Figures and Tables

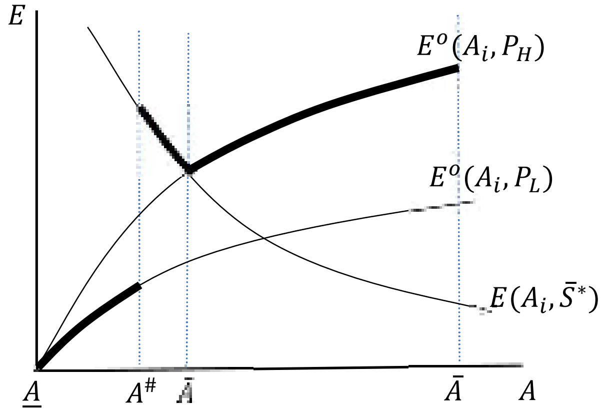
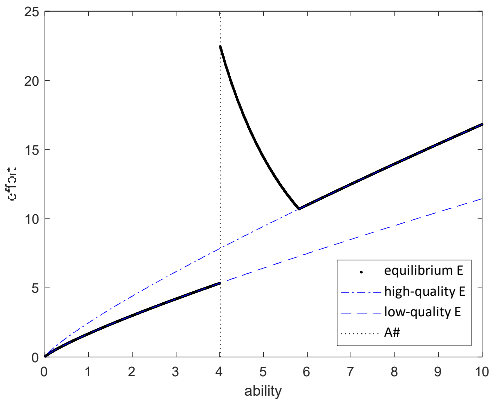
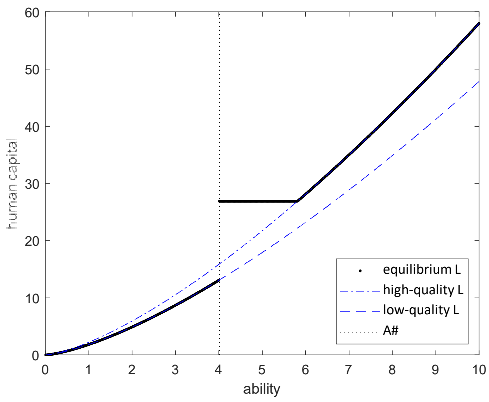


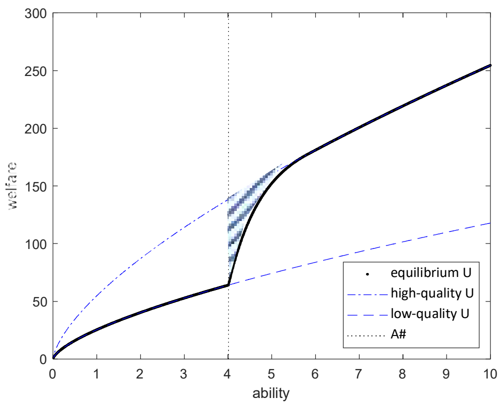
Figure 1 Student effort under competitive equilibrium



(a) Effort



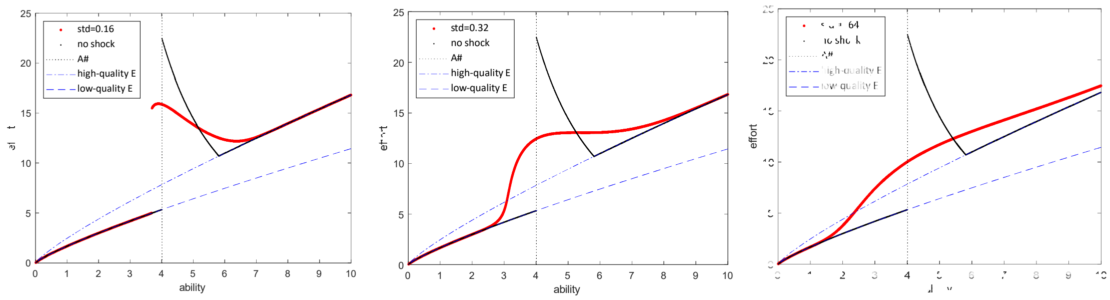
(b) Score (pre-college HC)



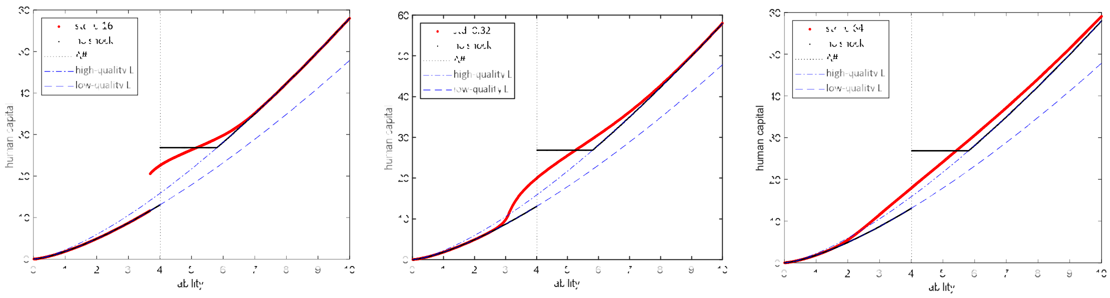
(c) Utility

Figure 2 Competitive equilibrium outcome (without random perturbations)

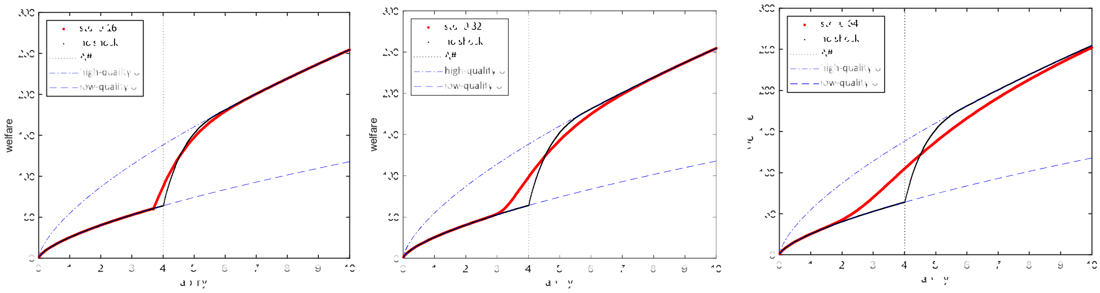
Note: Student number is 1,000, with abilities following a uniform distribution between 0 and 10. The quota of high-quality college is 600. The quality ratio between high-quality college and low-quality college is 4. For more details on parameters, see Table 1.



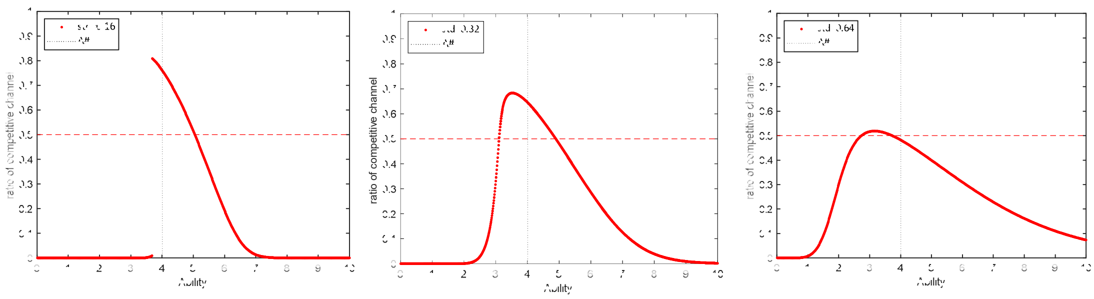
(a) effort



(b) pre-college HC



(c) welfare

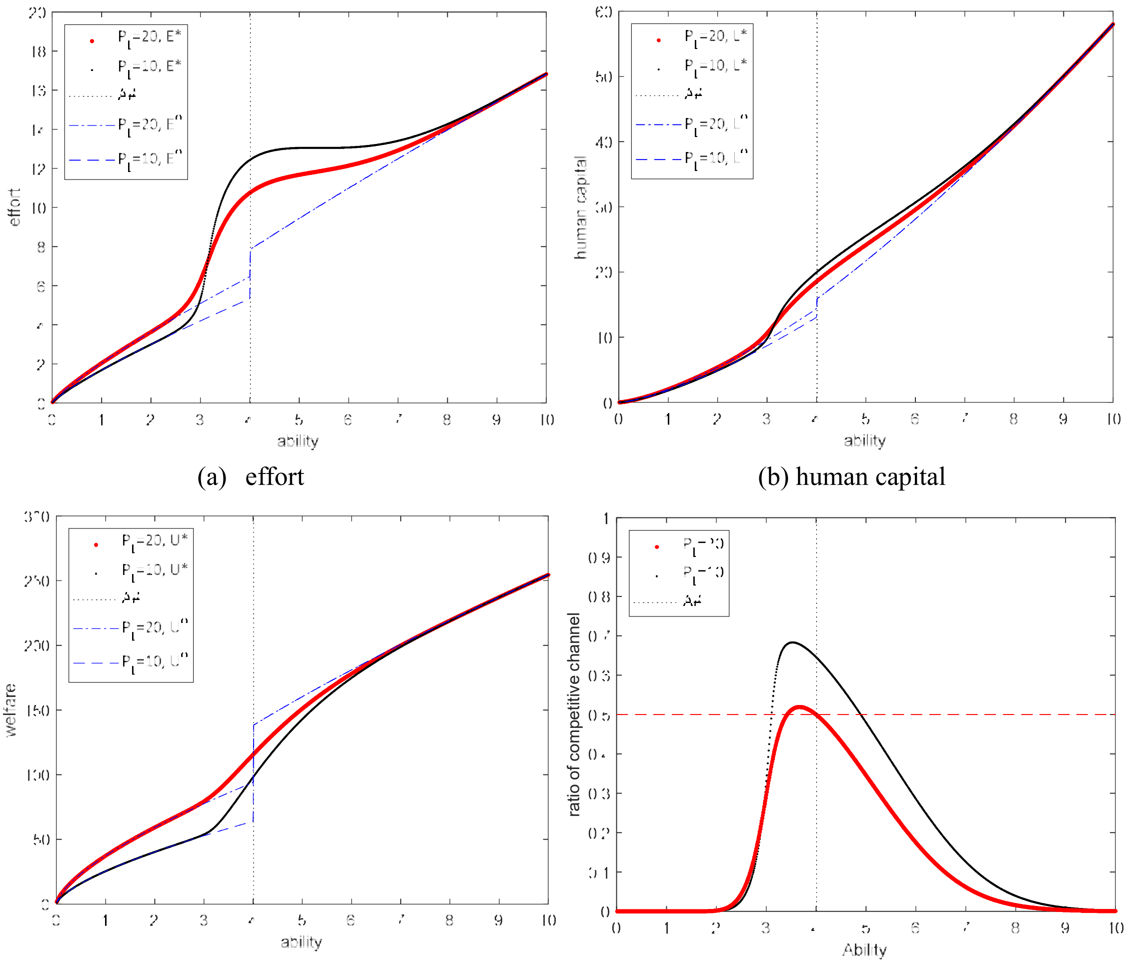


(d) competitive channel

Figure 3 Competitive equilibrium outcome (with random perturbations)

($\sigma_\varepsilon =$ 0.16 (left), 0.32(middle), 064 (right))

Note: Student number is 1,000, with abilities following a uniform distribution between 0 and 10. The quota of high-quality college is 600. The quality ratio between high-quality college and low-quality college is 4. For more details on parameters, see Table 1.

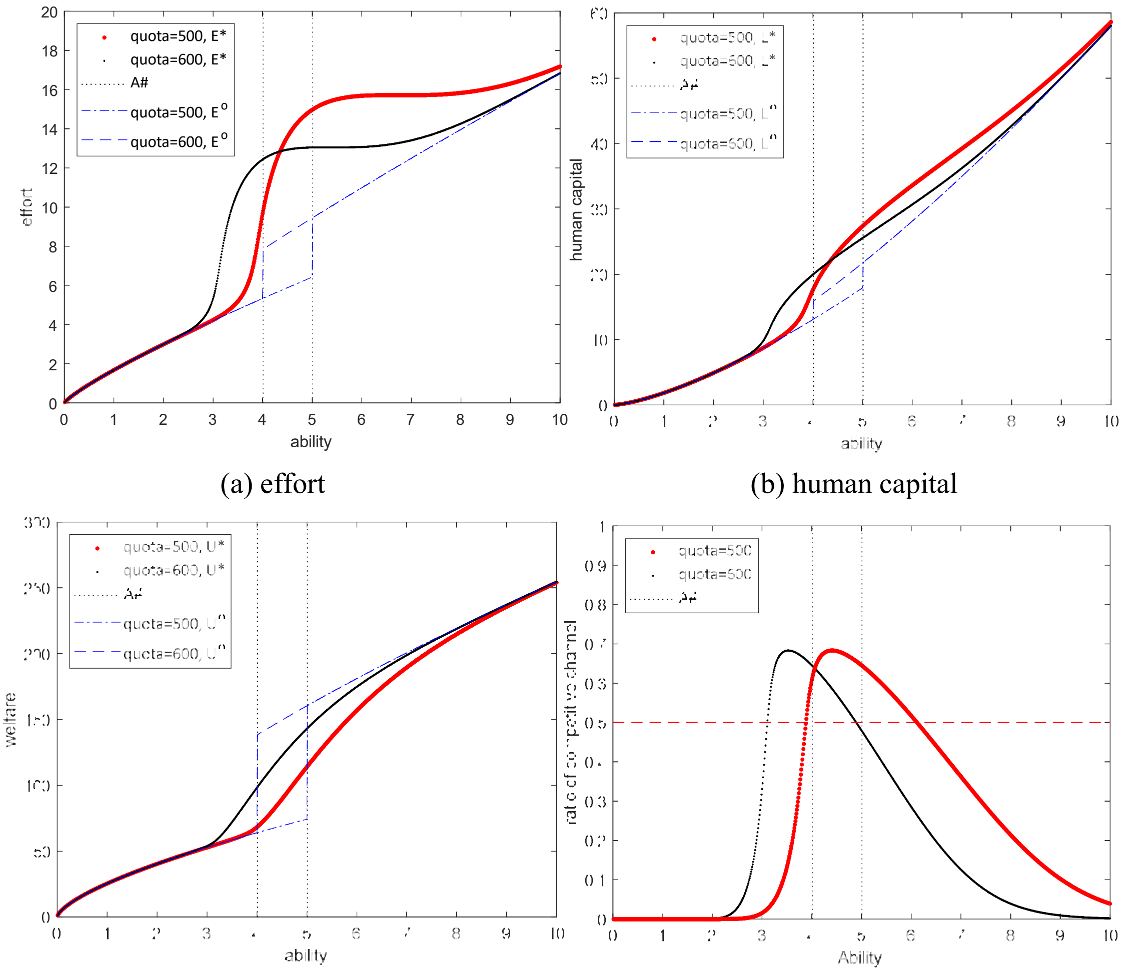


(a) effort (b) human capital

(c) welfare (d) competitive channel ratio

Figure 4 Comparative statics: Decreasing quality gap

Note: Student number is 1,000, with abilities following a uniform distribution between 0 and 10. The quota of high-quality college is 600. The quality of high-quality college is 40, and the quality of low-quality college is 20 (red) or 10 (black). Random perturbation is $\sigma_\varepsilon = 0.32$. For more details on parameters, see Table 1.

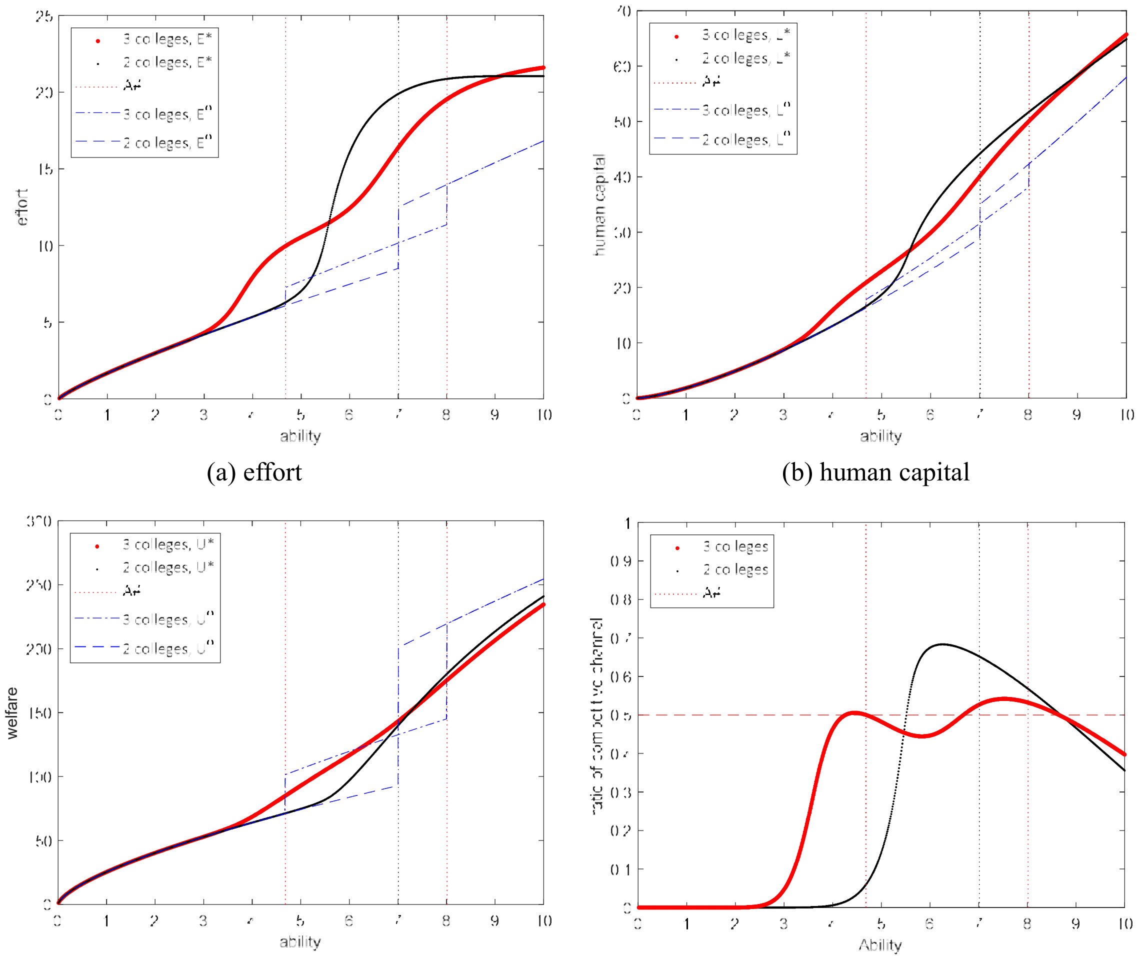


(a) effort (b) human capital

(c) welfare (d) competitive channel

Figure 5 Comparative Statics: Decreasing high-quality college quota

Note: Student number is 1,000, with abilities following a uniform distribution between 0 and 10. The quality ratio between high-quality college and low-quality college is 4. The quota for the high-quality college is 500 (red) or 600 (black). Random perturbation is $\sigma_\varepsilon = 0.32$. For more details on parameters, see Table 1.

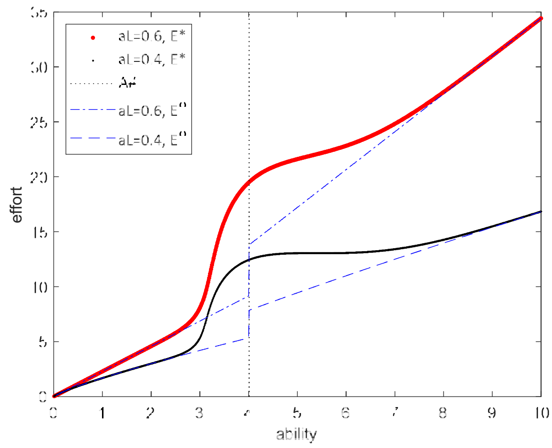


(a) effort (b) human capital

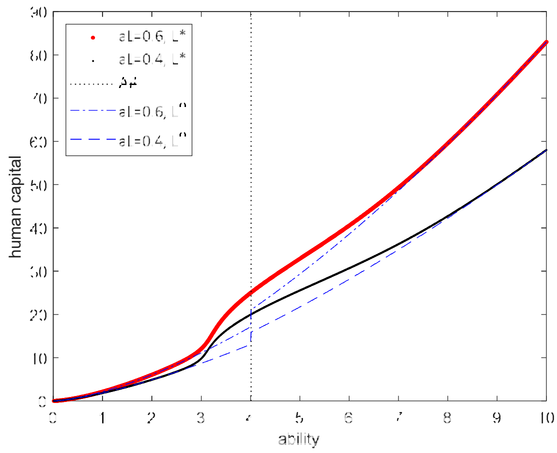
(c) welfare (d) competitive channel

Figure 6 Comparative statics: Adding one median-quality college

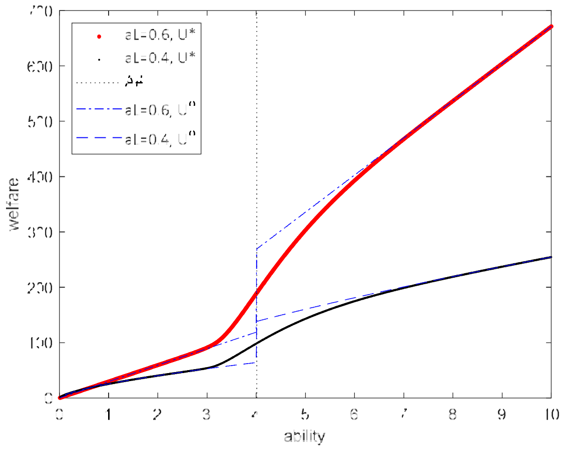
Note: Student number is 1,000, with abilities following a uniform distribution between 0 and 10. The black-colored curve shows the 2-college case where the quality ratio between high-quality college and low-quality college is 4. The red-colored curve shows the 3-college case by adding one median-quality college. Random perturbation is $\sigma_{\varepsilon} = 0.32$. For more details on parameters, see Table 1.



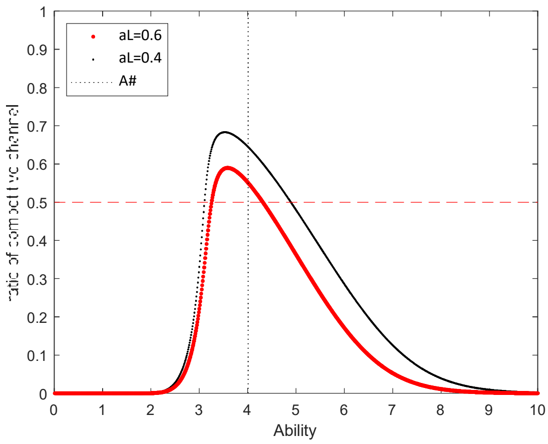
(a) effort



(b) human capital



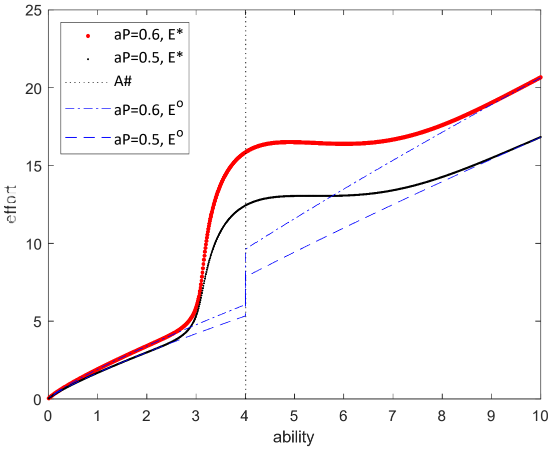
(c) welfare



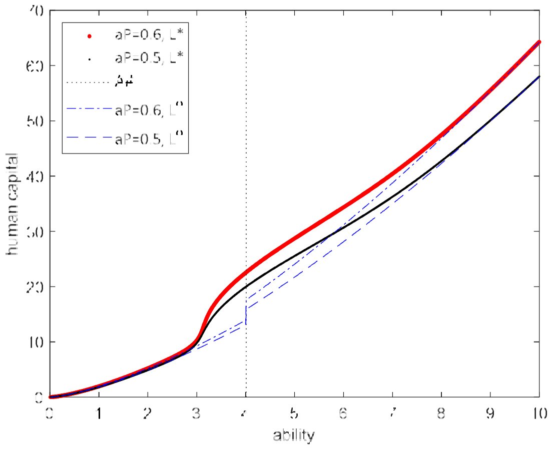
(d) competitive channel ratio

Figure 7 Comparative statics: Increasing the power of human capital in wage equation

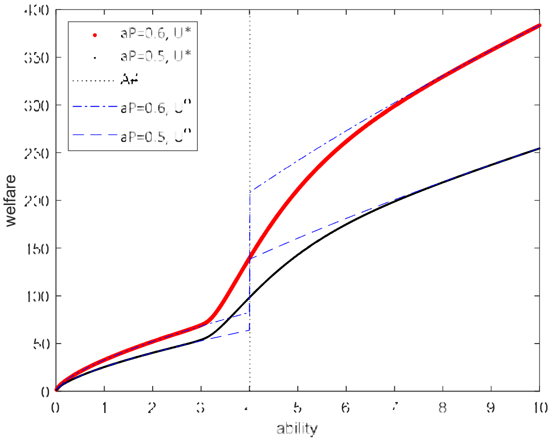
Note: Student number is 1,000, with abilities following a uniform distribution between 0 and 10. The quality ratio between high-quality college and low-quality college is 4. The power of human capital (α_L) is 0.6 (red) or 0.4(black). Random perturbation is $\sigma_\varepsilon = 0.32$. For more details on parameters, see Table 1.



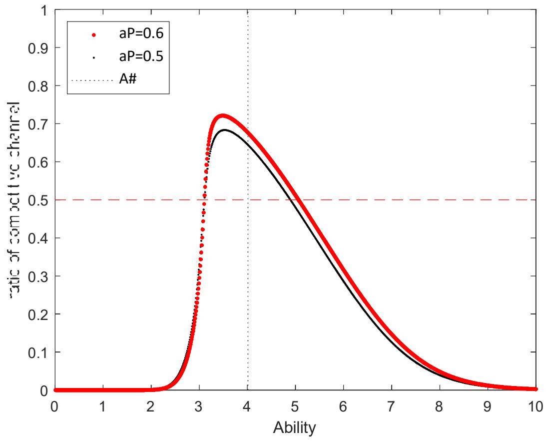
(a) effort



(b) human capital



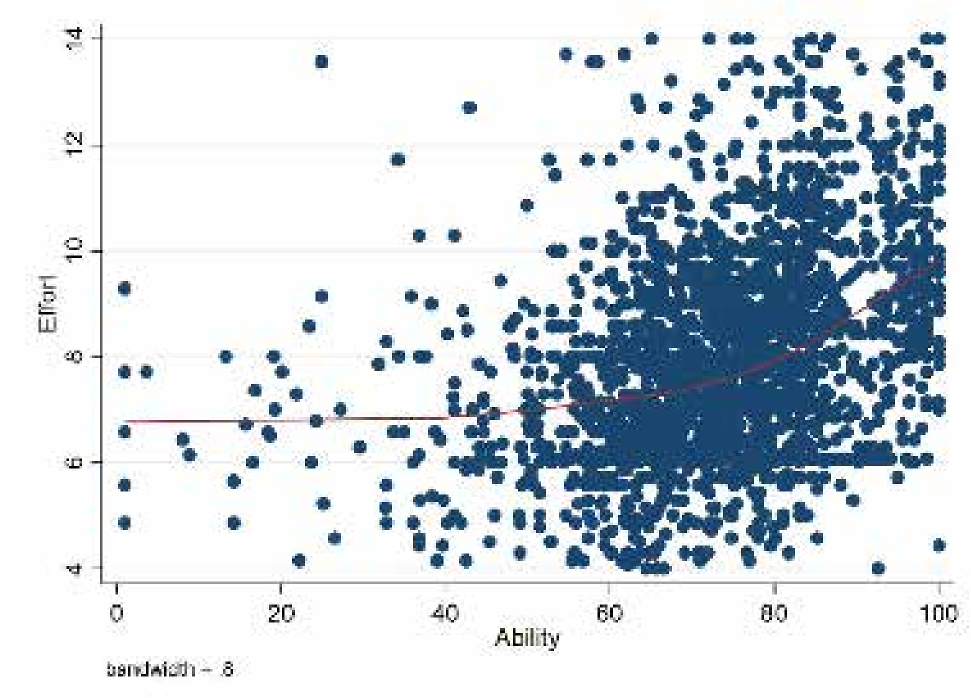
(c) welfare



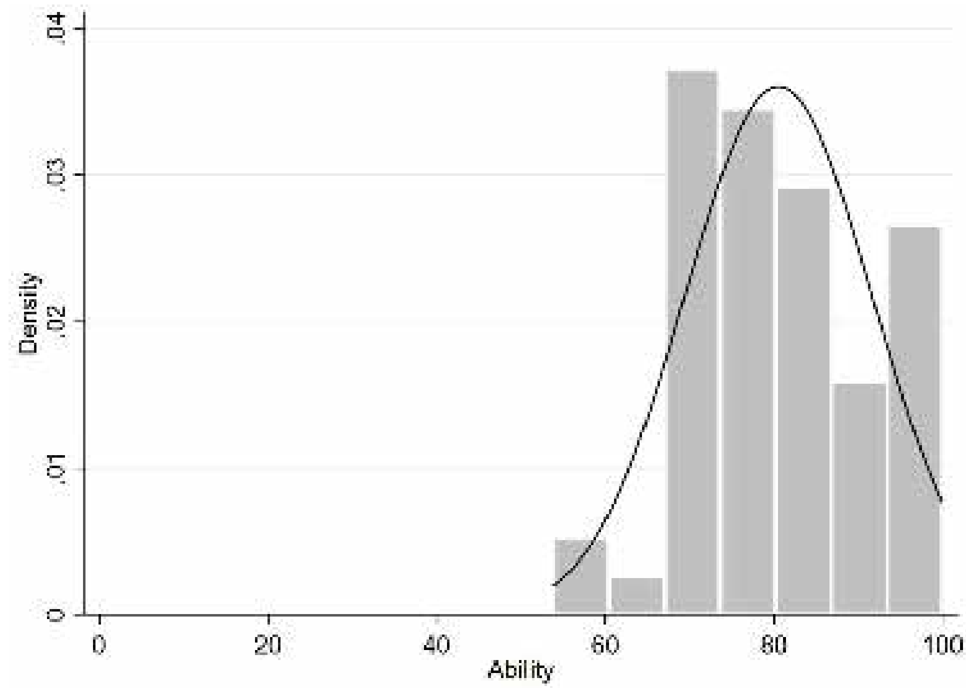
(d) competitive channel ratio

Figure 8 Comparative statics: increasing the power of college quality in wage equation

Note: Student number is 1,000, with abilities following a uniform distribution between 0 and 10. The quality ratio between high-quality college and low-quality college is 4. The power of college quality (α_P) is 0.6 (red) or 0.5(black). Random perturbation is $\sigma_\varepsilon = 0.32$. For more details on parameters, see Table 1.



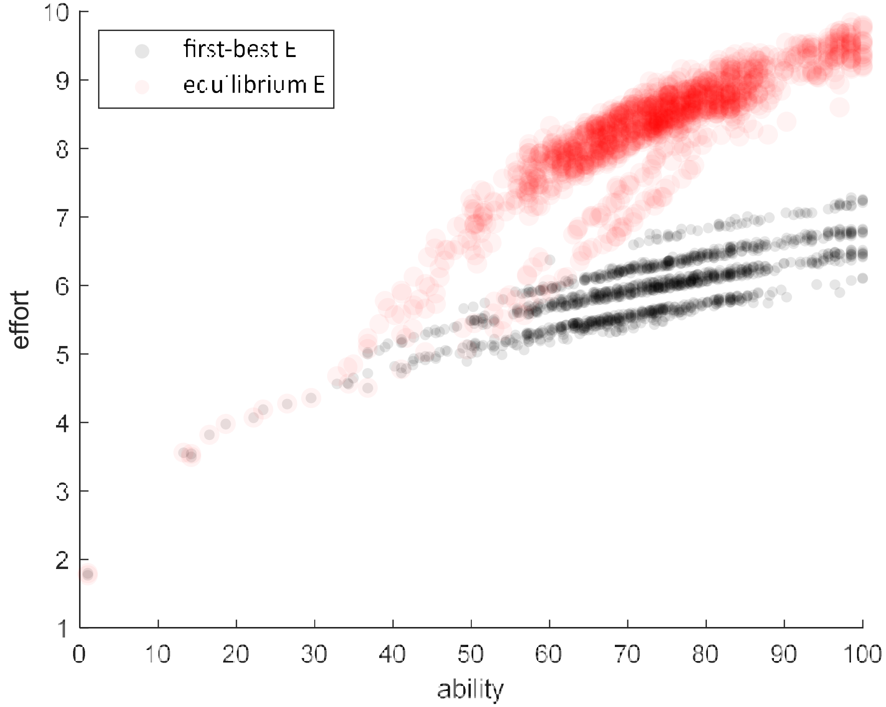
(a) student ability and daily study time (effort)



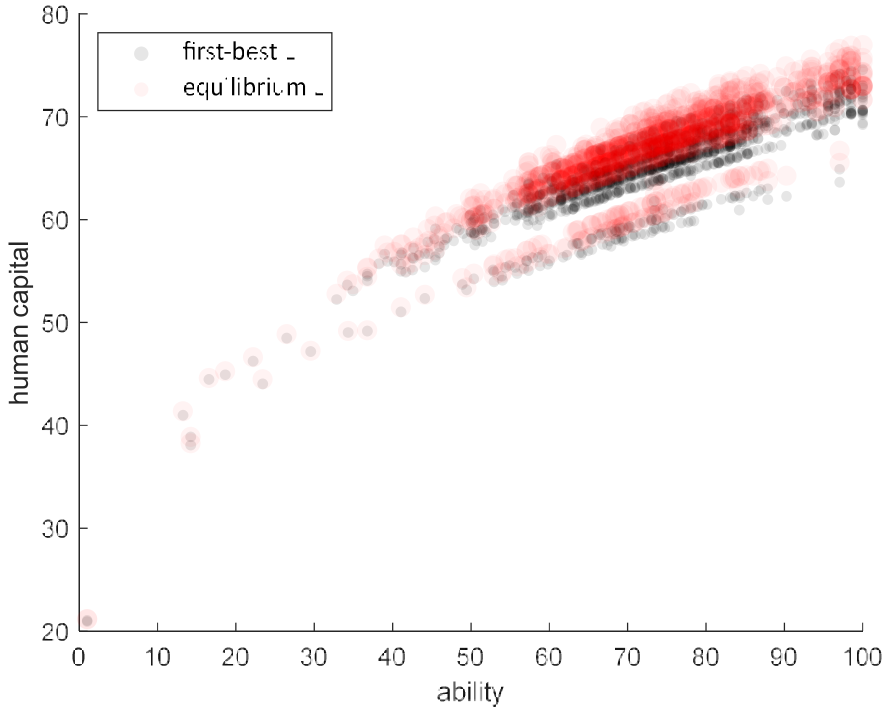
(b) ability distribution of students enrolled in 211-colleges

Figure 9 The relationship of student ability and daily study time

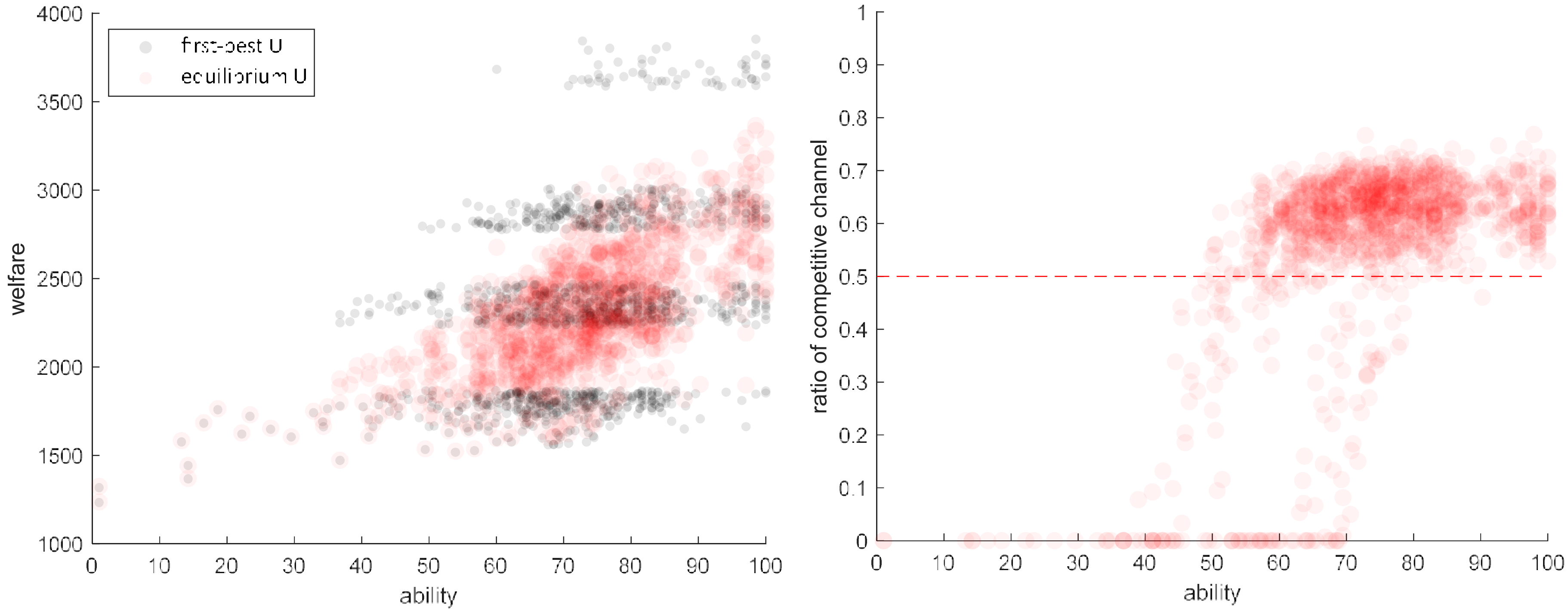
Note: Data source is CFPS (2010-2020). Data are censored by eliminating the highest and lowest 1% of the study time. Locally weighted regression scatter smoothing plot (Lowess model) is in part (a). The correlation coefficient between student ability and study time is 0.308 and significant at 1% level.



(a) effort



(b) human capital



(c) welfare

(d) competitive channel ratio

Figure 10 Simulated equilibrium under the current system

Note: The simulated sample is from CFPS, with a sample size of 1,248 (Table A7, Column (3)). For the simulated parameters, see Table 5.

Table 1 Parameter setting in numerical example

	Meaning	Value
Human capital production function	Constant term	$\alpha_0 = \sqrt{2}$=1.41
	Power of student ability (A)	$\alpha_1 = 1$
	Power of student effort (E)	$\alpha_2 = 0.5$
	Random perturbation	$\varepsilon \sim N(0, \sigma_\varepsilon^2)$, $\sigma_\varepsilon = 0$, 0.16, 0.32, 0.64
Cost function	Power of student effort (E)	$\mu = 2$
Wage determination equation	Constant term	$\alpha_w = 7 * 2^{0.1} = 7.50$
	Power of college quality (P)	$\alpha^P = 0.5$
	Power of human capital (L)	$\alpha^L = 0.4$
Other Parameters	Number of Students	$N = 1000$
	Student ability distribution	$A \sim \mathrm{U}(0,10)$
	College qualities	$P_H = 40;\ P_L = 10$
	Quota proportion of the high-quality college	$q_H = 0.6$

Table 2 Relationship between student effort and college admission rate (CFPS data)

Dependent variable: Average daily study time (hours)				
	(1)	(2)	(3)	(4)
Adm_rate	3.923	19.83***	-0.286	14.87**
	(2.318)	(4.785)	(3.315)	(6.442)
Adm_rate^2		-62.48***		-51.05***
		(16.26)		(16.46)
Ln(Finc)	0.072	0.069	0.057	0.057
	(0.076)	(0.076)	(0.077)	(0.077)
ParentEdu	0.014	0.019	0.018	0.019
	(0.021)	(0.021)	(0.021)	(0.021)
Male	-0.501***	-0.499***	-0.485***	-0.483***
	(0.117)	(0.119)	(0.119)	(0.120)
Urban	0.109	0.095	0.092	0.087
	(0.201)	(0.199)	(0.203)	(0.200)
Grade 2 (vs 1)	0.191	0.200	0.228*	0.231*
	(0.125)	(0.126)	(0.125)	(0.125)
Grade 3 (vs 1)	0.794***	0.794***	0.863***	0.863***
	(0.143)	(0.141)	(0.145)	(0.145)
Prov FE	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	Yes
N	2,660	2,660	2,660	2,660
R^2	0.061	0.066	0.069	0.072

Note: The data source is CFPS (2010-2020) and the sample includes all high school students. Admission rate is the number of students admitted by batch-1 colleges divided by the total number of college entrance exam takers. Standard errors (in parentheses) are clustered at the individual level in Columns (1)-(2) and are clustered at the provincial level in Columns (3)-(4). *p<0.1，**p<0.05，***p<0.01.

Table 3 Estimating human capital production function

	Dependent variable: Ln(S)				
	All students	The last test in middle school as ability	Years between two tests ≥ 6		
	(1)	(2)	(3)	(4)	(5)
Ln(A)	0.481***	0.559***	0.132*	0.122	0.123*
	(0.11)	(0.10)	(0.07)	(0.0798)	(0.06)
Ln(E)	0.041*	0.034*	0.039***	0.0770	0.042***
	(0.02)	(0.02)	(0.01)	(0.0848)	(0.01)
Ln(Finc)	0.001	0.001	-0.002	-0.00156	-0.005
	(0.01)	(0.01)	(0.01)	(0.00660)	(0.01)
Ln(ParentEdu)	0.035***	0.031***	0.023***	0.0242***	0.019***
	(0.01)	(0.01)	(0.01)	(0.00670)	(0.01)
Ln(FamEduExp)					0.008**
					(0.00)
Male	-0.009*	-0.009*	0.008	0.00661	0.014*
	(0.01)	(0.01)	(0.01)	(0.00930)	(0.01)
Urban	0.014	0.012	0.007	0.00248	0.012
	(0.01)	(0.01)	(0.01)	(0.0135)	(0.01)
Constant	2.121***	1.800***	3.692***	3.654***	3.726***
	(0.44)	(0.39)	(0.31)	(0.224)	(0.26)
N	2,468	2,468	1,306	1268	1,248
R^2	0.41	0.468	0.068	0.056	0.094

Note: Data source is CFPS (2010-2020). Robust standard errors clustered at the provincial level are in parentheses. *=p<0.1, **=p<0.05, ***=p<0.01.

Table 4 Estimating wage equation

Dependent variable: Ln(w)		
	Unweighted	Weighted
	(1)	(2)
College211	0.444***	0.468***
	(0.035)	(0.040)
Tier1	0.279***	0.297***
	(0.034)	(0.040)
Tier23	0.162***	0.180***
	(0.031)	(0.038)
Ln(L)	0.499***	0.452***
	(0.058)	(0.090)
Ln(Finc)	0.045***	0.046***
	(0.008)	(0.010)
Ln(ParentEdu)	0.027	0.026
	(0.025)	(0.023)
Male	0.098***	0.102***
	(0.017)	(0.018)
Urban	-0.003	-0.004
	(0.019)	(0.021)
Constant	5.353***	5.526***
	(0.231)	(0.364)
N	3148	3145
R^2	0.149	0.152

Note: The data source is CCSS 2013. Robust standard errors clustered to the provincial-track level are in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Table 5 Estimated parameter values

	Parameter	Value		Parameter	Value
Human capital production function	α_0	Exp(3.692)	**Wage equation**	α_w	Exp(5.526)
	α_1	0.132		α_{P1}	0.468
	α_2	0.039		α_{P2}	0.297
	α_{31}	-0.002		α_{P3}	0.180
	α_{32}	0.023		α_L	0.452
	α_{33}	0.008		α_{Q1}	0.046
	α_{34}	0.007		α_{Q2}	0.026
	σ_ε	0.075		α_{Q3}	0.102
	$\hat{b}$	0.536		α_{Q4}	-0.004
Cost function	μ	4.052			

Note: The parameters for human capital production function are estimated from CFPS data, see column (4) Table 4. The parameters from wage determination function are estimated from CCSS data, see column (2) Table 5. The cost function parameter is estimated by minimizing the sum of square of distance between equilibrium efforts in the model and the observed efforts of all sampled students. To predict true human capital, we converting all the coefficients in human capital production function (including constants) by using $\hat{b}$ as $\widehat{ln(\alpha_0)}' = \frac{1}{\hat{b}} * \widehat{Ln(\alpha_0)} + \left(1 - \frac{1}{\hat{b}}\right) * Ln(100), \widehat{\alpha_i}' = \frac{1}{\hat{b}} * \widehat{\alpha_i}, i = 1,2,3k.$

Table 6 Counterfactual supply-side policies

Counterfactual Policy		College quality $(\alpha_{P1,}\alpha_{P2},\alpha_{P3},\alpha_{P4})$	College quota (Σ=1248) (q_1, q_2, q_3, q_4)
	(1) benchmark	(0.468, 0.297, 0.180, 0)	(59, 303, 489, 397)
reduce quality gap among colleges	(2) eliminate gap between tier 1&2	(0.325, 0.325, 0.180, 0)	(59, 303, 489, 397)
	(3) eliminate gap between tier 2&3	(0.468, 0.2248,0.2248, 0)	(59, 303, 489, 397)
	(4) eliminate gap between tier 3&4	(0.468, 0.297, 0.0994, 0.0994)	(59, 303, 489, 397)
Increase high-quality college quota	(5) tier 1 by 10%, with quality reduced by 10%	(0.4212, 0.297, 0.180, 0)	(65, 303, 489, 391)
	(6) tier 1&2 by 10%, with quality reduced by 10%	(0.4212, 0.2673, 0.180, 0)	(65, 333, 489, 361)
	(7) tier 1-3 by 10%, with quality reduced by 10%	(0.4212, 0.2673, 0.162, 0)	(65, 333, 538, 312)
Add one median-quality college	(8) $\alpha_{P5} = (\alpha_{P1} * q_1 + \alpha_{P2} * q_2)/(q_1 + q_2)$, $q_5 = (q_1 + q_2)/3$	(0.468, 0.325=α_{P5}, 0.297, 0.180, 0)	(39, 121=q_5, 202, 489, 397)
	(9) $\alpha_{P5} = (\alpha_{P2} * q_2 + \alpha_{P3} * q_3)/(q_2 + q_3)$, $q_5 = (q_2 + q_3)/3$	(0.468, 0.297, 0.225=α_{P5}, 0.180, 0)	(59, 202, 264=q_5, 326, 397)
	(10) $\alpha_{P5} = (\alpha_{P3} * q_3 + \alpha_{P4} * q_4)/(q_3 + q_4)$, $q_5 = (q_3 + q_4)/3$	(0.468, 0.297, 0.180, 0.099=α_{P5}, 0)	(59, 303, 326, 295=q_5, 265)

Change number of students	(11) reduce the number of students by 10%	(0.468, 0.297, 0.180, 0)	(59, 303, 489, 272)
	(12) increase the number of students by 10%	(0.468, 0.297, 0.180, 0)	(59, 303, 489, 522)

Note: The benchmark college quality parameters (α_P) are estimated from wage determination equation and shown in Table 5. The college quotas (q) are drawn from CFPS sample (Table A7, column (3)).

Table 7 Simulated equilibrium under counterfactual policies

Counterfactual policy		Average effort (hours/day)	Average human capital (0-100)	Average utility (RMB/ month)	First best average utility (RMB/ month	Dead-weight loss (%)	Competitive channel ratio (0-1)	Cutoff scores (0-100)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
First best and Benchmark	(0) first best (of benchmark)	5.94	64.61	2351	2351	0	0	71.31, 67.10, 63.17
	(1) benchmark	8.23	66.59	2315	2351	1.53	0.564	78.39, 70.58, 63.72
Reduce Quality gap	(2) eliminate quality gap between tier 1&2 colleges	8.11	66.52	2316	2348	1.35	0.543	[78.12], 70.48, 63.70
	(3) eliminate quality gap between tier 2&3 colleges	8.00	66.46	2318	2344	1.12	0.524	78.15, [70.38], 63.61
	(4) eliminate quality gap between tier 3&4 colleges	7.74	66.27	2314	2339	1.07	0.428	78.37, 70.37, [63.26]
Increase quota of high-quality colleges	(5) increase quota of tier 1 college by 10%	8.18	66.56	2315	2348	1.41	0.557	77.96, 70.44, 63.61
	(6) increase quota of tier 1&2 colleges by 10%	8.07	66.49	2313	2342	1.26	0.538	77.87, 69.89, 63.10
	(7) increase quota of tier 1-3	8.02	66.47	2311	2339	1.21	0.524	77.86, 69.86, 62.28

	colleges by 10%							
Add one median-quality college	(8) add one college between tier 1&2 colleges	8.22	66.58	2315	2350	1.48	0.575	79.73, [74.55], 70.57, 63.71
	(9) add one college between tier 2&3 colleges	8.21	66.58	2315	2350	1.47	0.593	78.37, 72.30, [68.19], 63.71
	(10) add one college between tier 3&4 colleges	8.20	66.57	2314	2348	1.46	0.594	78.39, 70.56, 65.98, [61.51]
Change number of students	(11) reduce the total number of students by 10%	8.21	66.55	2359	2394	1.47	0.554	77.93, 69.87, 62.19
	(12) increase the total number of students by 10%	8.23	66.59	2282	2318	1.56	0.574	78.73, 71.11, 64.82

Note: The simulated sample is from CFPS, with a size of 1,248 (Table A7, Column (3)). For the simulated parameters for benchmark and the first best, see Table 5. For simulated parameters for counterfactual policies, see Table 6. Numbers in the square brackets show the cutoff scores of merged or added colleges.