constant, but a function of state variables. Specifically, the estimate for λ is given by

$$\begin{aligned}\widehat{\lambda} &= \widehat{\lambda}(a_t, h_t^s, h_t^w, age_t, educ, type; \widehat{\zeta}) \\ &\equiv 1 - \frac{\widehat{\underline{a}}_t}{a_t} \\ &= 1 - \frac{\underline{a}(h_t^s, h_t^w, age_t, educ, type; \widehat{\zeta})}{a_t}\end{aligned}$$

By incorporating the parameter estimates into the above equation, we have

$$\begin{aligned}&\widehat{\lambda}(a_t, h_t^s, h_t^w, age_t, educ, type; \widehat{\zeta}) \\ &= 1 + \frac{1}{a_t} \cdot (\exp[9.256 + 0.201 \cdot I(educ = college) + 0.355 \cdot I(h_t^s = 1) \\ &+0.063 \cdot h_t^w - 0.357 \cdot \frac{(h_t^w)^2}{100} - 0.028 \cdot (age_t - 20) + 0.207 \cdot \frac{(age_t - 20)^2}{100}]\Big) .\end{aligned}$$

As a example, this value becomes $\widehat{\lambda} = 1.62$ if I plug into the function $h_t^w = 6$ (the median for years of paid employed work; see Table 3), $age_t = 29$ (the median age) and $a_t = 20008$ (the median net worth) for the non-college educated who has never become a self-employer.

7.2.5 Type and the Propensity to Become a Self-Employer

It is interesting to notice that Type 1 individuals (comprising 28.0% of the simulated sample) and Type 2 individuals (comprising 72.0% of the simulated sample) are not much different in terms of risk attitude: the difference between the estimated values of Type 1's CRRA constant and of Type 2's is very small. Hence, as opposed to the

intuition behind the modeling by Kihlstrom and Laffont (1979) and by Kanbur (1979), heterogeneity in risk attitude may play a limited role in explaining the observed characteristics of the actual data.[60] The positive values of the estimated γ_2^w and γ_2^w suggest that Type 1 individuals are less productive both as a self-employer and as a wage worker. Type 2 individuals, who are more productive both in self-employment and in wage employment, are more likely to become a self-employer. Looking at Table 12 that shows the transition rates for Type 1 and for Type 2, we find that the transition rates from and to self-employment and from and to full-time paid employment are higher for Type 2 individuals. This may suggest that the decision to transit into self-employment may not be motivated by innate comparative advantage of one mode over another. The persistence of part-time paid employment, however, is higher while there is little difference as to the persistence of non-employment.

8 Counterfactual Experiments

One of the attractive features of using a dynamic structural model in empirical studies is that they can be used to predict the effects of (*ex ante*) counterfactual changes in exogenous variables. In this section, I conduct three counterfactual experiments to see their effects on the formation and continuation of self-employment. For each experiment, I first simulate behavior under the appropriately defined scenario, and then compare counterfactual behavior with the baseline behavior.

[60] Relatedly, Blanchflower and Oswald (1998, p.30) state that the classical writings such as Knight (1921), Schumpeter (1939) and Kirzner (1973) emphasized that "attitude to risk is not the central characteristic that determines who becomes an entrepreneur."

Table 12: Transition Matrices for Labor Supply Decisions by Type

Predicted Overall

Labor supply (t-1)		Labor supply (t) SE	PE Full-time	PE Part-time	NE
SE					
	Row %	84.6	9.6	0.6	5.2
	Column %	91.7	5.7	0.8	8.1
PE, Full-time					
	Row %	4.2	85.6	6.6	3.5
	Column %	4.6	50.8	8.8	5.5
PE, Part-time					
	Row %	2.4	52.2	44.9	0.5
	Column %	2.6	31.0	59.4	0.7
NE					
	Row %	1.0	21.0	23.4	54.5
	Column %	1.1	12.5	31.0	85.6

Type 1

Labor supply (t-1)		Labor supply (t) SE	PE Full-time	PE Part-time	NE
SE					
	Row %	78.2	12.3	1.8	7.8
	Column %	90.7	8.6	1.8	10.7
PE, Full-time					
	Row %	4.3	72.7	11.4	11.7
	Column %	5.0	51.2	11.5	16.0
PE, Part-time					
	Row %	2.8	42.6	54.1	0.5
	Column %	3.3	30.0	54.6	0.6
NE					
	Row %	0.9	14.5	31.9	52.7
	Column %	1.0	10.2	32.2	72.6

Type 2

Labor supply (t-1)		Labor supply (t) SE	PE Full-time	PE Part-time	NE
SE					
	Row %	86.7	8.8	0.2	4.3
	Column %	92.4	4.7	0.3	7.1
PE, Full-time					
	Row %	4.2	90.2	5.0	0.7
	Column %	4.4	48.8	8.2	1.1
PE, Part-time					
	Row %	2.0	62.0	35.5	0.5
	Column %	2.1	33.6	58.5	0.8
NE					
	Row %	1.0	23.7	20.0	55.3
	Column %	1.1	12.8	33.0	91.0

Table 13: Changes in the Characteristics on Entry into and Exit from Self-Employment (Experiments 1-3)

	Baseline		Exp. 1		Exp. 2		Exp. 3	
	NC	C	NC	C	NC	C	NC	C
Ever experience of self-employment (%)	26.76	24.43	23.77	21.76	34.40	31.54	51.51	47.82
First entry into self-employment occurs in less than or equal to first eight decision years (%)	59.34	69.20	65.97	64.55	69.31	67.29	84.49	87.86
Exit from self-employment in a year (%)	28.09	23.61	19.60	21.51	45.98	38.81	42.22	32.25

Note: "NC" ("C") denotes "non-college educated"("college educated").

8.1 Experiment 1: Relaxation of the Borrowing Constraints

The first experiment is to relax the estimated borrowing constraints to evaluate the extent to which they affect self-employment. Specifically, I make the asset floor $\underline{a}_t$ negatively very large to see the effects (1) on the percentage of those with experience of self-employment during the same periods that are covered by the NLSY79, (2) on the distribution of decision period when first entry into self-employment occurs, and (3) on the one-period transition rates. I consider the case of $\underline{a}_t = -\$30,000$ for any state (Experiment 1).[61]

Column "Exp.1" in Table 13 displays how the change affects the three key statistics in Table 1. Interestingly, relaxing the borrowing constraints weakens the propensity to enter into self-employment both for the non-college educated and for the college educated. This result presumably comes from the fact that there are two effects of relaxing the borrowing constraints: one is the direct effect that improves consumption smoothing over the life cycle, and the other is the indirect effect that makes individuals more likely to be non-employed. If we look at the age profiles of labor supply under the Baseline and under Experiment 1 in Table 13, we find that the rates of non-employment under Experiment 1 are likely to be higher before age 30 but are likely to be lower after age 30. Table 14 also compares the predicted age profile of self-employment with the one under the Baseline and under Experiment 1. The proportions of self-employers are now higher from the average 7 percent in the simulated sample to around 9 to 10 percent for the two experiments. The rates of increase are 30 to 40 percent, which is comparable to Evans and Jovanovic's (1989) results.

Turning back to Table 14, it is noteworthy that college-educated self-employers are now likely to enter into self-employment later than under the Baseline while non-college educated self-employers are likely to enter into self-employment earlier. The exit rates after the first year from self-employment improves more for the non-college educated while the effects on the college educated is weaker. These results together seem to suggest the differences of the effects of borrowing constraints by the level of schooling. Relaxation of borrowing constraints facilitates both the formation of and the continuation of self-employed businesses especially for the non-college educated. Lastly, Figures 16-19 show that there are no significant differences between incomes under the Baseline and those under Experiment 1.

[61] The estimated lower bound is between \$10,000 and \$18,000 for most of state variables.

Table 14: Changes in the Age Profiles of Labor Supply (Experiment 1)

	Self-Employment		Paid Employment		Non-Employment	
Age	**Baseline**	**Exp. 1**	**Baseline**	**Exp. 1**	**Baseline**	**Exp. 1**
20	0.9	1.1	78.6	78.9	20.7	20.0
21	3.5	2.2	79.4	78.4	16.6	19.4
22	5.7	5.8	79.8	72.4	11.1	21.8
23	7.4	6.2	82.1	76.6	9.4	17.2
24	7.8	7.3	89.3	81.4	9.1	11.4
25	7.2	6.9	88.4	83.6	7.2	9.5
26	6.5	6.4	90.1	84.8	7.7	8.7
27	6.7	6.5	89.3	84.2	8.3	9.3
28	8.0	8.2	84.6	81.3	9.9	10.5
29	7.7	8.0	78.4	81.6	9.8	10.4
30	7.3	7.8	77.0	81.8	9.9	10.4
31	7.1	7.8	75.4	81.5	10.7	10.7
32	7.2	7.9	76.9	83.7	10.5	8.4
33	7.5	8.2	76.9	84.5	10.8	7.4
34	7.8	8.8	77.3	83.9	11.7	7.4
35	7.2	9.5	76.5	84.3	12.9	6.2
36	7.3	10.6	79.9	83.6	12.4	5.8
37	7.3	11.5	84.2	83.4	13.3	5.1
38	8.0	11.9	84.3	83.0	14.6	5.1
39	8.4	12.6	85.3	82.2	15.3	5.2

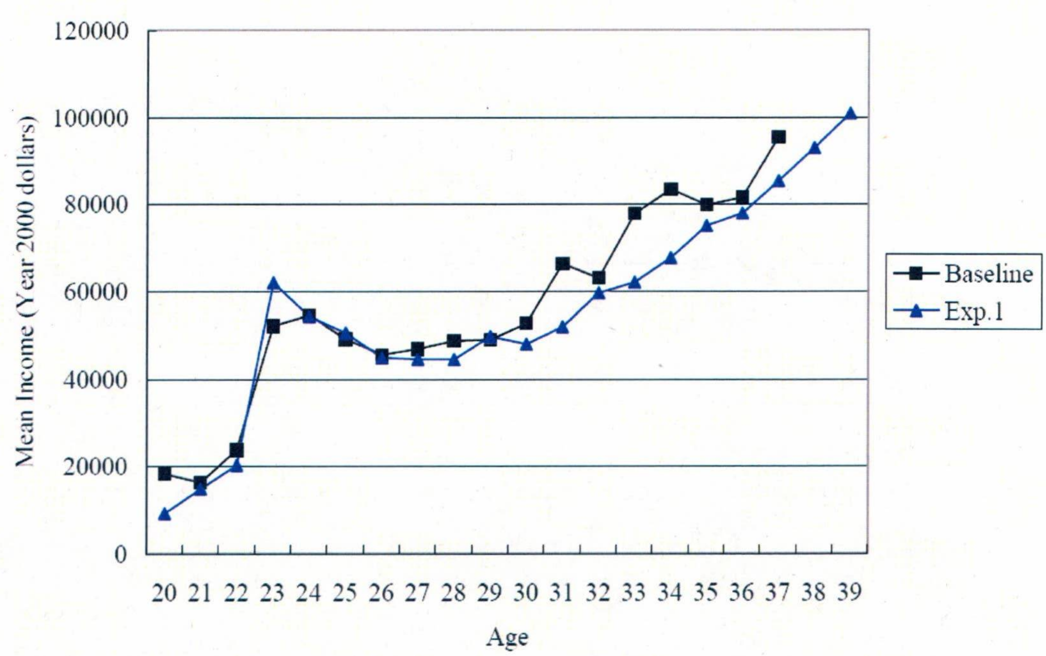

Figure 16: Changes in the Age Profiles of the Mean Income from Self-Employment (Non-college; Experiment 1)

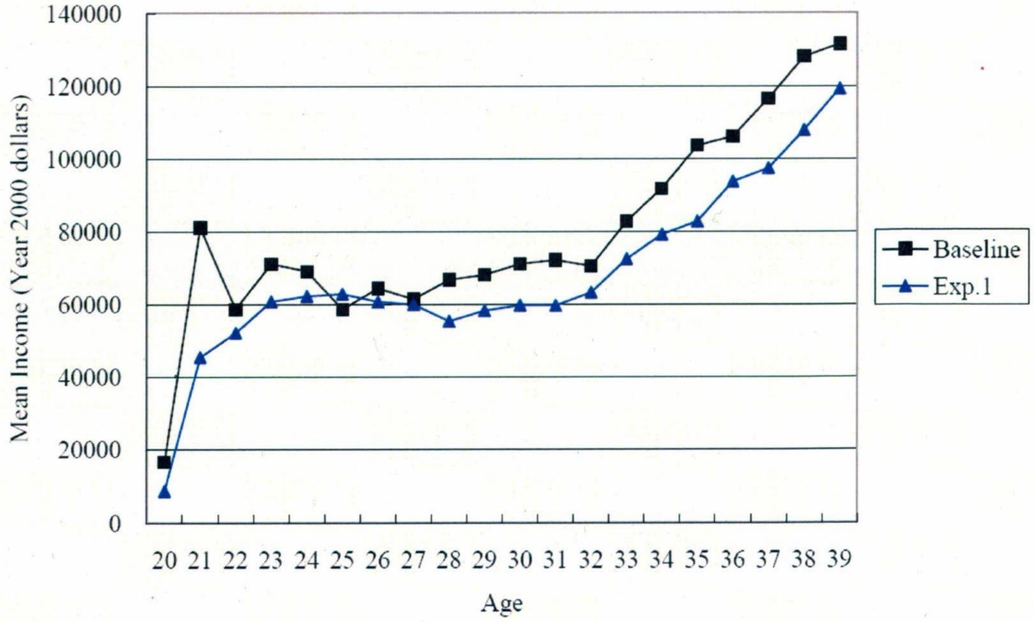

Figure 17: Changes in the Age Profiles of the Mean Income from Self-Employment (College; Experiment 1)

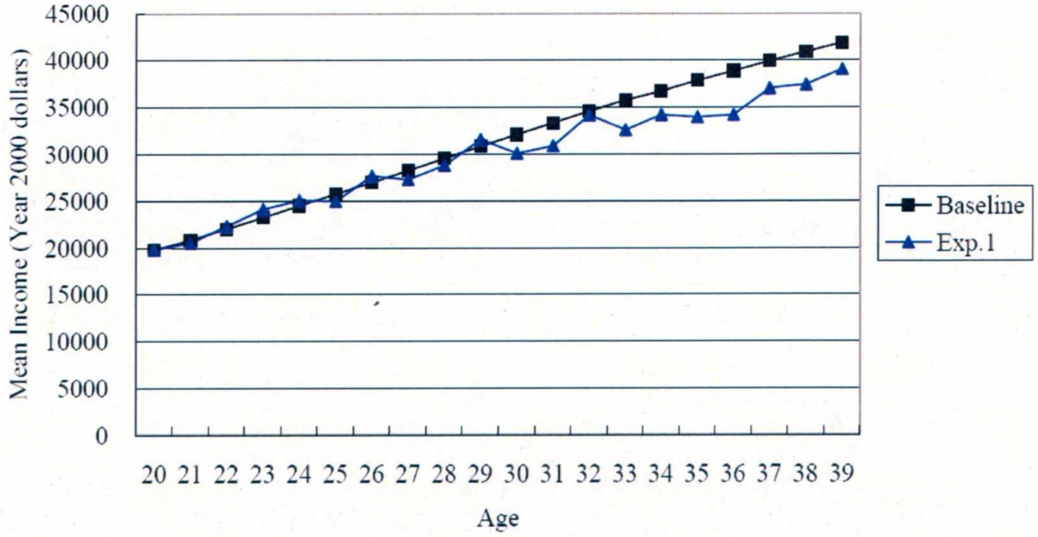

Figure 18: Changes in the Age Profiles of the Mean Income from Full-time Paid-Employment (Non-college; Experiment 1)

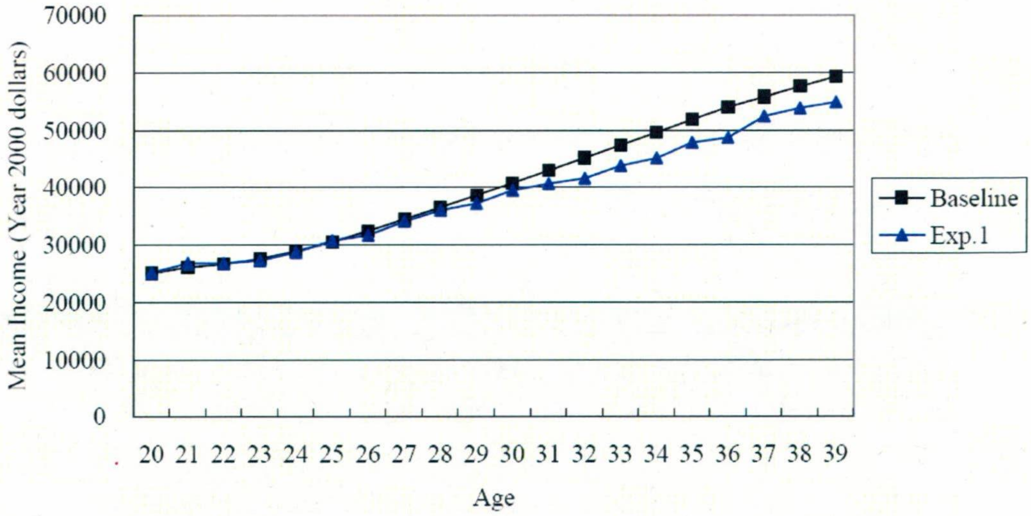

Figure 19: Changes in the Age Profiles of the Mean Income from Full-time Paid-Employment (College; Experiment 1)

8.2 Experiment 2: Injection to Business Capital

In this experiment, self-employers are subsidized through capital enhancement. Specifically, some amount of subsidy is given every year as long as individuals continue to be self-employed, and thus the amount of capital for self-employed is $(k_t^{\$} + subsidy)$ rather than $k_t^{\$}$ under the Baseline. I consider the case of $subsidy = \$30,000$.[62]

Column "Exp.2" in Table 13 displays how the subsidy scheme affects the three key statistics in Table 1. In comparison to the Baseline and Experiment 1, the ratios of those who ever experience self-employment for both schooling levels increase. While the college educated are now likely to enter into self-employment slightly later in their life, the non-college educated start self-employment earlier under the this subsidy scheme. The third item, however, shows that the subsidy scheme actually increases entries with shorter duration. In particular, nearly 50 percent of non-college educated self-employers exit in the first year.

Figures 20-23 show the changes in the age profiles of annual incomes of self-employment and of fullt-time paid employment. We find that increases of self-employment income are larger for the non-college educated. It is interesting to see that the mean incomes from full-time paid employment are now greater both for the non-college educated and for the college educated. This is presumably due to the decreases in the ratios of wage workers in the age profile of labor supply (Table 15).

8.3 Experiment 3: Injection to Self-Employment-Specific Human Capital

Most existing programs aims to assist unemployed workers to become self-employers through training. The next experiment is the one that gives indirect incentives for self-employment

[62] This and next experiments should be taken as thought experiments rather than as real policy experiments because the main program conducted today by the U.S. Small Business Administration (SBA) to support small businesses is by loan guarantees, and the SBA does not provide direct loans to start or expand a business.

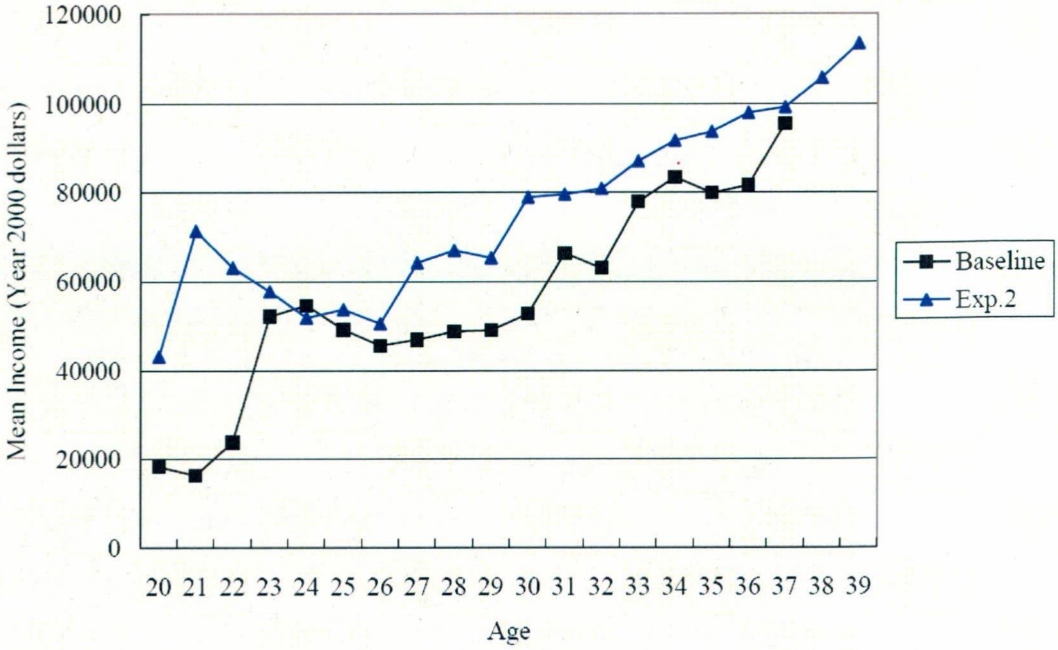

Figure 20: Changes in the Age Profiles of the Mean Income from Self-Employment (Non-college; Experiment 2)

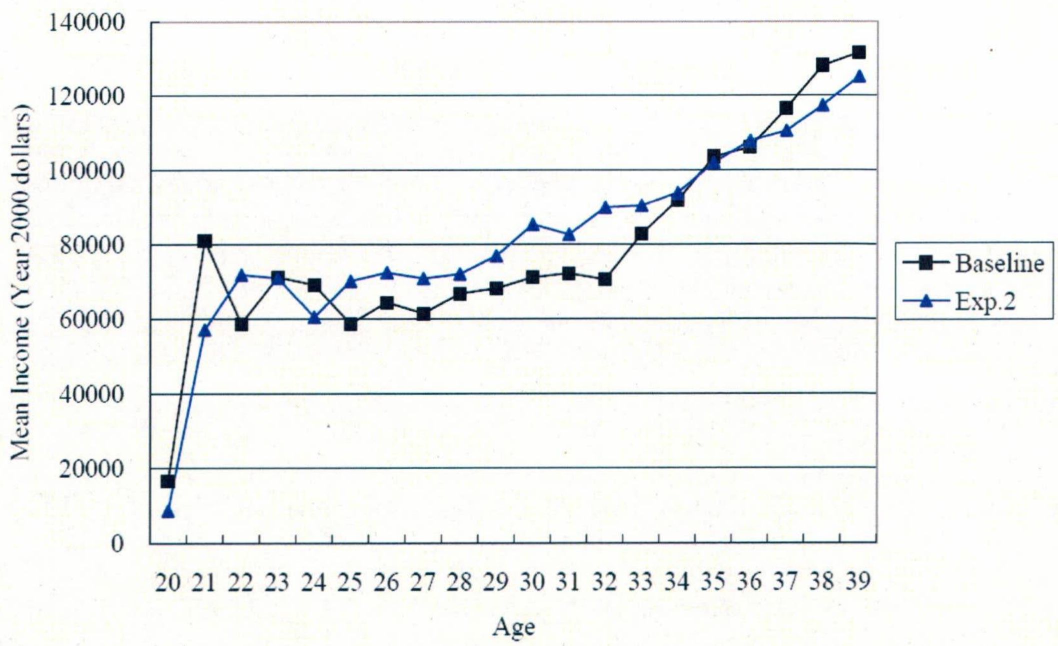

Figure 21: Changes in the Age Profiles of the Mean Income from Self-Employment (College; Experiment 2)

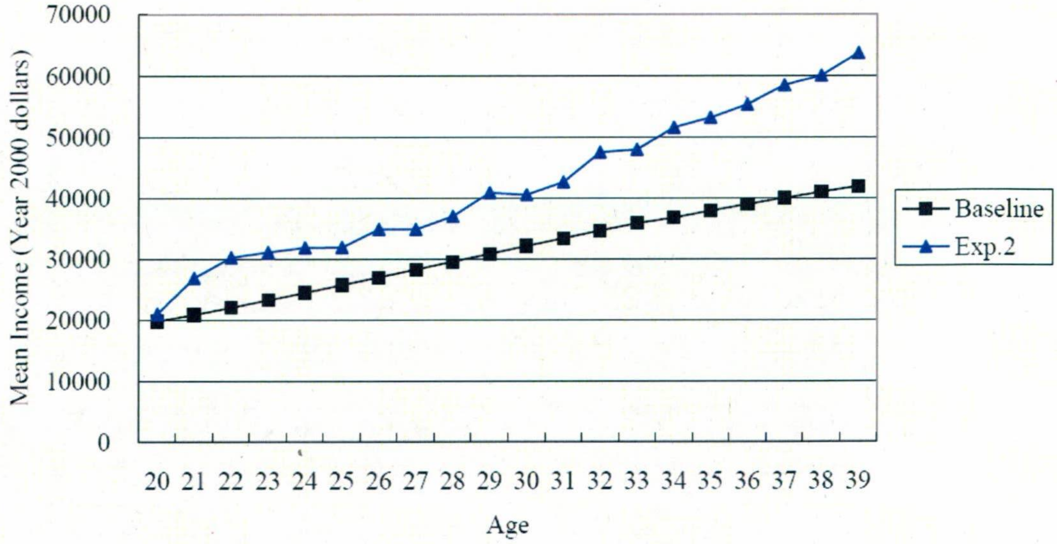

Figure 22: Changes in the Age Profiles of the Mean Income from Full-time Paid-Employment (Non-college; Experiment 2)

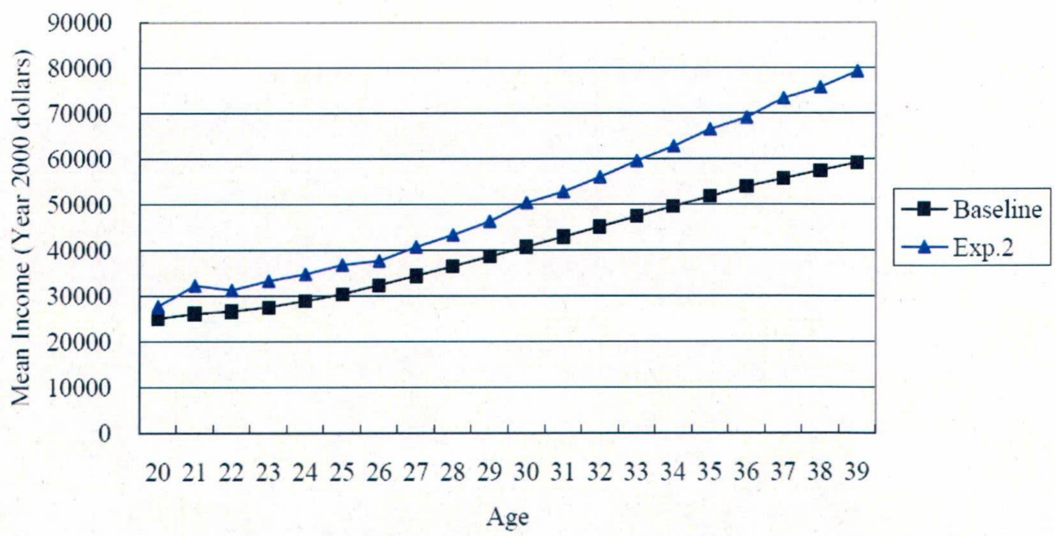

Figure 23: Changes in the Age Profiles of the Mean Income from Full-time Paid-Employment (College; Experiment 2)

Table 15: Changes in the Age Profiles of Labor Supply (Experiment 2)

	Self-Employment		Paid Employment		Non-Employment	
Age	**Baseline**	**Exp. 2**	**Baseline**	**Exp. 2**	**Baseline**	**Exp. 2**
20	0.9	1.3	78.6	78.8	20.7	19.9
21	3.5	4.4	79.4	72.9	16.6	22.7
22	5.7	6.4	79.8	79.5	11.1	14.1
23	7.4	8.3	82.1	77.3	9.4	14.4
24	7.8	9.0	89.3	77.7	9.1	13.3
25	7.2	9.9	88.4	76.0	7.2	14.1
26	6.5	8.9	90.1	80.6	7.7	10.5
27	6.7	8.2	89.3	80.6	8.3	11.2
28	8.0	8.0	84.6	78.7	9.9	13.3
29	7.7	7.6	78.4	77.8	9.8	14.6
30	7.3	6.7	77.0	78.5	9.9	14.8
31	7.1	6.9	75.4	80.7	10.7	12.5
32	7.2	6.4	76.9	80.3	10.5	13.3
33	7.5	6.2	76.9	78.2	10.8	15.6
34	7.8	6.7	77.3	75.3	11.7	18.0
35	7.2	7.0	76.5	71.0	12.9	22.0
36	7.3	7.6	79.9	71.1	12.4	21.3
37	7.3	8.1	84.2	70.0	13.3	21.9
38	8.0	8.7	84.3	70.6	14.6	20.8
39	8.4	9.5	85.3	62.4	15.3	28.1

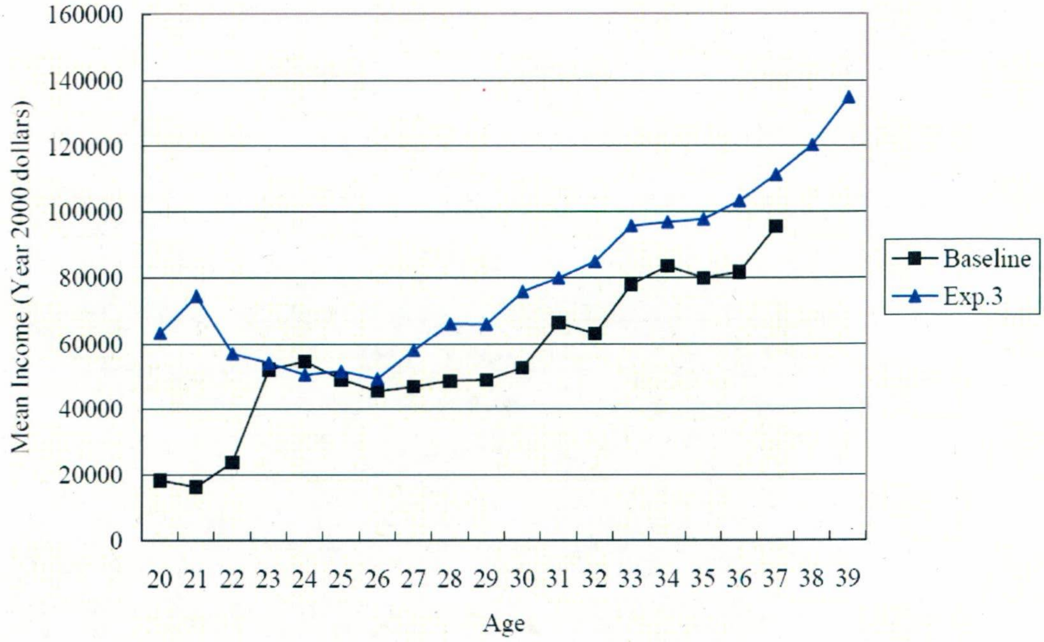

Figure 24: Changes in the Age Profiles of the Mean Income from Self-Employment (Non-college; Experiment 3)

before entering the labor force. Specifically, it sees the effects of enhanced entrepreneurial skill (by, for example, training offered to improve skills that are useful for self-employed work). Conceptually, it corresponds to changing $\widehat{\gamma}_0^s = 1.839$ to a higher value. I set it equal to the constant part for the wage-sector specific human capital, $\widehat{\gamma}_0^w = 2.322$ (Experiment 3).

Column "Exp.3" in Table 13 displays how the counterfactual stuation affects the three key statistics in Table 1. The directions in the changes are similar to those by Experiment 2: it induces more individuals to enter into self-employment earlier in their life-cycle, but they are likely to exit from self-employment sooner. Figures 24-27 show that the directions in the changes in the mean incomes are also similar with Experiment 2. The changes in the age profile of labor supply are also similar (not shown) with slightly higher ratios of the self-employed.

9 Concluding Remarks

In this paper, I have proposed and estimated a life-cycle model of entrepreneurial choice and wealth accumulation. The data for estimation are taken from the cohort of young white males from the NLSY79. Inclusion of the nonpecuniary benefits of self-employment in the model is important to accurately replicate the age patterns of labor supply. The estimated model is used to conduct counterfactual experiments. My results suggest that both a reasonable subsidy and an enhancement of human capital specific to self-employment would have a small impact on self-employment in the labor supply cohort up to age 39 in terms of either the age profile of self-employment or whether individuals have ever experienced self-employment. In contrast, a moderate relaxation of the borrowing constraints encourages entries into self-employment, and the average duration of self-employment is longer. Although the counterfactual experiments in this study are conducted with the use of US

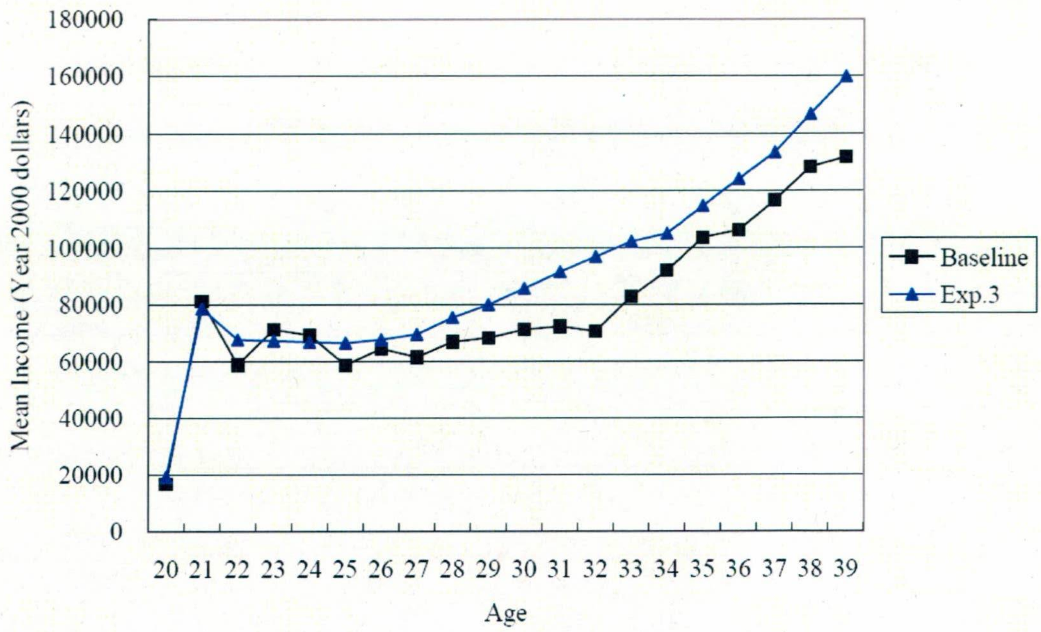

Figure 25: Changes in the Age Profiles of the Mean Income from Self-Employment (College; Experiment 3)

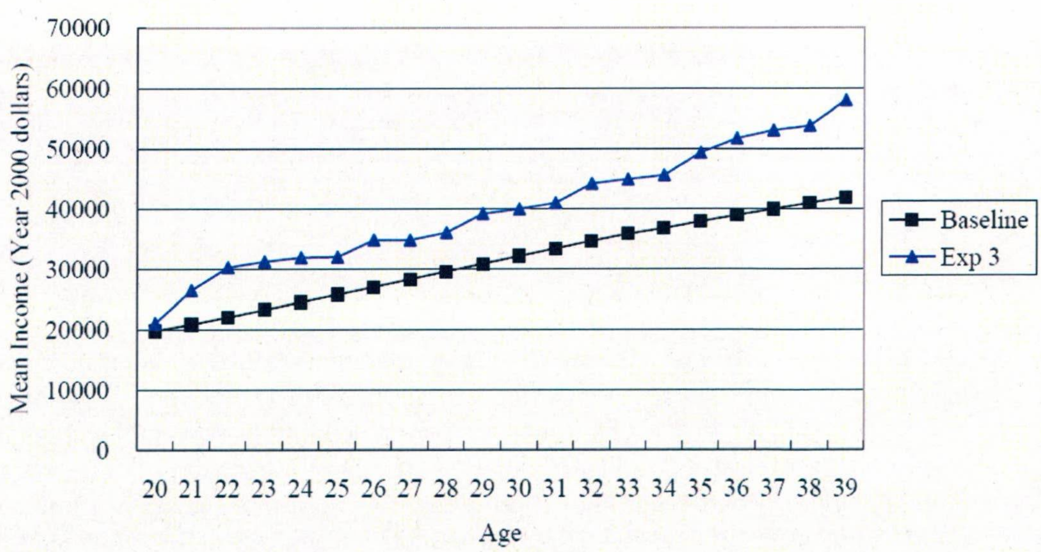

Figure 26: Changes in the Age Profiles of the Mean Income from Full-time Paid-Employment (Non-college; Experiment 3)

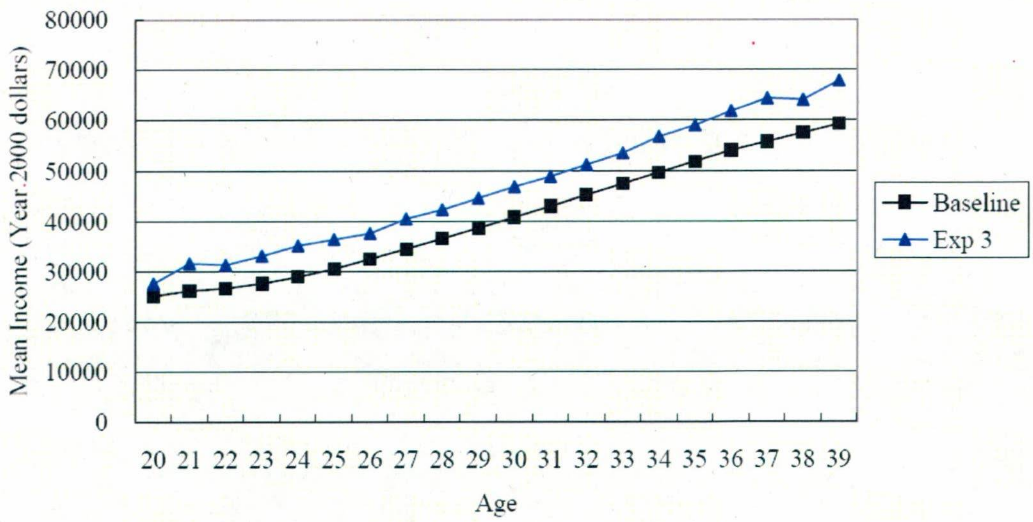

Figure 27: Changes in the Age Profiles of the Mean Income from Full-time Paid-Employment (College; Experiment 3)

data, insights from the experiments would be useful in considering what determines the dynamic characteristics of self-employment. However, as with any empirical work, the present study should not be taken as definitive, and further research remains to be undertaken, in particular, to determine how different setups and specifications would alter the conclusions.

The sample I use for this study consists of white males. To members of minority groups who are seeking self-employment opportunities, public assistance programs for self-employment would be relevant, as such individuals may face more severe borrowing constraints than white males. Self-employment assistance programs sometimes target specific groups; for example, structurally unemployed or displaced workers and social assistance recipients. The framework in the present study can be extended to study nonwhite self-employment.

The present study has considered the model of a single agent. It would be interesting and important to study various self-employment issues by extending Blau's (1987) two-sector (entrepreneurial and corporate) general equilibrium model to a dynamic model.[63] A more fundamental question is what entrepreneurial skills are.[64] Family background would also play an important role.[65] These and other interesting questions about self-employment are

[63] Blau's (1987) time series regression analysis found that six effects could explain most of the increase in the proportion of male workers that were self-employed from 1973 to 1982. In order of importance, the six effects were: an increase in a self-employment factor productivity index; a decrease in the marginal tax rate at a real income of $7000; an increase in the social security benefit level; a decrease in the real minimum wage; and a decrease in the age of the male labor force.

[64] Using data from a list of MBA alumni at a business school, Lazear (2005) found that, on average, those alumni with experience of self-employment (in an incorporated business) took courses from a broader area of specialties. The difference between the number of courses taken in the student's specialty and the average number of courses taken in other fields. A study by White, Thornhill and Hampson (2006) used data collected from 31 MBA students with significant prior involvement in new venture creation and from 79 other students with no new venture startup experience, and found that a *testosterone* (an endocrine hormone) effect upon behavior (new venture creation) is partially mediated by the psychological (risk propensity).

[65] For example, Hundley (2006) found evidence that whereas both parental self-employment and family income significantly increased a man's propensity to become a self-employer, the former effect is amplified by the latter.

left for future research.

Appendix A Exact Functional Forms Used to Simulate the Model

In this appendix, I show the exact forms of functions as they are used in *simulating* the life-cycle model. Essentially, when trying to match the model predictions to the empirical patterns in many dimensions, I am concerned about disentangling as precisely as possible unobserved heterogeneity from such effects as the life-cycle and the hysteresis effects on labor choice and saving decisions. Age variables and lagged choice variables in the utility function and the human capital functions are thus added to prevent the overstatement of the unobserved heterogeneity.

To *estimate* the model (by the simulated unconditional maximum likelihood method), I need parameterization for classification and measurement to connect simulated data and the observed data. I also need to determine the rule on how to fill in missing initial asset observations. The exact specification for these parts is given in Appendix C. The total number of the parameters in the current specification that are necessary for simulation (given an initial condition) is 64. In what follows, $I(\cdot)$ is an indicator function that assigns one if the term inside the parenthesis is true and zero otherwise.

A.1 Preference: Time Discount Factor, β, and the Utility Function, $u(\cdot;\epsilon_t^{ls},\epsilon_t^{lw};\mu)$ (16 Parameters)

I assume that β is a common constant for all individuals. The utility function is given as the following CRRA form augmented by τ_t^s, l_{t-1}^w and age_t:

$$
\begin{aligned}
u_t &= u(c_t, l_t^s, l_t^w, \tau_t^s, l_{t-1}^w, age_t,\ type; \epsilon_t^{ls}, \epsilon_t^{lw}; \mu) \\
&= \frac{c_t^{1-\mu_0(type)}}{1-\mu_0(type)} \\
&\quad - \underbrace{[\mu_{1,s}(type,\tau_t^s) + \epsilon_t^{ls}] \cdot I(l_t^s > 0)}_{\text{disutility from self-employment}} \\
&\quad - \underbrace{[\mu_{1,w}(type) + \epsilon_t^{lw}] \cdot [I(l_t^w \text{ is } \textit{full-time}) + \mu_{1,w,part} \cdot I(l_t^w \text{ is } \textit{part-time})]}_{\text{disutility from paid-employment}} \\
&\quad - \underbrace{\mu_{2,full}(type) \cdot I(l_t^s > 0\ \&\ l_t^w \text{ is } \textit{full-time}) - \mu_{2,part}(type) \cdot I(l_t^s > 0\ \&\ l_t^w \text{ is } \textit{part-time})}_{\text{disutility from dual job holding}} \\
&\quad + \underbrace{\mu_{3,s\to s}(type,\tau_t^s) \cdot I(l_t^s > 0\ \&\ \tau_t^s \geq 1)}_{\text{utility benefit from staying in SE}},
\end{aligned}
$$

where

$$\mu_0(type) = \mu_{00} + \mu_{01} \cdot I(type = 2)$$

is the parameter for relative risk aversion parameter ("prudence"),

$$
\begin{cases}
\mu_{1,s}(type,\tau_t^s) = \mu_{10,s} + \mu_{11,s} \cdot I(type=2) + \mu_{12,s} \cdot \tau_t^s \\
\mu_{1,w}(type) = \mu_{10,w} + \mu_{11,w} \cdot I(type=2)
\end{cases}
$$

are parameters for labor disutility and

$$\begin{cases} \mu_{2,full}(type) = \mu_{20,full} + \mu_{21,full} \cdot I(type = 2) \\ \mu_{2,part}(type) = \mu_{20,part} + \mu_{21,part} \cdot I(type = 2) \end{cases}$$

are additional disutility if he works as a self-employer and as a wage worker in the same period. Parameter $\mu_{3,\cdot,\cdot}$, interpreted as utility benefits from staying in self-employment when he has been a for τ_t^s years is

$$\mu_{3,s\to s}(type, \tau_t^s) = \mu_{40,s\to s} + \mu_{41,s\to s} \cdot I(type = 2) + \mu_{42,s\to s} \cdot \tau_t^s$$

A.2 Constraints: Lower Bound for Financial Net Worth, $\underline{a}(\cdot;\zeta)$, and the Consumption Floor, $c_{\min}$ (8 Parameters)

The borrowing constraint requires that net financial assets not fall below some nonpositive lower bound. I allow the constraint to evolve as a function of the individual's level of education, work experience and (unobserved) type as well as age:

$$\begin{aligned} \underline{a}_t &= \underline{a}(h_t^s, h_t^w, age_t, educ; \zeta) \\ &= -\exp[\zeta_0 + \zeta_1 \cdot I(educ = college) + \varsigma_2 \cdot I(h_t^s = 1) \\ &+ \zeta_3 \cdot h_t^w + \zeta_4 \cdot \frac{(h_t^w)^2}{100} + \zeta_5 \cdot (age_t - 20) + \zeta_6 \cdot \frac{(age_t - 20)^2}{100}] \end{aligned}$$

for $t = 2, ..., T$.[66] The interpretation is that education, work experience and type serve (through human capital) to forecast future earnings potential. The dependence of the lower bound on age is in expectation of getting a better fit. Together with $c_{\min}$, which I assume is a common constraint for all individual, the lower bound gives the upper and lower bounds for c_t and for a_{t+1}:

$$\begin{cases} c_{\min} \le c_t \le y_t + (1-\delta)k_t + (1+r)a_t - \underline{a}_{t+1} \\ \underline{a}_{t+1} \le a_{t+1} \le y_t + (1-\delta)k_t + (1+r)a_t - c_{\min} \end{cases}$$

where c_t and a_{t+1} are related through the budget constraint:

$$c_t + a_{t+1} = y_t + (1-\delta)k_t + (1+r)a_t.$$

[66] Remember that the initial financial net worth a_1 must also satisfy $a_1 \ge \underline{a}_1$; otherwise, the upper bound for k_1, $a_1 - \underline{a}_1$, could be negative. In this study, I simply look at the minimum of financial net worth in the first decision period for each cell in the following table.

age 20	−15840.0 (obs.no. = 138)
ages 21-23	−38308.0 (obs.no. = 177)
ages 24-25	−21864.0 (obs.no. = 72)

Why I do not collapse the table by education is that the numbers of observations with college for age 20 and of those with noncollege for ages 24-25 are very small. I simply assume that the lower bound, $\underline{a}_1$, for each age level is 1.20 × minimum (negative for all of the three case) of observations (−19008.0 for age 20, −45969.6 for ages 21 to 23 and −26236.8 for ages 24 and 25).

A.3 Human Capital, $\overline{\Psi}^m(\cdot;\gamma^m)$, $m = w, s$, and the Rental Price for Part-time Paid-Employment, R^p (17 Parameters)

In the present study, I consider mode-specific human capital, that is, I distinguish human capital that is used for paid-employment ($\overline{\Psi}_t^m$) and that for self-employment ($\overline{\Psi}_t^s$). For paid-employment, the market value of human capital is given by

$$\begin{aligned} R^j \cdot \overline{\Psi}_t^w &= R^j \cdot \overline{\Psi}_t^w(h_t^w, h_t^s, age_t, educ, type; \gamma^w) \\ &= R^j \cdot \exp[\underbrace{\gamma_0^w}_{\text{constant}} + \underbrace{\gamma_1^w \cdot I(educ = college)}_{\text{schooling}} + \underbrace{\gamma_2^w \cdot I(type = 2)}_{\text{type}} \\ &+ \underbrace{\gamma_3^w \cdot h_t^w + \gamma_4^w \cdot \frac{(h_t^w)^2}{100}}_{\text{accumulated ``own'' experience}} + \underbrace{\gamma_5^w \cdot I(h_t^s = 1)}_{\text{ever experience in SE}} \\ &+ \underbrace{\gamma_6^w \cdot (age_t - 20)}_{\text{age effect}}] \end{aligned}$$

where R^j is the *rental price* for j-time paid employment. Since the relative price matters, I normalize $R^f = 1$. For self-employment, the value of human capital is given by

$$\begin{aligned} \overline{\Psi}_t^s &= \overline{\Psi}_t^s(h_t^s, \tau_t^s, h_t^w, age_t, educ, type; \gamma^s) \\ &= \exp[\underbrace{\gamma_0^s}_{\text{constant}} + \underbrace{\gamma_1^s \cdot I(educ = college)}_{\text{schooling}} + \underbrace{\gamma_2^s \cdot I(type = 2)}_{\text{type}} \\ &+ \underbrace{\gamma_3^s \cdot I(h_t^s = 1)}_{\text{ever ``own'' experience}} \quad \underbrace{+\gamma_4^s \cdot \tau_t^s + \gamma_5^s \cdot \frac{(\tau_t^s)^2}{100}}_{\text{accumulated ``own'' experience \textit{in a row}}} + \underbrace{\gamma_6^s \cdot h_t^w + \gamma_7^s \cdot \frac{(h_t^w)^2}{100}}_{\text{accumulated experience in PE}} \\ &+ \underbrace{\gamma_8^s \cdot (age_t - 20)}_{\text{age effect}}]. \end{aligned}$$

As in the standard literature on human capital formation, the productivity of human capital for mode $m = s, w$ in period t depends on his attained education (γ_1^m) and the quadratic form of his past experience in mode m ($\gamma_{2,m'}^m$ and $\gamma_{3,m'}^m$). Notice that the specification above assumes that unobserved heterogeneity with respect to entrepreneurial skills/talents is used only in self-employment ($\gamma_{1,type}^s$). Additional terms are to capture an effect of work experience in the other mode (γ_4^m), an age effect (γ_5^m), a benefit from staying (γ_6^m), and a first-year experience effect (γ_7^m).

A.4 Entrepreneurial Production Function, $f([\overline{\Psi}_t^s l_t^s], k_t, \epsilon_t^{ys}; \alpha)$ (2 Parameters)

As explained in the main text, the parametric form for f is

$$\begin{aligned} & f([\overline{\Psi}_t^s l_t^s], k_t, \epsilon_t^{ys}; \alpha) \\ &= [\overline{\Psi}_t^s l_t^s]^{1-\alpha} k_t{}^{\alpha} \exp(\epsilon_t^{ys}), \end{aligned}$$

where I consider the dependence of schooling on capital returns, α:

$$\alpha = \alpha_0 + \alpha_1 \cdot I(educ = 1).$$

A.5 Type Proportions, $\Pr(type, initial\ conditions; \eta)$ (4 Parameters)

In the current specification, I assume that there are two unobserved types, $type = 1,2$. The type probabilities are logistic functions of the initial conditions. Specifically, they are written by

$$\Pr(type = 2;\ initial\ conditions,\ \eta) = \frac{\exp\begin{pmatrix} \eta_{0,2} + \eta_{1,2} \cdot I(educ = college) \\ +\eta_{2,2} \cdot \frac{a_1}{10000} + \eta_{3,2} \cdot I(\underline{age} \geq 23) \end{pmatrix}}{1 + \exp\begin{pmatrix} \eta_{0,2} + \eta_{1,2} \cdot I(educ = college) \\ +\eta_{2,2} \cdot \frac{a_1}{10000} + \eta_{3,2} \cdot I(\underline{age} \geq 23) \end{pmatrix}}$$

and $\Pr(type = 1;\ initial\ conditions,\ \eta) = 1 - \Pr(type = 2;\ initial\ conditions,\ \eta)$.[67]

A.6 Variances and the Covariances of the Period-by-Period Disturbance, $\sigma^2_{\epsilon,\cdot}$ (6 Parameters)

The period-by-period disturbances to labor disutility, $\epsilon^l_t = (\epsilon^{ls}_t, \epsilon^{lw}_t)$ and $\epsilon^y_t = (\epsilon^{ys}_t, \epsilon^{yw}_t)$, and to the borrowing constraint, ϵ^a_t, are observed in the beginning of each period. I assume that ϵ^l_t, ϵ^y_t and ϵ^a_t are independently and identically distributed. For ϵ^l_t and ϵ^y_t, I assume serial independence across t, and I allow correlation between ϵ^{ls}_t and ϵ^{lw}_t. Specifically, I assume $\epsilon^l_t \sim N(0, \Sigma^l)$, where the variance-covariance matrix, Σ^l, is given by

$$\Sigma^l = \begin{pmatrix} \sigma^2_{\epsilon,ls} & \cdot \\ \sigma_{\epsilon,(ls,lw)} & \sigma^2_{\epsilon,lw} \end{pmatrix}.$$

Similarly, I assume $\epsilon^{ys}_t \sim N(0, \Sigma^y)$ where

$$\Sigma^y = \begin{pmatrix} \sigma^2_{\epsilon,ys} & \cdot \\ \sigma_{\epsilon,(ys,yw)} & \sigma^2_{\epsilon,yw} \end{pmatrix}.$$

A.7 Quasi-Terminal *Emax* Function, $Emax_{T^*}(\cdot; \kappa_{T^*})$ (5 Parameters)

To ease computational burden, the terminal period is set to be $T^* < T$.[68] I assume the following specification:

$$\begin{aligned} Emax_{T^*} &= Emax_{T^*}(a_{T^*+1}, educ, h^s_{T^*+1}, h^w_{T^*+1}; \kappa_{T^*}) \\ &= \kappa_{T^*,1} \cdot a_{T^*+1} + \kappa_{T^*,2} \cdot \frac{(a_{T^*+1})^2}{10000} \\ &\quad + \kappa_{T^*,3} \cdot I(educ{=}college) \\ &\quad + \kappa_{T^*,4} \cdot h^s_{T^*+1} + \kappa_{T^*,5} \cdot h^w_{T^*+1} \end{aligned}$$

Remember that this form of specification is one of many other alternatives. Future work should elaborate more on this issue.

[67] Variables related to family background and psychological characteristics could be included. In the present study I do not use them because I want to keep the numbers of individuals in the sample as large as possible so that the number of the self-employed does not become smaller in each age. For example, I need to drop 90 out of the current 1,916 individuals.

[68] Specifically, I set $T^* = 30$ for all individuals.

Appendix B Details on the Construction of the Data

The aim of this appendix is to show how the data for estimation,

$$X = \{X_i\}_{i=1}^N$$
$$= \{((l_{i,t}^s, l_{i,t}^w), (y_{i,t}^s, w_{i,t}), a_{i,t+1}, k_{i,t}, age_{i,t})_{t=1}^{\widehat{T}_i}, a_{i1}, educ_i\}_{i=1}^N,$$

is constructed from the original 1979 cohort of the National Longitudinal Survey of Youth (NLSY79).[69] The first and last calender years for which information from data is utilized are 1979 and 2000.[70]

Remember that the decision period in my life-cycle model is a *calender year* (job durations, for example, are measured in terms of years), while various information is available on a weekly or monthly basis. So, the original data must be arranged to match the length of the decision period in the model. Other modifications are necessary to accommodate the data to the life-cycle model.

After showing the construction of the age variable in B.1, I show my rules on how to determine the first decision period for each individual in B.2. B.3 explains the restriction on the person dimension. I then give details on the construction of the main variables in B.4. For the sake of presentations, these processes are explained in order, but the actual process of data construction was not implemented in this order because, for example, the restriction on the person dimension needs some information on variables constructed in B.4.

B.1 Constructing Age Variable, $age_{i,t}$

Respondent's data of birth was asked twice: in survey years 1979 and 1981. Although it is suspicious that there are some misreportings of birth year,[71] I simply use the 1979 information. First, I calculate an individual's age in months at each interview date by

interview date (month/year) - DOB (month/year) in the 1979 survey.

Then, I compute his age in months as of January of the interview year simply by

age at interview date (in months) - (interview month -1) (in months).

B.2 Determining Age in the First Decision Period

Remember that schooling decision is not modelled in the life-cycle model. The individual in the model starts decisions one year after when he completed schooling. His year of schooling is taken as an exogenous variable in the dynamic model, and it does not change over time. To follow each individual in the data from his first decision period, I need to determine when he is considered to have finished schooling. Again, note that the decision period of my choice

[69] All the original data was retrieved online at the "NLS Investigator" (http://www.nlsinfo.org/web-investigator/).

[70] While I use the survey rounds up to 2002, I do not take 2002 as the final year because the year 2002 survey did not collect information on assets. The reason why I also use the year 2002 survey is that it covers weekly labor status information after the interview date in 2000.

[71] I found 14 observations (out of the core while male sample; 2,439 observations) are suspected to have misreported his birth year. The number is small, so this problem should be minor. Also note that all original 2439 respondents answered the question in 1979. For details, see the documation attched to the data (will be available in due course).

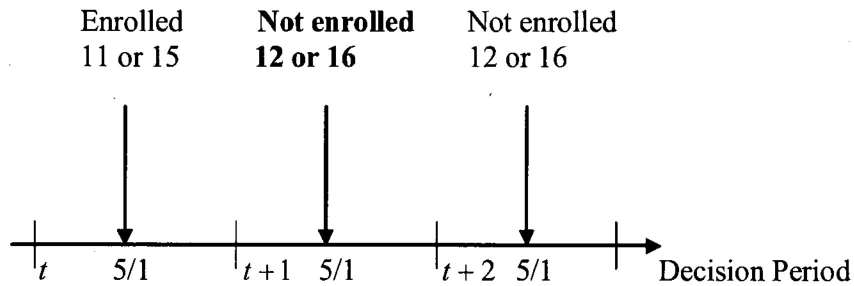

Figure 28: Typical Pattern of Transition from Schooling to Work/Non-Work

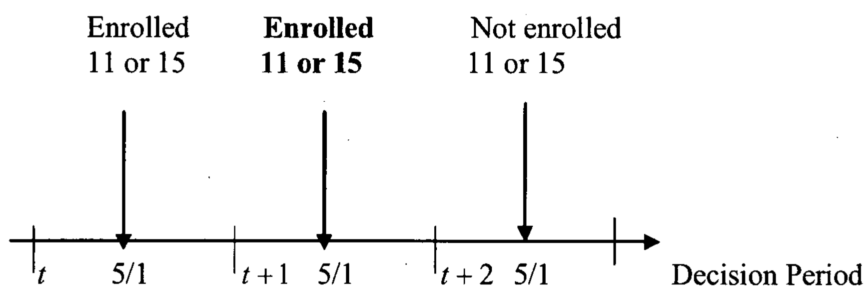

Figure 29: An Ambiguous Case of Transition

is a *calender* year. I must therefore be careful about the differences between calender years and *school* years, because it matters to the transition from schooling to work/non-work.

At each survey round, each respondent's school enrollment status as of May 1 of each year is available. I collapse the four categories in the original data into two as follows:

$$\begin{cases} \text{Enrolled} \leftarrow \text{"Enrolled in high school" or "Enrolled in college"} \\ \text{Not Enrolled} \leftarrow \text{"Not enrolled, completed less than 12th grade"} \\ \qquad\qquad\qquad\quad \text{or "Not enrolled, high school graduate."} \end{cases}$$

Using this information on "school enrollment status as of May 1st" (denoted by "$enrollment_{i,t}$"), however, alone may not be precise. This problem is depicted as follows.

In most cases, we expect to observe patterns as shown in Figure 28. The horizontal arrow shows the timeline, and each partition corresponds to one period (t, $t+1$ and $t+2$). For each period, the status of school enrollment (the first row) and the accumulated year of schooling (the second row) are observed. The story that would be the most plausible to the information in this figure is the following. The individual graduated from school at some point after May 1st of period t, obtaining one more year of schooling. This is found by the information in period $t+1$: he is not enrolled in school and his year of schooling increased from the one in the previous year If that is the case, it is natural to assume that year $t+1$ is the first decision period.

Now, suppose that we observe a pattern as in Figure 29. In this case, it would be natural to assume that he stopped schooling in year t, even though he was reportedly enrolled in school year $t+1$ because he did not obtain one more year of schooling.

To avoid this type of ambiguity, I need to look at *changes* in his "Highest Degree Completed as of May 1st" (denoted by variable "$completed_{i,t}$") to determine the first decision period. Essentially, I want to find when his $completed_{i,t}$ stopped to increase. The first decision period should be one year after the year when his $completed_{i,t}$ firstly stopped to increase.

When I see an individual's $completed_{i,t}$ go up again after years of constant $completed_{i,t}$, I judge whether or not he is considered to have temporary left for additional schooling (see Restriction PD-4 in B.3). More formally, I adopt the following rules to determine when the first decision period is.

Rule 1. If individual i is not enrolled in school as of May 1st of year t ($enrollment_{i,t} = 0$), then I say he is not enrolled in school in year t ($in\_school_{i,t} = 0$).

Rule 2. If individual i is enrolled in school as of May 1st of year t ($enrollment_{i,t} = 1$), then I say he is enrolled in school in year t ($in\_school_{i,t} = 1$) if

$$completed_{i,t+1} > completed_{i,t},$$

and he is not enrolled in school year t ($in\_school_{i,t} = 0$) if

$$completed_{i,t+1} \leq completed_{i,t}.$$

In this way, for each individual, any calender year is categorized into "1" (attended school) or "0" (did not attend school), as long as "school enrollment status as of May 1" and "highest grade completed as of May 1" are available for that year. For most of individuals, "1"s appear in a row when young and "0"s in subsequent years. This case has no difficulty in determining the first decision period. For other individuals, I decided whether or not he is judged to have temporary left for additional schooling by looking at the computed hours worked and the monthly information on school enrollment.[72]

B.3 Restriction on the Person Dimension, N

I employ the following steps to restrict on the person dimension (PD) for the data used for this study.

- Restriction PD-1. I extract the *white male* part in the *core random* sample.[73] This reduces the initial sample size 12,686 to 2,439.

- Restriction PD-2. Next, to drop individuals who have served in the *military*, I look at the "Weekly Labor Force Status" Section (from Week 0 in 1978 to Week 52 in 2000).[74] If an individual's labor force status is "military" in any week since January 1, 1978, then he is excluded from the sample. I find 268 observations have ever been in the military (10.99%). This reduces the sample size 2,439 to 2,171.

- Restriction PD-3. Using the defined occupation for individual i in year t (see B.4.1. below), I exclude *professionals* and *farmers*. First, I exclude individuals who experienced any of the following occupations in any year t: "Accountants and auditors," "Lawyers and judges," "Health diagnosing occupations," and "Farming, forestry & fishing occupations." I then find, among the nonmilitary experienced, 361 observations have ever experienced professional or farmer.[75] I put 259 of them back to the sample if they were

[72] Monthly attendance record is available after January, 1980.

[73] One can retrive all necessary data online by filitering "R0173600 $\leq$ 2" ("R0173600" is the sample identification code).

[74] The 1979 year survey covers weekly labor status in 1978 as well.

[75] Among those already excluded in Restriction CS-2, 26 observations have ever experienced professional or farmer.

a professional or a farmer before the first decision period, if they are judged to have temporally worked for such jobs, or if they were in such occupations as pharmacist and registered nurse. The number of excluded individuals is now 102 (4.70%). This process reduces the sample size 2,171 to 2,069.

- Restriction PD-4. I drop individuals who are judged to have temporary left for adult schooling (24 individuals),[76] to have started working late (26 years old or older; 80 individuals) as well as 2 individuals whose first decision period is judged 14 years old. I also exclude individuals if it is difficult to determine the first decision period (6), or if no survey years are covered when working (41). The number of excluded individuals is 153 (7.39%), and this process reduces the sample size 2,069 to 1,916.

My final sample consists of 1,916 white males with a total of 32,166 person-year observations.

B.4 Construction of the Main Variables

Here I show how the main variables are constructed from the NLSY79. First, I utilize the weekly information is available whether or not someone was self-employed.

B.4.1 Defining Self-Employment and Wage-Employment

To obtain information on whether an individual is a self-employer or a wage worker in a year, I look at the "Class of Worker" Section (up to 2002 survey year). For each "job"[77] (up to five jobs for each survey year) a respondent reports whether he[78]

(1) worked/works for a private company or individual for wages, salary, or commission,
(2) was/is a government employee,
(3) was/is *self-employed* in his/her own business, professional practice or farm, or
(4) was/is working without pay in a family business or farm,

for that job. In the NLSY79, the respondent is classified (by his/her answers to the job classification questions) as *self-employed* if[79]

> "he or she owned at least 50 percent of the business,
> was the chief executive officer or principal managing partner of the business, or
> was supposed to file a form SE for Federal income taxes"

or he or she identifies himself or herself as

> "an independent contractor, independent consultant, or free-lancer."

Using this information, I associate each job with information on whether the respondent worked as a self-employer or as a paid-worker for that job. Specifically, a job is attached

[76] I made judgment by looking at calculated hours worked, changes in the highest degree completed, and the monthly school enrollment information.

[77] All references to a "job" should be understood as references to an employer.

[78] Category (1) includes individuals working for pay for settlement houses, churches, unions, and other private nonprofit organizations until 1994 when these begun to be independently coded.

[79] See ftp://www.nlsinfo.org/pub/usersvc/NLSY79/NLSY79%202004%20User%20Guide/79text/cow.htm.