Economics of Education Review 53 (2016) 159–163

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

"(Un)informed College and Major Choice": Verification in an alternate setting

Nick Huntington-Klein*

Department of Economics, California State University Fullerton, 800 N. State College Blvd., Fullerton, CA 92831-3599, United States

ARTICLE INFO

Article history:
Received 7 November 2015
Revised 8 March 2016
Accepted 18 March 2016
Available online 30 March 2016

JEL Classification:
D83
I23
I26

Keywords:
Costs
Demand for schooling
Rate of return
Information

ABSTRACT

In their recent paper "(Un)informed College and Major Choice: Evidence from Linked Survey and Administrative Data," Hastings, Neilson, Ramirez, & Zimmerman (2016) provide an informal costly-information model, linking family background to students' beliefs about educational costs and benefits. They verify predictions of their model using a data set of beliefs about college institutions and majors among Chilean college applicants and students. I test some of those same predictions using a data set of beliefs about college institutions and different levels of college education among high school students in the United States. I verify their predictions, with some exceptions, supporting the use of their costly-search model.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The decision to pursue education has significant labor market implications. To approach the decision rationally, a student must consider the costs and benefits of each available option. However, mounting empirical evidence shows that student expectations of costs and benefits are noisy and vary across students.

Hastings, Neilson, Ramirez, and Zimmerman (2016) (hereafter HNRZ) contribute to this literature with their large-scale study of student beliefs in Chile. They gather beliefs about the earnings and costs associated with different college institutions and majors from Chilean college students and applicants. The authors aim to provide descriptive evidence on the characteristics of these beliefs, and show how they relate to student choice.

HNRZ is a part of Proyecto 3E. Proyecto 3E is a study of college and career choice in Chile, carried out with the support of the Chilean government. The Proyecto 3E database follows fifteen cohorts of high school graduates through college and is linked to administrative Chilean government data. The use of novel interventions as well as depth and size of the data set (the HNRZ sample includes 7382 students) gives Proyecto 3E the potential to be one of the most fruitful studies of college and curriculum choice in the modern literature, and to improve the Chilean education system.

HNRZ is not the first product of Proyecto 3E. Previous work has produced causal estimates from regression discontinuity of the returns to different majors, which have commonly been elusive and difficult to identify in prior literature (Hastings, Neilson, & Zimmerman, 2013). Other work has examined the effect of student loan caps based on the earnings of prior graduates in a given major, and how the caps may guide low-income students to choose higher-earnings majors (Beyer, Hastings, Neilson, & Zimmerman, 2015). Hastings, Neilson, and Zimmerman (2015) reports the results of an earnings belief intervention experiment, a field experiment similar to the lab

* Tel.: +1 6572783918.
E-mail address: nhuntington-klein@fullerton.edu

http://dx.doi.org/10.1016/j.econedurev.2016.03.008
0272-7757/© 2016 Elsevier Ltd. All rights reserved.

experiment of Wiswall and Zafar (2015). Student loan applicants are presented with expected earnings and cost information about available degrees. While the information has little effect on matriculation, low-socioeconomic status students in the treatment group are more likely to enroll in degrees with higher earnings net of costs.

HNRZ continues the research agenda, and provides one of the largest scale studies of student beliefs about college costs and benefits. HNRZ find that student beliefs are on average accurate for costs, but that beliefs are noisy and vary heavily over students. While cost expectations are on average accurate, earnings expectations are not. On average, students overestimate early-career earnings among the graduates of their first-choice institutions by 39.3%. They also find evidence that, while beliefs about earnings are not strongly linked to matriculation, they are related to major choice. Students with higher expectations of earnings are likely to choose majors that on average lead to lower earnings and graduation rates, and higher loan default rates. Students who overestimate costs are less likely to matriculate in any institution, including the program they declare as being their first choice, and are more likely to drop out.

Importantly, HNRZ stand out from the rest of the literature on student expectations by providing an informal theoretical framework for differences in student beliefs based on the costs of gathering information. Students who value financial outcomes less or who must pay a high cost to gather information are likely to gather less information.

Based on this model, HNRZ highlight five predictions, detailed in Section 3. Each prediction is an implication of differing student preferences for information or differing costs of information across groups. HNRZ evaluate these predictions using Chilean Proyecto 3E and administrative data and find support for them. As they note, there is a need for the model to be tested in alternate settings. In this paper I evaluate four of these predictions using the data set from Huntington-Klein (2015) (hereafter HK),[1] which comes from a 2012 survey of high school juniors and seniors in the United States and focuses on differing levels of education rather than major choice.

I find strong support for two of the predictions, partial support for another two, and also find differences in informational access across groups as expected.

2. Comparing HNRZ and HK

This paper aims to test the predictions of the HNRZ model using the data set from HK. The HK data set comes from a survey of 1,224 high school juniors and seniors at thirteen high schools in King County, Washington, which includes Seattle. The survey was administered using paper and pencil, and a study representative was always on hand to answer questions. The survey elicited student beliefs about earnings and the probability of employment conditional on the level of college education, as well as anticipated tuition and aid at four well-known Washington State colleges and university systems.[2] The data set also includes attitudes towards education and basic demographic characteristics. More information about survey design and administration, sample statistics, and response rates are in Huntington-Klein (2015).

The HK data provides an excellent opportunity to test the predictions of the HNRZ model. The HNRZ model is not designed to be specific to the Chilean context, but rather rests on basic principles of costly information. As such, we would expect the predictions to hold wherever students face information costs. The HK data can be used to test the model in a very different setting, while also maintaining several features that are similar to the HNRZ data, which makes HK a good candidate for comparison.

One useful feature of the HK data is that the sample consists of high school students who are not yet in college. Student beliefs in HK are prospective, and are taken from students who may have thought about or researched college but who have not yet attended. Like the college applicants in HNRZ, these students gather information through means other than actual college experience. So, the information gathering processes in both samples may plausibly be described by the same costly-information model.

Another convenient feature of the HK data is that it includes several variables that are present in the HNRZ data but not all data sets of student beliefs. HK and HNRZ both elicit student information sources, asking where they learned about college and careers. HK and HNRZ also ask students about their attitudes towards education and how important earnings are in making educational choices. Students are asked about their first-choice educational options, which allow both data sets to similarly handle costs and earnings, since beliefs may be more or less accurate depending on whether or not the student expects the information to be relevant to their future. Finally, both HK and HNRZ distinguish between a student's expected earnings for *themselves* as opposed to earnings for the *typical person*.[3]

Lastly, the HK data were collected in 2012, before the HNRZ theory or results were made public. The design of the HK survey cannot have been influenced by knowledge of the HNRZ data it is to be compared against.

Some features of the HK data make the comparison less straightforward. First, the HK data does not follow students through college, and so the HNRZ prediction that students who arrive at degree programs with inaccurate expectations should be less likely to graduate cannot be tested. Second, there may be fundamental differences in the way that students collect information about *levels* of college education as opposed to college major. If these differences are great enough, then the HNRZ predictions about earnings expectations would not be expected to hold, since HK elicits earnings expectations conditional on level of education, as opposed to major in HNRZ.

[1] Their fifth prediction, concerning college dropout, cannot be tested here because the HK data set does not follow students through college.

[2] The University of Washington, Washington State University, Western Washington University, and Seattle Community College.

[3] HNRZ also include earnings for the typical person in their gender and test score group.

Table 1
Errors in cost and earnings expectations.

	% Did not respond	Median error	IQR	Observations
(A) Cost expectation errors, first-choice institution				
All students	20.2	40.6	131.2	282
Low-SES	31.2	38.6	153.4	125
High-SES	10.7***	44.9	110.0***	149
(B) Cost expectation errors				
All students	24.1	70.4	193.3	1224
Low-SES	33.9	70.4	238.0	531
High-SES	14.7***	70.4	158.0***	650
(C) Earnings errors, typical graduate at first-choice education level				
All students	7.7	38.1	92.2	1199
Low-SES	10.5	23.5	96.8	514
High-SES	4.8***	47.1***	83.2***	643
(D) Earnings errors, own earnings at first-choice education level				
All students	6.3	56.8	113.3	1077
Low-SES	8.2	46.3	116.9	438
High-SES	4.8**	64.6***	104.4***	606

* indicates statistical difference between low-SES and high-SES students at the 10% level, determined by 5000 bootstrap iterations. Students with missing data were dropped from the relevant analysis. This table mimics Table 2 in HNRZ.

** indicates statistical difference between low-SES and high-SES students at the 5% level, determined by 5000 bootstrap iterations.

*** indicates statistical difference between low-SES and high-SES students at the 1% level, determined by 5000 bootstrap iterations.

3. Evaluating the HNRZ predictions

(1) Students who face higher search costs (e.g. students from low-SES, college-inexperienced families, neighborhoods, or schools) should have less accurate expectations about college characteristics.

I report errors in student expectations across socioeconomic status (SES) in the HK sample in Table 1. SES is measured by whether the student self-reports ever having received Free or Reduced Price Lunch (FRPL). Results are very similar if instead comparing students with above- or below-median self-reported GPA, or students with or without a parent with a bachelor's degree. HNRZ suggest low-SES students have higher information costs, an assertion examined directly in the *Information Sources* subsection below. In all cases, error is calculated as 100 × (Student report – benchmark)/benchmark.

Panels A and B compare student expectations of tuition, fees, and book costs at four local college and university systems against the actual in-state listed charges. Panel C compares earnings expectations for the "typical Washington State thirty-year old" graduate with their first-choice level of education against National Longitudinal Survey of Youth (NLSY) 1997 cohort respondents aged 29–31 in Washington State with that same level of education. Panel D compares expectations for one's own earnings with their first-choice level of education against NLSY respondents, matched by education, gender, and GPA quintile.

Students overestimate costs at their first-choice institution by 40.6% at the median, much higher than in HNRZ, who find on average no overestimation.

However, differences between students are consistent with HNRZ. Low-SES students do not have meaningfully higher errors at the median, but they are more likely to skip the tuition question, showing less confidence. Low-SES students also exhibit more dispersed beliefs, reported as the inter-quartile range (IQR) of the error distribution. Further, under the assumption that students who did not respond would have given less informed answers had they responded, the higher non-response rate for low-SES students may be biasing the IQR downwards relative to high-SES students. So the difference may actually be understated. The students who did not report a first-choice education level (and so were dropped from both panels C and D) were also more likely to be low-SES, further biasing the difference downwards. Panel A limits the sample only to those who chose one of the four colleges for which cost expectations are elicited as their first choice. This group is made of seniors and is about four percentage points less likely to be low-SES and has GPAs about .2 higher than the full sample, which includes both those not planning to go to college and those planning to go out of state to elite colleges. This subsample is not wildly different from the full sample, and so it is unlikely that there are major selection pressures on these results.

Students overestimate typical-graduate earnings, and their own earnings to a higher degree.[4] Again, low-SES students show greater spread in their responses, and the difference in spread may again be understated because of the higher nonresponse rate for low-SES students. As a caveat, it is not clear from these differences that the high-SES students are necessarily better-informed. The distribution is tighter for high-SES students, but this comes along with higher median levels of overestimation.

(2) Students who place relatively low value on earnings or costs in college choice should also have less accurate expectations about those attributes.

The HK survey asked about "good reasons for you to attend a college." 20.1% did not choose "Get a better paying job" as a good reason. 65.5% chose it as *a* good reason. A further 14.4% selected it as *the most important* reason. I compare earnings expectations across these three groups.

The HNRZ prediction is supported by the HK data: caring more about earnings is associated with fewer typical-graduate questions skipped (19.6% vs. 3.7% vs. 2.2%), tighter distributions (IQRs of 110.5 vs. 88.6 vs. 84.4), and non-monotonically with lower errors (medians of 37.1 vs. 12.2 vs. 20.9). Results are similar for own earnings, and the lowest category is in all cases significantly different from the other two at the 1% level.

(3) Students should have more knowledge of degree programs closer to their own interests.

I test predictions 3 and 4 using first-choice institution and education level rather than major. In HNRZ, prediction 3 was evaluated by looking at student claims about how much they know about a particular field of study outside of their first choice. They found support for the prediction. The HK data does not have a variable directly asking stu-

[4] Overestimation may be somewhat accounted for if, unprompted, students are incorporating inflation into expectations to account for the twelve years between NLSY earnings at the time of the survey and earnings when they themselves are thirty, but at an inflation rate of 2% this only addresses 26.8% of overestimation.

dents how much they know about other options, and so here I examine whether students hold more accurate beliefs about their first choice relative to other options. It is reasonable to expect that improved knowledge should be reflected in reduced errors, and this approach to testing the prediction may avoid some issues inherent in using self-reported knowledge levels.

Errors in cost estimates (see Table 1) are smaller for first-choice institutions than any institution. The IQR is also smaller, 131.2 for first choices vs. 193.3 for all four.

The same pattern does not hold for financial aid (Huntington-Klein & Blume, 2013) or for earnings expectations. For typical-graduate earnings expectations, the IQR is 92.2 for first-choice vs. 91.6 for any education level. For one's own earnings, the IQR is 113.3 vs. 112.4.

One possible explanation for this difference in results is that the level of variance in per-student aid and earnings for a particular U.S. institution is so high that there is less incentive to pay special attention to their first-choice option, but that costs are more predictable and so better described by the model. This is, of course, speculative.

Another interpretation of these results is that the prediction is only partially confirmed, and may not be robust to different methods of testing, given the different methods of testing this prediction here and in HNRZ.

(4) Students who have less accurate expectations about earnings or costs or who do not value these outcomes when making degree choices should be more likely to enroll in degree programs where past students have performed poorly along these dimensions.

For students who chose one of the seven Washington State four-year public college as their first choice, I regress the observed average quarterly earnings in the first year after graduation from the first-choice institution on quartile indicators for typical-graduate earnings errors, and on earnings emphasis as described for the test of prediction 2. Observed earnings come from Unemployment Insurance data, linked to administrative records at Washington State four-year public colleges by the Education Research and Data Center. Results are robust to the inclusion of gender, GPA, and FRPL controls, and to the use of Pell grant recipient earnings ten years after matriculation as reported in the College Scorecard.[5]

The first part of prediction 4, about students with less accurate expectations, is mildly supported. Students with lowest- or highest-quartile errors choose lower-earnings colleges than those in the middle two quartiles, (approximately \$300 difference between any of 1st/4th and 2nd/3rd, $p < .1$, $N = 253$). Since lowest-quartile errors are the most accurate, unusual errors, rather than large errors, are associated with low-earnings institutions.

The second part of the prediction, about students who do not value earnings or cost outcomes, is also mildly supported. Those with more emphasis on earnings attend higher-earnings colleges (\$105 and \$347 difference for *a* good reason and *most important* reason, respectively, compared to neither), but the difference is not statistically significant. With only 258 observations in this test, precision may be an issue.[6]

Table 2
Sources of information about college.

	All students	Low-SES	High-SES
Parent	70.5	55.9	83.5***
Other family	50.0	44.1	55.2***
Teacher	69.7	68.4	71.5
Other school staff	53.2	51.6	55.1
College rep.	45.0	46.0	45.1
Other adult	47.3	40.9	53.1***
Friend	58.1	50.3	64.9***
College fair/event	37.2	37.9	37.1
Visit to a college	51.1	44.6	56.9***
Printed materials	35.9	28.6	42.9***
Internet	52.9	41.6	63.1***
TV or movies	23.6	21.8	25.7
Observations	1224	531	650

* indicates statistical difference between low-SES and high-SES students at the 10% level, determined by 5000 bootstrap iterations. This table mimics Table 5 in HNRZ.
** indicates statistical difference between low-SES and high-SES students at the 5% level, determined by 5000 bootstrap iterations. This table mimics Table 5 in HNRZ.
*** indicates statistical difference between low-SES and high-SES students at the 1% level, determined by 5000 bootstrap iterations.

The weak confirmation may be due to differences in student preferences between Chile and the United States, or the subsample of United States students who attend in-state public university. The result may also be due to the relatively small number (7) of different institutions to choose between, as opposed to the large number of potential majors in HNRZ. With fewer options, there is less variance in earnings between different plausible options, and less room for option-specific preferences that are not related to finances to cancel out or be accounted for over the sample.

3.1. (Additional Prediction) information sources

Table 2 reports rate of use for sources of information about college. Results are similar when looking at sources of information about work.

In line with the HNRZ model, students with higher SES or higher GPAs make use of more sources of information than their peers. In general, differences are larger for background-specific sources such as family, friend group, access to the internet, and the opportunity to visit colleges. Differences are lower for school-based resources likely to be more equitably distributed across groups.

4. Discussion

Students face complex decisions about their education and are unlikely to have full information about the

[5] collegescorecard.ed.gov.

[6] Using College Scorecard data to expand the sample to include students who choose other colleges as their first option, these differences grow to \$993 and \$3317, respectively, and *most important* reason becomes significant ($p < .05$, $N = 568$). However, this analysis broadens the pool of institutions significantly (comparing community colleges to, for example, Harvard) and so selectivity of institutions and the ability to travel for college may explain a large portion of the difference. Smaller-sample results, focusing on a set of in-state colleges that students may realistically choose between, may be preferred. This expanded analysis does not resolve the odd results for the first part of prediction 4.

consequences of their actions. While empirical results concerning both the beliefs of students and barriers to information have been accumulating, there has been little useful theory.

In addition to the impressive Proyecto 3E data gathering effort, generating one of the largest data sets of student beliefs, HNRZ offer a theory of belief formation that is likely to be applicable in other contexts. This model provides a framing for the literature on student beliefs and allows the causal determinants of belief formation to be better studied and understood.

Theoretical development here is important given the general growing interest in information-based behavioral intervention. Informational interventions targeting low-SES students, who are emphasized in the theory since they face different information costs, include work by Hoxby and Turner (2013), who inform high-performing low-income students about their potential to attend and afford selective colleges, garnering a large behavioral response. Bettinger, Long, Oreopoulos, and Sanbonmatsu (2012) inform students and families about the availability of financial aid for college, although the information alone was ineffective unless paired with aid in filling out the financial aid application. Hastings et al. (2015), as previously mentioned, inform students about the costs and earnings associated with their educational options, finding that low-income students have the largest marginal response to the intervention. Resting these sorts of interventions (and the study of student choice more broadly) on a coherent theory of belief formation tells us why these interventions seem to work, lets us understand when they would not be expected to, and points towards how they can be improved.

In this paper I find general support for the predictions of the HNRZ model, both confirming that low-SES students appear to face higher information costs when learning about college, and supporting many of the implications of differing information costs between high- and low-SES students. The support found in the HK data for the HNRZ model is a recommendation for the model, given its success in predicting beliefs in vastly different settings (Washington State vs. Chile) and even for differing decision domains (level of education vs. choice of college major).

However, the support is not perfect. The prediction that students should have more information about degree options closer to their own interests is supported here for beliefs about costs, but not financial aid or earnings. This may come down to the differing contexts, in which the differences between institutions in financial aid and earnings in the United States are noisy enough that the incentives to search for information are not well represented by the model. There is also weak support for the prediction that students with less accurate expectations of earnings should be more likely to enroll in low-earnings educational options. Such a tendency may be more difficult to find in HK when considering the choice between a small number of institutions, as opposed to a large number of college majors. In each case, keeping these potential alternate explanations in mind, the final result is that while the replication largely confirms the original paper, it is not perfect and there may be reason to adjust the model in the face of these results. In the case of the prediction concerning matriculation choice, it may be worthwhile to adjust the model by looking into the layers of decision making between student beliefs and student choice, such as the influence of parents, as in Attanasio and Kaufmann (2014).

Just as the original study is not enough to confirm the model completely, one replication is not enough to decide these matters either. The HNRZ model is broad enough that there is potential to test it in many settings. Ideally, identical surveys conducted in multiple locations and across several different levels of educational decision-making would allow for a thorough test of the model's generalizability. A more pragmatic recommendation is that the attitude and information variables necessary to test the model could be easily included in other studies of student beliefs.

A more thorough understanding of the model has potentially huge value, and not only for the development of policy interventions. More broadly, the study of educational choice has long had to deal with an insufficient model of student beliefs. As HK finds, the common use of observational data as a proxy for student beliefs is not sufficient. HNRZ offer a better approach that can move the field forward.

Acknowledgments

Support was provided by Institute of Education Sciences Grant #R305B090012. I thank Mitchell Livy and anonymous reviewers for comments, and the Washington Education Research and Data Center for access to Washington State earnings data.

References

Attanasio, O. P., & Kaufmann, K. M. (2014). Education choices and returns to schooling: mothers' and youths' subjective expectations and their role by gender. *Journal of Development Economics, 109*, 203–216.

Bettinger, E., Long, B. T., Oreopoulos, P., & Sanbonmatsu, L. (2012). The role of simplification and information in college decisions: results from the H&R block FAFSA experiment. *The Quarterly Journal of Economics, 127*(3), 1205–1242.

Beyer, H., Hastings, J. S., Neilson, C. A., & Zimmerman, S. D. (2015). Connecting Student loans to labor market outcomes: policy lessons from Chile. *American Economic Review: Papers & Proceedings, 105*(5), 508–513.

Hastings, J. S., Neilson, C. A., Ramirez, A., & Zimmerman, S. D. (2016). (Un)informed college and major choice: evidence from linked survey and administrative data. *Economics of Education Review*. (forthcoming).

Hastings, J. S., Neilson, C. A., & Zimmerman, S. D. (2013). Are some degrees worth more than others? Evidence from college admission cutoffs in Chile. NBER Working Paper No. 19241.

Hastings, J. S., Neilson, C. A., & Zimmerman, S. D. (2015). The effects of earnings disclosure on college enrollment decisions. NBER Working Paper No. 21300.

Hoxby, C., & Turner, S. (2013). Expanding college opportunities for high-achieving, low income students. SIEPR Discussion Paper No. 12-014.

Huntington-Klein, N. (2015). Subjective and projected returns to education. *Journal of Economic Behavior & Organization, 117*, 10–25.

Huntington-Klein, N., & Blume, G. H. (2013). Are college-bound students aware of what's "Promised" to them? Evidence from Washington State. In *Proceedings of the 88th Annual Conference of Western Economic Association International, WEAI*.

Wiswall, M., & Zafar, B. (2015). Determinants of college major choice: identification using an information experiment. *Review of Economic Studies, 82*(2), 791–824.