wealth distribution moments and the intergenerational wealth mobility moments. Note that in order to invert the variance-covariance matrix, we use seven wealth moments (dropping the first one) to avoid perfect collinearity. Denote the variance-covariance matrix of the wealth distribution moments as $\mathcal{V}_{T_1}$, and that of the wealth mobility moments as $\mathcal{V}_{T_2}$, where T_1 and T_2 are the number of observations in each of the two samples.[4] We assume that there is no correlation in the error structure between the two samples. The optimal weighting matrix W_{T_1,T_2} would be the inverse of the concatenated block-diagonal variance-covariance matrix, that is,

$$W_{T_1,T_2} = \begin{bmatrix} \mathcal{V}_{T_1} & \mathbf{0} \\ \mathbf{0} & \mathcal{V}_{T_1} \end{bmatrix}^{-1}$$

We bootstrap 10,000 times for each of the variance-covariance matrix, and for each bootstrap we use half of the original sample to calculate the bootstrapped sample moments. As the wealth distribution moments are much more precisely estimated than the mobility moments, the weights on the former are around 3 orders of magnitude higher than the latter.

MSM results. In the first step, we use the same matrix we use in the baseline as the weighting matrix. We denote the first-step estimate as $\hat{\theta}_1$. Using $\hat{\theta}_1$ as the initial guess, we repeat the estimation procedure with the new optimal weighting matrix calculated earlier, $\widehat{W_{T_1,T_2}}$. Denote the second-step estimate as $\hat{\theta}_2$.

Appendix C - Table 4: Model fit: MSM

	wealth distribution							
percentile	0-20	20-40	40-60	60-80	80-90	90-95	95-99	99-100
wealth share (data)	-0.002	0.001	0.045	0.112	0.120	0.111	0.267	0.336
wealth share (model)								
(1) baseline	0.049	0.077	0.111	0.110	0.110	0.076	0.142	0.325
(2) optimal weighting W	0.047	0.076	0.108	0.103	0.106	0.073	0.140	0.346

	social mobility				
percentile	0-20	20-40	40-60	60-80	80-100
transition diagonal (data)	0.36	0.24	0.25	0.26	0.36
transition diaginal (model)					
(1) baseline	0.349	0.197	0.201	0.210	0.340
(2) optimal weighting W	0.287	0.242	0.261	0.277	0.380

[4] Note that we use two different data samples for calculating wealth distribution and mobility moments. The former comes from the SCF 2007 cross-sectional sample, while the latter comes from the SCF 2007-2009 panel subsample.

Appendix C - Table 5: Parameter estimates: MSM

	preferences				
	σ	μ	A	β	T
	[2]	0.5993	0.0006	[0.97]	[36]
		(0.3854)	(1.8317)		
	rate of return process				
state space	0.0010	0.0094	0.0257	0.0574	0.0841
	(0.0042)	(0.0044)	(0.0148)	(0.0518)	(0.1633)
transition diagnonal	0.0507	0.3067	0.1379	0.2200	0.0215
	(0.1196)	(0.2597)	(0.3167)	(0.0424)	(0.0141)
statistics	$\mathbb{E}(r)$	$\sigma(r)$	$\rho(r)$		
	3.01%	2.72%	0.198		
	(0.02%)	(0.01%)	(0.253)		

Notes: Standard errors in (); fixed parameters in [].

References

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States." *The Quarterly Journal of Economics*, 129(4): 1553–1623.

Guvenen, Fatih. 2016. *Quantitative Economics with Heterogeneity: An A-to-Z Guidebook.* Princeton University Press.

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante. 2010. "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006." *Review of Economic Dynamics*, 13(1): 15–51.

Kennickell, Arthur B, and Martha Starr-McCluer. 1997. "Household Saving and Portfolio Change: Evidence from the 1983-89 SCF Panel." *Review of Income and Wealth*, 43(4): 381–99.

Klevmarken, Anders, Joseph P. Lupton, and Frank P. Stafford. 2003. "Wealth Dynamics in the 1980s and 1990s: Sweden and the United States." *Journal of Human Resources*, 38(2).

Miranda, Mario J., and Paul L. Fackler. 2004. *Applied Computational Economics and Finance.* Vol. 1 of *MIT Press Books*, The MIT Press.